Domain Specificity of Achievement Emotions

Dissertation
zur Erlangung des akademischen Grades
des Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Antonie Paula Magdalena Collier

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion

Fachbereich Psychologie

Tag der mündlichen Prüfung: 19.12.2011
1. Referent: Prof. Dr. Thomas Götz
2. Referentin: Prof. Dr. Anne Frenzel

Danksagung

Für die fachliche sowie persönliche Unterstützung beim Erstellen dieser Arbeit möchte ich mich an dieser Stelle herzlich bei mehreren Personen bedanken.

Zuallererst gilt mein Dank Prof. Dr. Thomas Götz für die Betreuung meiner Promotion, bei der ich mich immer auf die hilfreiche Unterstützung verlassen konnte. Unsere Gespräche waren für mich fachlich wie persönlich eine große Bereicherung. Ferner möchte ich mich auch herzlich bei meiner Zweitgutachterin Prof. Dr. Anne Frenzel bedanken sowie bei den Mitarbeitern des Lehrstuhls an der Universität Konstanz. Insbesondere danke ich Dr. Ulrike Nett für ihre (Gast-) Freudschaft sowie ihren fachlichen Rat, der mir immer neue Impulse und Antrieb für diese Arbeit gab. Dieser Dank gilt ebenso Dr. Birgit Wimmers und Elena Daschmann. Stellvertretend für die zahlreichen Studenten und Hilfskräfte, die an den Studien beteiligt waren, danke ich Sandra Benzing für die reibungslose und angenehme Teamarbeit bei der Planung und Durchführung der Studien.

Auch möchte ich mich bei meiner Familie und meinen Freunden, die mich in dieser Zeit unterstützt haben, für die Hilfe und den Zuspruch bedanken. And last but not least, I want to thank Selam for the precious support.

Table of Contents

Summary

The need of acquiring good learning skills has broadened in recent years well beyond its mere relevance to students or trainees. In fact in an ever-changing professional context, lifelong learning has become increasingly important. Similarly, for the acquisition of knowledge and the development of motivation, learning and achievement emotions are central. Indeed emotions have not only an impact on the learning process, but also on the psychosomatic health of learners and their personal development. This is to say that emotional experience has a meaning to the choice of study or professional career, and to lifelong learning.

It has been known for sometimes now that, having rather a domain-specific organization, emotional experience tend to have no cross-domain connections, at least in the school instruction context. However explanations on why emotions are such organized are rare. In consequence, with the broader aim of finding causes for the domain-specific experience of emotions, this dissertation will first try to find out which characteristics, or which categories of characteristics, better distinguish domains, or school subjects, from one another (Study I). These categories will be used to describe perceived differences and resemblances between school subjects that will subsequently be related to similarities and dissimilarities of students´ emotional experience in them (Study II).

It is to be noted additionally that studies on domain specificity of achievement emotions are mostly based on questionnaire measurements, which generate data on memory-based emotions depicting habitually experienced emotions. However it is probable that emotions are strongly dependent on the situational conditions. Therefore another aim of this work is, additionally to examining domain specificity of achievement emotions with the help of two different assessment methods, to compare these two methods. Hence, two separate methods of assessment were used on samples of students to gauge their emotional experience. Study II used a memory-based questionnaire to this purpose and assessed habitual emotional experience, as opposed to the real-time experienced emotions in classrooms that were assessed using an experience sampling method in Study III.

As a result, this dissertation presents three studies that approach the subject in hand using different angles, different assessment methods.

A first interview study (Study I) direct toward students enquired about the characteristics that they felt best described different school subjects, and built categories out of the obtained characteristics. Therefore 40 students, exactly half of whom were female, were probed, on which properties they felt most closely match their portrayal of specific school subjects. Out of their answers, 13 categories were developed with the help of inductive categorization, to which all obtained characteristics could be assigned.

The second study, a questionnaire study, developed scales for these categories with which one could capture the degree to which a specific category characterizes a specific school subject. These scales were assessed, together with emotions, with the help of a questionnaire to which 1709 students, 49.2% of whom were female, from Grade 8 and Grade 11 answered. This second study confirmed the assumed domain specificity of achievement emotions, as well as existence of even stronger domain specificity for the sub-sample composed of older students compared to the younger students. Moreover it has been shown that subjects intuitively perceived as being more similar were, on the one hand characterized more similarly, and on the other hand experienced emotionally more similarly, at least in students´ perception.

In the last study, 121 students, 50.4% of whom were female, all 8th and 11th graders were surveyed by means of the experience sampling method. As in the questionnaire study, the domain specificity of achievement emotions and the assumed age related differences have been confirmed by the experience sampling study. However it was not observed in this study that perceived similarities in school subjects translated into perceived similarities in their emotional experience. This in turn implied that there is a difference between the perceived emotions as measured by memory-based studies, and the emotional experience assessed in real-time during lessons as done by the experience sampling method. Memory-based enquiries are much more dependent on the students' image of the school subjects, whereas enquiries on emotions made during lessons are less influenced by personal opinions on the school subject, and more dependent on the current situational context.

Additionally the following results are noteworthy:

The built categorization system seems to be an adequate instrument to describe perceived differences and similarities of school subjects. Since, also these differences and similarities in characterizations seem to relate to the observed differences and similarities in their perceived emotional experience, as measured by the questionnaire study. Moreover, the domain-specific organization of achievement emotions was confirmed, as well as the fact that these emotions get even more differentiated with age. These results were observed for both the questionnaire method and the experience sampling method. The combination of the several assessment methods can, and should, be seen as a particular strength of this work, in addition to the very innovative character of the interview study. The results of the investigation enable a better understanding of why emotions are perceived domain-specific, and suggest that a particular attention must be given to the image building of school subjects. And that since it seems that the perception students have of a subject has a particular influence on how they experience it emotionally. Finally, implications that arise from these results concerning whether their implementation in practice or for further research will be discussed in this dissertation.

Zusammenfassung

Lernen ist nicht nur für Schüler, Auszubildende oder Studenten von Bedeutung, es wird auch im beruflichen Kontext immer wichtiger, lebenslang Wissen zu erwerben. Dabei sind Lern- und Leistungsemotionen grundlegend für die Motivationsbildung und maßgeblich für den Wissenserwerb. Sie haben nicht nur eine bedeutende Wirkung auf Lernprozesse, sondern auch Einfluss auf die psychosomatische Gesundheit von Lernern und ihre persönliche Entwicklung. Das emotionale Erleben ist zudem auch für die Wahl des Studienganges beziehungsweise die berufliche Karriere sowie für das lebenslange Lernen von Bedeutung.

Für den schulischen Unterricht ist seit einiger Zeit bekannt, dass emotionales Erleben nicht fächerübergreifend ist, sondern domänenspezifisch empfunden wird und somit domänenspezifisch erfasst werden muss. Allerdings sind Erklärungen, warum Leistungsemotionen domänenspezifisch organisiert sind, rar. Mit dem übergeordneten Ziel, mögliche Gründe für das domänenspezifische Erleben von Leistungsemotionen zu finden, wird in einem ersten Schritt versucht, Schulfächer zu charakterisieren beziehungsweise Kategorien zu finden, in welche diese Eigenschaften eingeteilt werden können (Studie I). Diese Kategorien sollen dazu dienen, Gemeinsamkeiten und Unterschiede von verschiedenen Schulfächern aufzudecken und diese in Verbindung mit den Ähnlichkeiten beziehungsweise Unähnlichkeiten der Schulfächer bezüglich ihres emotionalen Empfindens zu setzen (Studie II). Des Weiteren basieren Studien über die Domänenspezifität von Emotionen meist auf Fragebogenerhebungen, welche Antworten generieren, die auf Erinnerungen basiere, also die für gewöhnlich empfundenen Emotionen erfragen. Jedoch ist anzunehmen, dass Emotionen stark moment- und situationsabhängig sind. Deshalb ist ein weiteres Hauptziel dieser Dissertation, Domänenspezifität von Leistungsemotionen mit Hilfe von zwei verschiedenen Erhebungsmethoden zu untersuchen und diese schließlich zu vergleichen. Zum einen wurde eine erinnerungsbasierte Fragebogenmethode (Studie II) durchgeführt, zum anderen eine Experience-Sampling-Erhebung (Studie III), welche zeitnah emotionale Empfindungen direkt während des Unterrichts erfasst. In dieser Dissertation werden somit drei Studien vorgestellt, die sich auf unterschiedliche Weise und mit unterschiedlichen Erhebungsmethoden dem zu behandelnden Thema nähern.

Durch eine Interview-Studie (Studie I) wurden Eigenschaften erfragt, mit denen Schüler verschiedene Schulfächer beschreiben und für diese Charakteristika wurden 13 Kategorien gebildet. Es wurden 40 Schüler und Schülerinnen (50% weiblich) gefragt, welche Eigenschaften das jeweilige Fach speziell für sie kennzeichnen. Aus den Antworten wurden mittels induktiver Kategorienfindung 13 Kategorien gebildet, denen alle Beschreibungen zugeordnet werden konnten.

In der zweiten Studie, der Fragebogenstudie, wurden zu den in Studie I gebildeten Kategorien Skalen entwickelt, welche die Ausprägung der jeweiligen Eigenschaften erfassen. Diese Skalen wurden zusammen mit Emotionen im schulischen Kontext erhoben, zu diesem Zweck wurde der Fragebogen von 1709 Schülern und Schülerinnen aus der 8. und 11. Jahrgangstufe, von denen 49,2% weiblich waren, beantwortet. Die zweite Studie bestätigt die vermutete domänenspezifische Organisation von Leistungsemotionen sowie die sogar noch etwas stärkere Domänenspezifität der Emotionen in der Stichprobe der älteren Schüler im Vergleich zu den jüngeren Schülern. Des Weiteren konnte gezeigt werden, dass Fächer, welchen intuitiv eine größere Ähnlichkeit zugeschrieben wird, einerseits ähnlicher charakterisiert, und andererseits auch emotional ähnlicher erlebt werden.

In der letzten Studie wurden mit der Experience-Sampling-Methode 121 Schüler (50,4% weiblich), wiederum aus der 8. sowie 11. Klasse, untersucht. Die Domänenspezifität von Leistungsemotionen sowie die vermuteten altersbezogenen Unterschiede konnten, wie schon in der Fragebogenstudie, auch in der Experience-Sampling-Erhebung bestätigt werden. Allerdings konnte in dieser Studie nicht beobachtet werden, dass ähnliche Fächer auch ähnlich emotional wahrgenommen werden. Das heißt, es besteht ein Unterschied darin, ob Emotionen erinnerungsbasiert eingeschätzt werden oder ob das emotionale Erleben direkt in der Unterrichtsituation erfasst wird. Auf Erinnerungen basierte Urteile sind sehr viel abhängiger davon, welche Meinung Schüler von den speziellen Schulfächern haben, wohingegen die in der Unterrichtssituation erfassten Emotionen weniger von dem Bild über das Fach beeinflusst werden, sondern stärker von der tatsächlichen Situation abhängen.

Zudem können folgende weitere Schlussfolgerungen aus den Ergebnissen gezogen werden: Das gebildete Kategorien-System der Charakteristika scheint ein geeignetes Instru-

ment zu sein, um empfundene Unterschiede und Gemeinsamkeiten von Schulfächern zu beschreiben. Diese Eigenschaften stehen auch in Verbindung mit den in der Fragebogenstudie gemessenen Emotionen der Schüler in den verschiedenen Schulfächern. Darüber hinaus wurden die domänenspezifische Organisation von Leistungsemotionen und deren noch stärkere Ausdifferenzierung mit dem Alter bestätigt. Zu diesen Ergebnissen kamen sowohl die Fragebogen-Studie als auch die Experience-Sampling-Studie. Die Kombination der verschiedenen Erhebungsmethoden kann als eine der besonderen Stärken der Arbeit gesehen werden, ebenso wie der innovative Ansatz der Interview-Studie. Die Ergebnisse der Studie, ermöglicht es besser zu verstehen, warum Emotionen domänenspezifisch erlebt werden, und legen nahe, der Bedeutung von Einstellungen zu Unterrichtsfächern besondere Beachtung zu schenken. Denn es scheint, als wirke das Bild, das Schüler von einem Schulfach haben, entscheidend auf die in diesem Fach erlebten Emotionen ein. Abschließend werden in vorliegenden Dissertation Möglichkeiten für weitere Forschungsansätze sowie die sich ergebenden Implikationen für die praktische Umsetzung in der Schule diskutiert.

1 Introduction

1.1 Domain Specificity of Achievement Emotions

In educational-psychology research, the emotional experience related to learning and academic achievement has been largely neglected, and that for a long time. In addition, current research in this area suggests that emotions ought to have a significant role in pupil's psychosomatic health, in their personality development, as well as for their educational experience (e.g. Bong & Skaalvik, 2003; Efklides & Volet, 2005; Linnenbrink, 2006; Schutz & Pekrun, 2007). Furthermore achievement emotions have an influence not only on the choice of study, or rather on the choice of career, but also on matters concerning lifelong learning (e.g. Gieseke, 2007; Nagy, Köller, & Heckhausen, 2005). On the one hand, in many cases emotions in the learning and teaching context, have been analyzed on a cross-domain level (e.g. Zeidner, 1998). On the other hand, some investigations (e.g. Goetz, 2004; Goetz, Frenzel, Pekrun, Hall, & Lüdtke, 2007; Hembree, 1990; Marsh & Yeung, 1996) pointed to a domain-specific organization of learning and achievement emotions. This suggests that the interrelations of emotional experiences in different domains (e.g. school subjects) are relatively weak; it also means that the strongest connection between achievement and emotions can be found when this constructs are drawn within the same domain. Indeed, up until, now the causes of the domain-specific experience of emotions in the learning and teaching context are unclear. Obviously a distinction between the different subjects is done in regard to the differences in their teaching content. But it might be also interesting to see whether different school subjects can be characterized and differentiated according to certain other qualities, and to find out if similarities in these characterizations are related to similarities in the emotions experienced when studying these subjects.

Moreover, educational surveys measure students' emotions mostly by self-report. Though there are different types of self-reports, they can be distinguished into online reports (that means reports of current feeling) and memory-based reports (see i.a. Robinson & Clore, 2002). Most domain specificity studies about academic emotions are memory-based questionnaires which assess dispositions, or the so called trait emotions, rather than the current emotional situation, also called state emotions, which are meas-

ured from online reports. The concepts of trait emotions and state emotions were introduced to the studies of psychology of emotions by Spielberger (1972), who himself referred to the work of Cattell and Scheier (1961). The latter were the first to suggest the distinction between state and trait emotions, referring to anxiety. Accordingly, state emotions describe a temporary state, an emotional experience at a specific point in time, whereas trait emotions measure temporally stable, or rather habitual, emotions.

However it is likely that emotions in particular are experienced in a specific manner, depending on the moment and the situation. An alternative evaluation method which consider the above observation, is the experience sampling method (Csikszentmihalyi & Larson, 1987; Hektner, Schmidt, & Csikszentmihalyi, 2007). Another advantage of this method is that the reactive effects are minimized through the real-time data gathering and the natural environment.

Comparing these two assessment methods, namely questionnaire study and experience sampling method, is also interesting because there seem to be a big difference between what persons feel and what they perceive of what they feel. Noncurrent questionnaire studies asses more a habitual perceived emotion or an image of how the students think they feel. Hence the characterization of the school subjects might be closely connected to the beliefs students have about the domains or how they perceive them. So do beliefs about any given domain or subject influence beliefs about perceived achievement emotions in this specific domain? And could the emotions be measured more realistically and less dependently from the images of the subjects, with a more adequate assessment method?

The aim of this thesis is to fill in these gaps and to gain a more comprehensive understanding of students' emotional experience. Therefore, first an interview study (Study I) was conducted, and from it characteristics of school subjects were established in order to find reasons that could explain the domain-specific organization of academic emotions. Then out of these characterizations, scales were set up. Afterwards, these scales were confronted to the emotions experienced during different school subjects, and their relations observed in a memory-based questionnaire study (Study II). A further experience

sampling study (Study III) served to examine the differences between these two evaluation methods.

1.1.1 Definition

The more one thinks about domain specificity of achievement emotions, the harder it gets to define. For all intents and purposes, a good definition should encompass all aspects of the subject in study. Taking this into account, the search for a broader definition comes invariably down to one simple observation: There is domain specificity of achievement emotions if emotions are experienced in a different manner for different school subjects.

Obviously this statement requires an illustrating example; let's say that if every student experiences the same level of joy in mathematics as in another subject like in history, then it could be deduced that this achievement emotion does not possess a domain-specific organization. However this is a very simple case-example, and does not contain all possible means of establishing the existence of domain specificity of achievement emotions; more will be said on the ways of proving domain-specific organization of emotions below. Therefore, studying domain specificity of achievement emotions comes as a question on whether emotional experience differs from domain to domain, or from subject to subject (e.g. school subjects), and on studying where these differences lie, if they exist.

One simple idea to investigate where these differences reside could be to compare the mean values of the single emotions of students in-between the different domains. Another way to look at it could be to observe the correlations of the single emotions in-between domains, the latter approach having the advantage of being more specific. Indeed, when trying to relate sets of numbers or values, there are two common ways of handling it. First and easiest would be to deduce the mean of each set of values and to compare the obtained means; A second more exhaustive way is to compare each specific member of one set of values to a corresponding member of a different set of values.

Another more important advantage of the latter method is that it averts two kinds of errors that in general describe the same thing. The first error could be in assuming domain specificity even if there is no actual domain-specific organization. Let's say for example that dissimilarity in mean values is established. Based on this single information one could

conclude that there is domain specificity. But this result cannot be a guarantee for low correlations. Now, let´s assume that actual calculated correlations turn out to be high in value suggesting non-domain-specific organization of the emotions. This in consequence will render our initial conclusion not valid. Indeed, if the values of specific emotions in one subject are linearly linked to the corresponding specific emotions in another subject for every student, not only values of computed correlation will be high, but one can also conclude that there is no domain specificity of emotions. The second error, which mirrors the first, would be to assume no domain specificity based on the only fact that mean values are similar, when there is actually domain specificity. This would happen if similarity in mean values is observed, while actual computed correlations are low, driving one to conclude that there is domain specificity. One simple explanation could be that the distribution of single values of emotions around the mean values are greatly different, hinting at substantial differences in variance even if means are supposed to be equal. That in turn could forbid high values of correlations between corresponding single values of two different domains.

In both cases – meaning whether one looks at the mean values or the correlations - the focus is on a comparison in-between the different domains, or subjects. This comparison follows as an answer to the central question of whether there exists any relation between one specific emotion in one domain, to the same specific emotion in another domain, or to put it simply, a between-domain relation.

Though there are different ways to define domain specificity of emotional experience, a simple but fundamental classification would be the “between-domain” and the “within-domain” differentiation. Both approaches are illustrated in the following Figures 1 and 2.

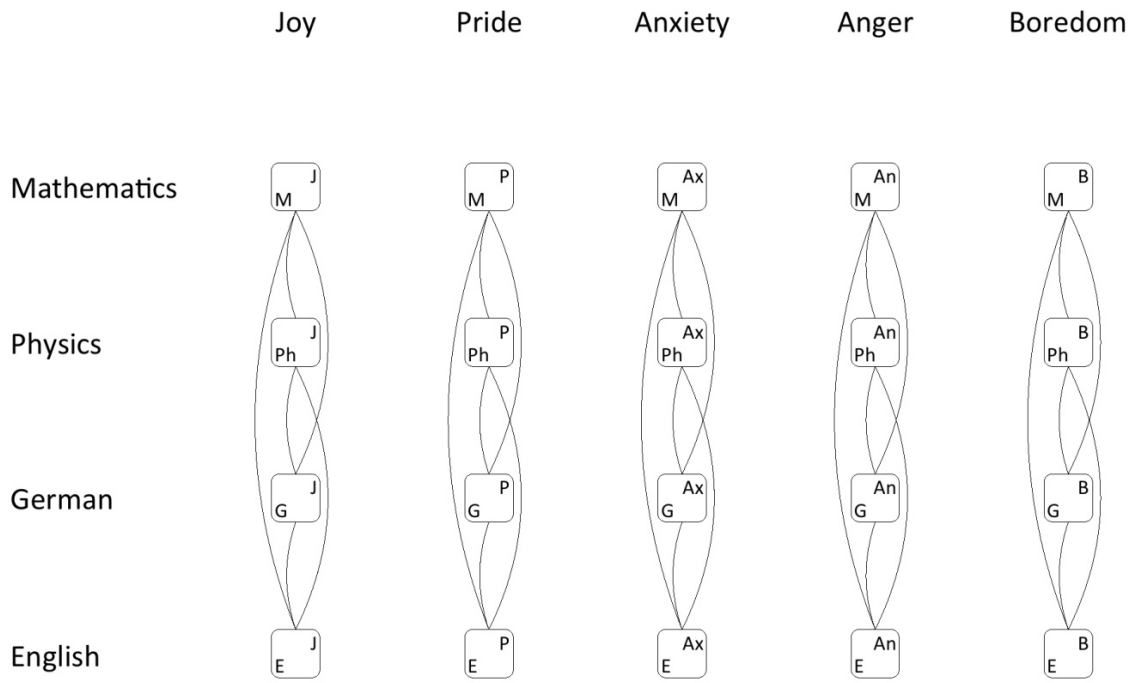


Figure 1: Between-Domain Relations of Achievement Emotions.

Note. Correlations are represented by the lines; Enjoyment (Ej), Pride (Pr), Anxiety (Ax), Anger (An), Boredom (B); Mathematics (M), Physics (P), German (G) and English (E).

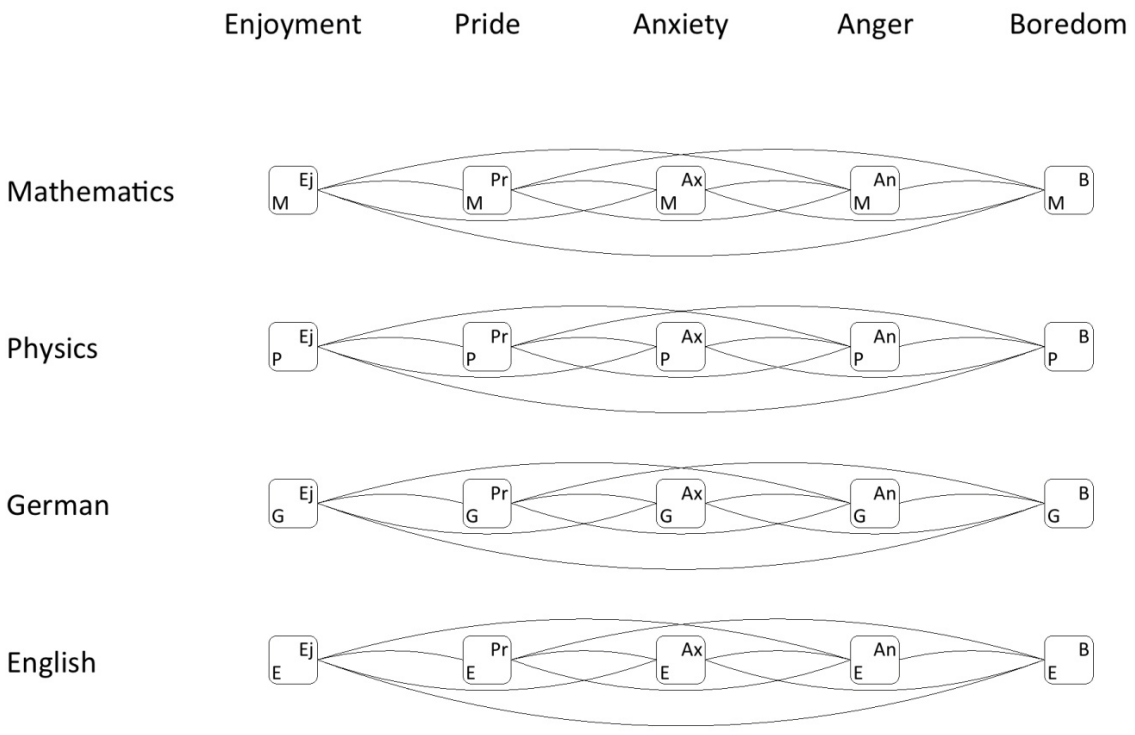


Figure 2: Within-Domain Relations of Achievement Emotions.

Note. Correlations are represented by the lines; Enjoyment (Ej), Pride (Pr), Anxiety (Ax), Anger (An), Boredom (B); Mathematics (M), Physics (P), German (G) and English (E).

Looking at domain specificity in the light of between-domain relations – as described above – seemed to be more straightforward compared to observations made from the within-domain relations point of view. This directness comes from the fact that the between-domain relations manner observes directly the between-domain differences of singled out emotions. The first figure describes this comparison process.

The latter manner however questions if the domains differ in their structure, or rather raises the question of the nature of the relations of the emotions within a subject, and compares the structural dynamic of the different domains as illustrated in Figure 2.

There exists however two methods to answer this within-domain relations question. On the one hand, studying the relations between the different emotions in one specific domain could be of interest, on the other hand, the study of correlations between emotions and other constructs (e.g. grades or self-concept) could be interesting.

For this thesis, observing domain specificity of achievement emotions is equated to observing the between-domain relations of emotions, and more specifically to observing the correlations of the singled out emotions in-between different domains.

A sure sign of domain specificity of achievement emotions is, first of all, relatively weak between-domain correlations. Another indication would be stronger correlations of singled out emotions between similar domains, or subjects, compared to correlations of the singled out emotions between dissimilar domains.

Finally, once that domain specificity of achievement emotions is established, one must insure that the assessments in these studies were conducted in a domain-specific manner. This in turn is shown through a high convergent and divergent validity of domain-specific assessed emotions. Looking at the relations between domain-specific emotions and their corresponding domain-specific achievement scores, high values of correlations support the convergent validity of the domain-specific emotion construct. It is also equally expected that the relations between emotions and achievement in similar domains are substantially large and that the relations between these variables in dissimilar subject are quite small, supporting the divergent validity.

1.1.2 Causes for Domain Specificity

The findings made about domain specificity of academic emotions are described in a more detailed manner in the following subchapter. But these findings raise the question of what their actual causes are. A lot can be said on the whys and wherefores.

First of all, it can be said that the social environment – e.g. family, pears, teachers or cultural background – as well as the age and the sex, have an influence on the emotions students experience in different school subjects and on their levels of interest in different domains. Likewise the classroom setting, the composition of the class (e.g. if the class stays the same or if the classmates change across subjects) and the course instructors can be assumed to moderate the strength of between-domain relations of academic emotions. Some of these aspects have already been examined, for example the impact of classmates (e.g. Frenzel, Pekrun, & Goetz, 2007) or the influence of the instructor (e.g. Goetz, Frenzel, Lüdtke, & Hall, 2011). The latter explored the moderating influence of having the same versus a different course instructor in mathematics and physics on the between-domain relations of the emotions enjoyment, pride, anxiety, anger and boredom. In line with other investigations the between-domain relations were found to be relatively weak, but these relations were stronger in classes which had the same instructor in both subjects, compared to classes which had different instructors for each subject.

This thesis concentrates on another possible reason for domain specificity of students' emotional experience, the domain by itself, and specifically the characteristics of the domains. There are only few studies in which characteristics of subjects are considered to explain differences in perceptions of subjects. Bachmair (1969) analyzed students' and teachers' attitudes towards the school subjects mathematics, German, English, geography, music and sports. The students assessed the domains among other things with regard to popularity, interest, and importance – the importance of the topics in general and the importance for life in general – but also to the experienced anxiety and the required level of effort needed to succeed.

The low popularity of science subjects was investigated by Merzyn (2008) with the help of student surveys, teachers' comments, and experts' judgments. In this publication Merzyn attributes the low popularity of mathematics, physics, and biology to several aspects:

Teaching methods, content of lessons, standing of the subject, perceived difficulty, and amount of subject matters.

Another study which assessed characteristics of science subjects is one made by Jenkins and Nelson (2005), where 1277 students, most of them age 14 to 15, were questioned on their views about science school subjects. Among other things the students had to rate the difficulty, the interest, the everyday usefulness, and the popularity of school subjects.

The image of mathematics and science was also analyzed by Kessel and Hannover (2007), as well as the influence of the analyzed image on the development of academic interest. In addition, in another publication, Kessel and Hannover (2006) defined three image factors, namely the perceived difficulty of the subject, the perceived gender connotation, and how strong the students assess the opportunity for self-realization in the subject.

Stodolsky and Grossman (1995) examined if the way teachers perceive their subjects, and the way they teach and plan their lessons, are related. They based their study on the observation of a number of teachers schooling in different subjects, whereby they concluded that: “five features of subject matter, and teachers’ perceptions of how these features apply to their own subject, seem especially salient. These features are: degree of definition, scope, or the number of distinct fields included in the school subject; degree of sequence; characterization of subject as static or dynamic; and the required or elective statue of the subject” (Stodolsky & Grossman, 1995, p.229).

In a survey by Jäger-Flor and Jäger (2008), a sample of individuals consisting mainly of students and teachers was asked about how they assess different aspects of the subject mathematics. Therefore the questions addressed the frequency of illustrations, everyday usefulness, difficulty, interestingness, logic, importance of the subject and also if mathematics is exciting or just a duty, as well as if the subject was considered for everyone or only experts. The authors observe in particular mathematics, even though they also assessed the popularity and the importance of a set of 21 school subjects using a 5-point Likert scale.

In their article, Stevens, Wineburg, Herrenkohl, and Bell (2005) provide an overview of the scientific paradigm in educational psychology, and more specifically in questions con-

cerning how the school subjects are related or rather differentiated. In their opinion there is a lack of studies on the similarities and dissimilarities of school subjects.

After discovering the domain-specific character of learning behavior most investigations examine only single domains, whereby the relations between the different domains cannot be detected. A classification of school subjects' attributes would be interesting, with regard to the genesis of domain specificity of achievement emotions. Such a characterization may also serve to find reasons for the largely domain-specific nature of students' emotional experience, though suggested by several studies. In particular, since there are only few investigations about the reasons for domain-specific organization of achievement emotions, especially ones that cover a number of different school subjects.

1.1.3 Strength

Current studies (e.g. Bong, 2001; Marsh & Yeung, 1996; Möller & Köller, 2001) show that constructs, for example the academic self-concept and the self-efficacy expectancies and also the achievement purposes and task-orientated valence, are not organized in a cross-domain manner. Appraisal-theories of emotions (e.g. Scherer, Schorr, & Johnstone, 2001) assume that expectations, self-concepts, personal values and achievement goals play an important role in the development of emotions. Pekrun's control-value-theory (2000, 2006) states that subjectively experienced control as well as valence have an influence on the development of learning and achievement emotions. Hence, one conclusion of the theory is that the connections of achievement emotions are weak between different subjects, because the constructs control and valence are domain-specific. Therefore it must be concluded that achievement emotions are also domain-specific. The domain specificity of achievement emotions has already been proved in several studies:
Marsh and Yeung (1996) analyzed data from the "National Educational Longitudinal Study of 1988", where they assessed prospective joy, perceived naturalness and anxiety in mathematics, natural science subjects, social studies and English. They have determined a domain-specific organization of emotions for children in the 8th grade. The meta-analysis of Seipp and Schwarzer (1991), in which the relationship between anxiety and performance was analyzed, also indicates the domain specificity of emotions. In another meta-analysis conducted by Hembree (1990) anxiety in mathematics (Mathematics Anxiety

Rating Scale/MARS; Richardson & Suinn, 1972) was correlated with the verbal and mathematical skills. Findings suggest a negative correlation between the anxiety scores (MARS) and the mathematical skills, but no correlation between anxiety and the verbal abilities. These observations are an indication that anxiety is not organized across domains. Pekrun, Goetz, Titz, & Perry (2002) investigated enjoyment in mathematics, languages, music and sports and found no correlations of the emotions between the different subjects. Also for the emotion of anxiety only low correlations were observed. The study by Goetz (2004) analyzed enjoyment, anger, anxiety, boredom and hopelessness in mathematics, German, English, music and sports. Goetz concludes, in line with recent findings, that the results indicate unequivocally a domain-specific organization of learning and achievement emotions. Moreover Goetz, Frenzel, Pekrun and Hall (2006), Goetz, Pekrun, Hall and Haag (2006) as well as Goetz, Cronjäger, Frenzel, Lüdtke, and Hall (2010) got to similar results. In addition, the investigation by Götz et al. (2007) which examined the between- and within-domain relations of students' academic emotions, having analyzed enjoyment, pride, anxiety, anger and boredom emotions in mathematics, physics, German and English experienced by students, draw analog conclusions.

1.1.4 Age-related Differences

In a further analysis, questions dealing with the existence of differences between different age groups with regard to domain specificity of achievement emotions could be of interest.

Concerning the within-domain relations of age-related differences, there are few, and not all consistent, observations made in the scientific literature. As to the comparison of different age groups, a study by Bong (2001) found stability in the within-domain relations for motivational constructs across both samples. Similar results were presented by Goetz et al. (2007), where the authors suggest that the within-domain relations of emotional experience are relatively stable over the different age groups. In addition, a study by Goetz et al. (2010) observed the within-domain relations between self-concept and achievement emotions on a sample out of two different class levels (Grade 8 and Grade 11). However, in contrast to the two previously mentioned investigations, the latter showed no stable strength of relations between the self-concept and the emotion con-

structs across the different age groups; or more precisely these relations were stronger for the sample of older students, compared to the sample of younger students.

There are also only a small number of studies concerning age-related differences in between-domain relations of academic emotions. In motivation research, the study by Bong (2001) indicates that motivational constructs are more differentiated in older students compared to younger ones. Bong investigated motivational construct in two different age groups (7^{th} and 11^{th} grade) and showed even stronger domain specificity in the older sample. Furthermore theory and previous research results, mention an increase in specificity, between childhood and adolescence, in students' interest related topics (see Krapp, 2002).

In the research area concerning students' self concept, Marsh and Ayotte (2003) analyzed data from students in different grade levels and observed that concerning self-concept the between-domain relations are also stronger for younger than for older students. In their differential distinctiveness hypothesis they stated that constructs which are weakly associated across domains in younger children, become more differentiated with age, this implies that the older group shows stronger between-domain differences. What's more, this hypothesis also suggests that for constructs which have strong between-domain relations as observed in younger children, decreases less, stay the same or even increase when observed in older students.

Finally, it can be said that there are not much studies that investigate age differences in achievement emotions. In the aforementioned study by Goetz et al. (2007) two age groups were examined (Grade 8 and 11) and even stronger domain specificity of the emotional experience was observed for the older students. The results of this investigation are consistent with Marsh and Ayotte's conclusions in their differential distinctiveness hypothesis transposed on emotions; meaning that the weak between-domain relations of emotions, pertaining to quantitative and verbal subjects from the 8^{th} grade, got even weaker or even negative in the 11^{th} grade. Additionally age-related decline was observed in the already relatively weak relations between the two language domains, German and English. In contrary Goetz et al. (2007) observed that the relatively strong correlations of

emotions between mathematics and physics in the younger sample, stayed as strong in the older sample, or got even stronger.

1.2 Research Goals

Former studies amply demonstrated the domain-specific organization of achievement emotions, and this dissertation in essence keeps up this tradition as it goes along this line of questioning. Consequently, the three empirical studies contained in this dissertation investigate the domain specificity of achievement emotions. Nevertheless what sets apart this particular work is that beyond just a discussion on the findings of empirical studies, it tries to uncover early solutions to the question of causes for domain specificity that earlier works have often overlooked.

Therefore, bearing in mind that the bigger aim of this dissertation is a foray into the causes for domain specificity which are as yet unclear, the first step towards that goal should be to outline attributes of domains, or school subjects. Once a list of describing attributes is established categories of these could be extracted. This first stride is made through an interview study (Study I) which develops these categories of characteristics.

Subsequently, out of these categories, scales are developed for studying properties of characteristics. The motivation behind this particular step is to discover differences and similarities of domains. In parallel, differences and similarities in the emotional experience of the different school subjects are measured, after which the relations of the categories and the emotional experience are looked upon. This basically describes the questionnaire study (Study II), to which a study on the domain-specific character of emotions can be added.

Another goal of this work is to compare different measurement methods. Thus, in a final stroke of pen, in addition to the questionnaire study, Study III analyzes the domain specificity of achievement emotions by means of the experience sampling method. This allows the comparison of, on the one hand memory-based reports, and on the other hand online reports; both concerning the investigation of domain specificity of achievement emotions.

To put it in a nutshell, the current dissertation answers in essence the following major research questions:

- Do students experience achievement emotions in a domain-specific way?
- Are there any age-related differences with regard to the domain specificity?
- Which attributes describe specific school subjects, and which categories can be found to summarize these characteristics?
- How are these categories related to the emotional experience, in the different domains?
- Do differences in measurement methods – particularly a questionnaire design and an experience sampling method – lead to the same results?

2 Study I – Interview Study: Characteristics of Domains

2.1 Research Aim and Question

In studying any given subject, the central part should be the explanation of the topic in hand. This in turn has the advantage of outlining the borders of the study and keeps it in the scope of the original problematic, but more importantly of insuring the exactness of the ensuing conclusions. In addition, as it was already stated in the introduction, the aim of this dissertation is foray into the causes for domain specificity of achievement emotions. To answer to this problematic one must have a clear view of the nature of domains and their characteristics.

For the sake of exactitude, this explanation must establish a differentiation between domains, sketching out borders separating them. Nonetheless there is more than one way to differentiate one domain from another or, in this case, one school subject from another. Though the most elementary differentiation one can make is by the topics dealt with in each subject, this dissertation looks at other ways of discriminating between school subjects. Are there other attributes by which these domains can be described and distinguished? Or more importantly, what are the attributes used by students to describe school subjects?

Thus, in an effort to define characteristics, to describe specific school subjects, the present study asks the following question, for each of the particular domains:

Which properties characterize this specific school subject in particular?

Common sense dictating that a number of characteristic can be taken into account to describe a specific school subjects, the aim of this first study will be to construct a categorization system of characteristics that can be later on used to describe specific domains.

2.2 Method

2.2.1 Participants and Data Collection

As stated above the main objective of this study is to find the characteristics by which students describe best school subjects. Hence it comes to mind that the easiest would be to question the students themselves, and thus an interview study was put in place.

Therefore students were recruited from four different German high schools in the area of the city of Konstanz. The average age of the 40 interviewed students was 15.79 (*SD* = 1.63), half of them being from Grade 8 (mean age = 14.24 years, *SD* = .33) and the other half from Grade 11 (mean age = 17.34 years, *SD* = .47), 50% of the interviewees being female, in each case. Interviews were conducted by trained testing personal, using half-structured interview questionnaires, and took place from February to March 2009. Participation was voluntary, and data collection was anonymous.

Approximately ten minutes were given for each interviewee to answer to a series of questions. Interviewers put a question to each interviewed student, saying: “Which properties characterize the school subject … in particular, in your opinion?” This question refers to a set of seven subjects, or namely mathematics, physics, German, English, biology, history and music. If students had not understood the exact wording of the question, or talked instead about their feelings, respective to the school subject, or discussed topics treated in it, the interviewer led back to the original question focusing on the characteristics of the subject.

2.2.2 Data Analysis

Out of the collected data a wide variety of characteristics were generated. The sheer number of characteristics described by students in their interviews, forces one to organize and classify them into categories, for the sake of readability and to avoid redundancy. The conformity of the classifications developed being central, the path of this study must go through testing it. One last step in the analysis should be the description of the collected data using these newly developed categories.

Development of Categories of Domain Characteristics

The development of categories breaks down to one simple idea, the quantification of qualitative data. And it is for this specific intent, that the procedure of qualitative content analysis (Mayring, 2003) was used. Indeed, this method helped in structuring, in evaluating the transcribed interviews, and in setting up a categorization system generated by inductive categorization.

One simple way to describe this procedure would be as follows: After determining the level of abstraction, answers to a first interview conducted were classified into categories. Inductively the next interviews were successively divided into the existing categories derived from previous interviews. In the case where answers to the interviews did not fit into the existing categories, the category system was extended to accommodate them. After five interviews the next step was tested deductively, either the formed categories fitted to the rest of the material as a whole or the system of categories was optimized in several induction-deduction-loops.

Interrater Reliability

Eventually categorization playing an important role in the general explanation of domains, ensuring the conformity of these classifications to the observed data plays a great part in warranting the accuracy of the conclusions drawn from the study.

The adopted way of measuring conformity of these classifications here is to deduce the degree of agreement among raters. Thus, bearing in mind that a high degree of agreement between raters is a sign of a good classification, the interrater reliability was calculated out of the collected data. Though different statistics are appropriate for different types of measurement the Fleiss' Kappa is the option adopted here. The Fleiss' Kappa in essence measures how well the observed agreement degree among raters does compared to what one might expect if all raters made their ratings randomly. And this is calculated as a ratio described here:

$$\kappa = \frac{p_0 - p_c}{1 - p_c}.$$

Where the numerator indicates the degree of agreement actually achieved above chance and the denominator gives the degree of agreement that is attainable above chance, and where p_0 and p_c indicate the actual and the randomly expected value of agreement among raters respectively. Landis and Koch (1977) proposed that a Kappa-value of $\kappa > .60$ is considered acceptable.

For the present study taking two raters into account, the Fleiss' Kappa is as high as .85, a reasonably high value.

Frequency Analysis

After ensuring the conformity of the categories to the actual data, the focus must be on describing this data using the newly formed categorization system. As a consequence, a frequency analysis of the named categories, themselves resulting from the interviews, was conducted. Describing shortly the procedure it can be said that, first if one same student made several statements with identical content these statements were counted as one single remark for the according category. Second, statements were not weighted accordingly to quantifiers used by the interviewee (e.g. very, little…). Thus all statements were treated identically. Ultimately this procedure helps in deducing how often specific categories were mentioned, an interesting result to look at.

2.3 Results

Now, as introduced in previous sections, the purpose of this study is the precise explication of domains; and it was discussed how discriminating between domains helped in this task. It was also stated how attributes used by students to describe school subjects, support this differentiation process.

Yet again as observed previously, an interview study was conducted in order to find out characteristics, or rather categories of characteristics, taken into account by students to describe school subjects. Presenting results of this interview study below, it was decided that after the description of the developed categories of characteristics, the collected data will be illustrated through a frequencies description. This frequency analysis encompass the complete set of data comprising of two ages groups and seven school subjects designated previously (see part 2.2.1). This analysis was conducted in a specific and sepa-

rated manner for the two different age groups, but also as a whole, both based on, the complete set of seven subjects, and in a separate and specific manner on a set of four selected subjects, namely mathematics, physics, German and English.

Hence the procedure of inductive categorization, which mechanisms were described above, was implemented to analyze the content of the interviews. Following the optimization of the categorization system in several induction-deduction-loops, thirteen categories presented in Table 1 below, were selected.

Table 1: *Categories of Characteristics of Domains*

1	amount of subject matter
2	variety in lesson content
3	importance of the grade
4	subject matter importance
5	effort
6	everyday usefulness
7	topic relation
8	level of difficulty
9	class discussion
10	perceived learning requirements
11	illustration
12	right solution
13	current topics

Though most of the categories named in Table 1 above are reasonably self explanatory, some need more clarifications. It is rather obvious that the described categories refer to only a handful of aspects, like topics, importance and so on.

The assessment of importance for example is split into two main categories, the importance of getting a good grade as assessed by the students, on the one hand, which is described in third position in the table, and the importance of the subject regardless of grades, as described in fourth position in the table, on the other hand. Even thought category six describing the everyday usefulness can also be considered as measuring importance.

The assessment of topics dealt within each subject is treated in different categories. Specifically, categories one and two concentrate on describing the number of subject matter dealt within each school subject, and their variety, respectively. As categories seven and thirteen try to illustrate each, the continuity of the topics presented in school subjects, or rather if successive topics are related or unrelated to each other, and the relation of topics presented in school subjects to current events, respectively. This four attributes are classified under the titles “amount of subject matter,” “variety in lesson content,” “topic relation” and “current topics” respectively.

Moreover, one other important subject assessed by the interviews was the teaching method or the design of lessons. Namely if lessons were prone to class discussion, prompting opinion exchange, as listed in category nine under “class discussion,” or as formalized by category eleven under “illustration” how high the frequency of illustrations in lessons is. More peripherally maybe, the fact that the questions or problems associated with specific school subject are perceived as having, one unique solution or many acceptable solutions, was of course of interest and looked upon in category twelve under “right solution”.

Additional aspects assessed by this study were the effort put into learning a subject by the student, the level of difficulty, and finally the perceived learning requirement or if the student perceives that learning a specific subject requires a natural talent or merely to acquire a particular learning behavior. These categories were recorded under “effort,” “level of difficulty,” and “perceived learning requirements” in lines five, eight and ten of Table 1 above. It is to bear in mind that these categories describe school subject as they were perceived by the students.

Subsequently a high value of Fleiss’ Kappa ensuring the high conformity of this classification to the data, a frequency analysis was conducted in an effort to illustrate the data set using these new found categories. For practical reasons the presentation of the results of the frequency analysis is separated in two main groups, mainly frequency analysis calculated on the complete set of students as a whole on the one hand, and in a separate and specific manner for the two different age groups on the other.

These two parts in turn were separated into two sub-groups each, one dealing with frequencies analyzed on the entire set of seven school subjects and another on a set of four school subjects resulting in a total of four sub-groups. Now this set of four subjects, namely mathematics, physics, German and English was selected mainly because these subjects were objectively perceived as being fundamental and widely studied, making results easily comparable to previous works; But more fundamentally maybe because these are the four subjects that will be considered later on in this dissertation.

Paramount to understanding the frequency analysis is the knowledge of its nature and of what it actually tries to describe. The obvious answer is that it can be described as a ratio calculated for each category looking at what proportion of the students actually specifically mentioned the corresponding category to describe school subjects. The corresponding ratios to each category for the four different sub-groups are summarized in figures 3 to 6. For example, Figure 3 shows the percentage of the responses describing the properties of the seven domains that specifically used the corresponding category to make the description. These percentages were calculated of course on the entire sample of students as a whole.

The observations of the figures point out a number of interesting facts. While Figure 3 reveals that proportion of specific mentions of categories varies strongly from category to category compared to Figure 4, these disparities between mention rates of different categories show little variation. Most frequently mentioned categories were found to be “level of difficulty” and “illustration”, with about 37% of the students’ referring to at least one of these categories to characterize at least one school subject. And in contrast less frequently mentioned categories were “importance of the grade”, “topic relation”, “right solution”, and “current topics” with less than a 3% score for each as far as their mention rate is concerned. And the remaining seven categories assigned a rate level comprised between 4 and 10%.

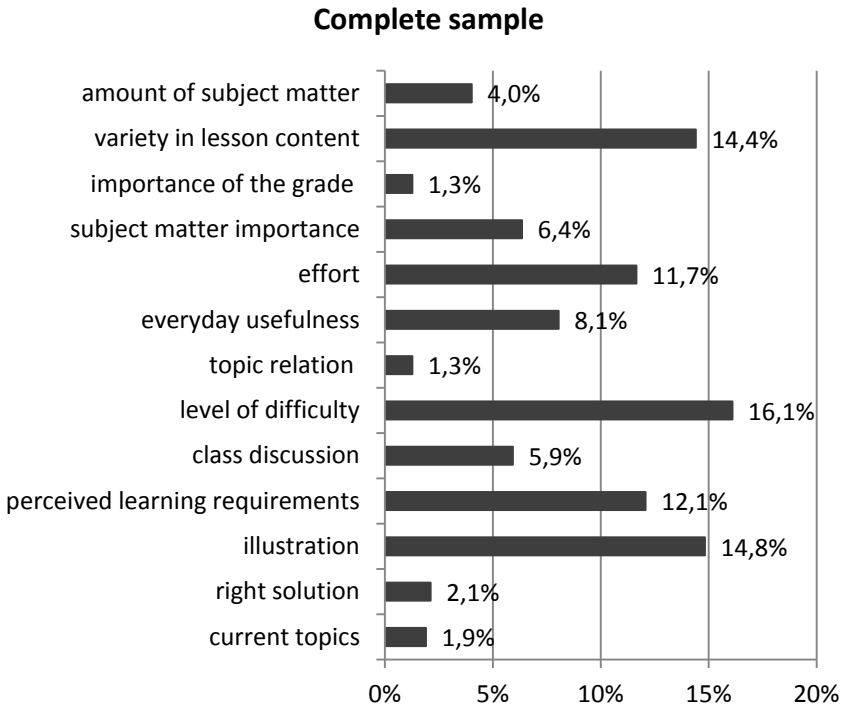


Figure 3: Frequency of Mentions of Categories Calculated on the Overall Seven Domains.

Note. In total 472 mentions for mathematics, physics, German, English, biology, history and music were recorded – multiple responses were possible.

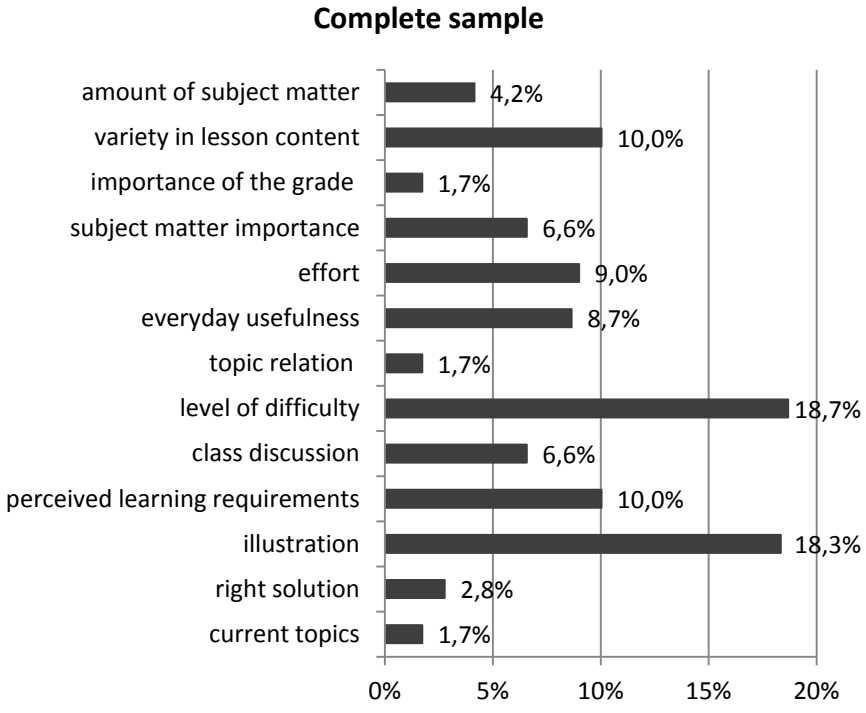


Figure 4: Frequency of Mentions of Categories Calculated on a Set of Four Selected Domains.

Note. In total 289 mentions for mathematics, physics, German and English were recorded – multiple responses were possible.

Figure 4 shows the percentage of the responses describing the properties of the selected domains, namely mathematics, physics, German, and English, that specifically used the corresponding category to make the description. The percentages were calculated on the entire sample of students as a whole. One notable difference here compared to Figure 3 is that the category labeled “variety in lesson content” scores also high bringing the proportion of students mentioning this category or at least “level of difficulty” or “illustration” specifically to describe at least one school subject to a total of 45%.

In the first part data were analyzed as a whole, in the next part the analysis will be made in a discriminatory manner; the discriminating factor being the grade. Figures 5 and 6 show the percentage of the responses describing the properties of all the domains, and the percentage of the responses describing the properties of the restricted domains respectively that specifically used the corresponding category to make the description. The percentages were calculated this time on the entire sample of students but in a differentiated manner according to the school grade of the student.

In addition, in Figure 5 and Figure 6 it is easy to observe the differences between the two class levels and some results are quite remarkable. For example, younger students – in comparison to the older students – referred more often to the variety in lessons or to the level of difficulty, but also more often specifically mentioned how strenuous lessons can be, or how high the frequency of illustration in the content of lessons is. In contrast, 8th graders did not think about the fact that for a specific school subject questions or problems can be considered as having a unique or a number of acceptable solutions, neither about the continuity of the successive topics studied in each subject, nor did they deal with the possibility of introducing current events in their daily lessons. In these three categories almost all mentions were made by 11th grade students, although in total not many answers took these aspects into account. Another interesting fact worth noting is that the sample composed of older students deals much more frequently with the importance of a subject and with the transferability of the content on everyday life.

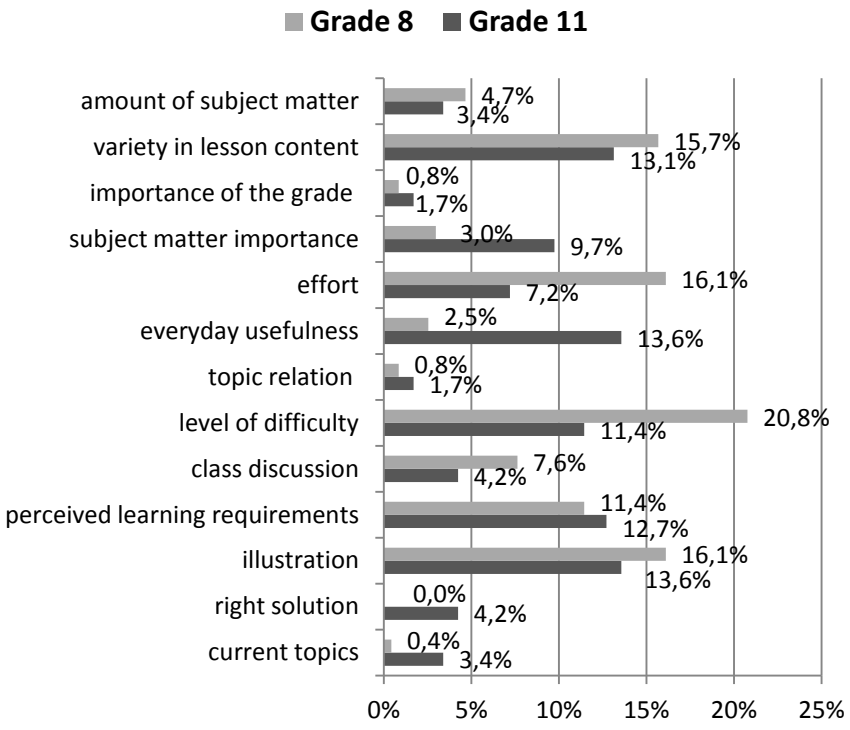


Figure 5: Frequency of Mentions of Categories Calculated on Overall Seven Domains Divided into Different Age Groups.

Note. In total 472 mentions for mathematics, physics, German, English; biology, history and music were recorded – multiple responses were possible.

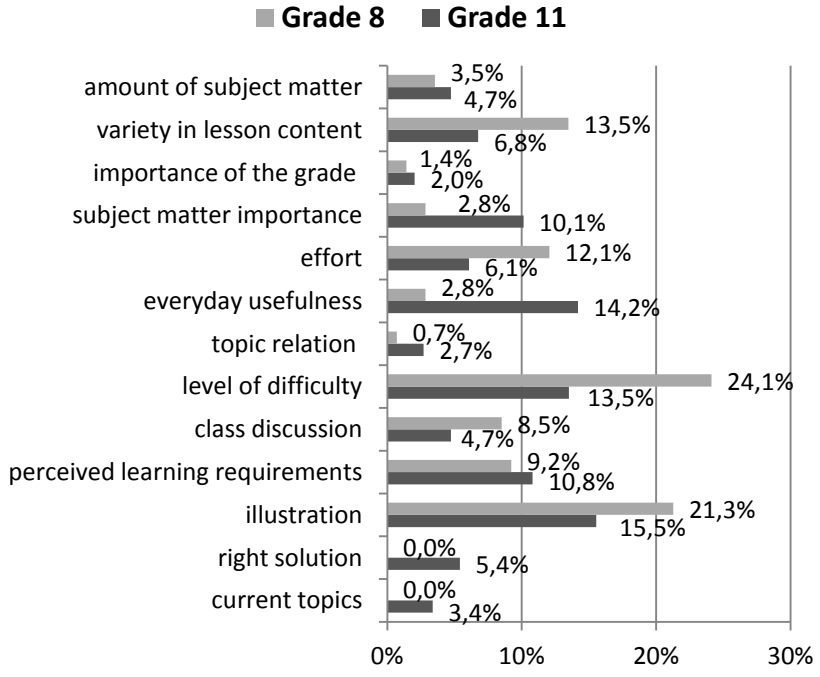


Figure 6: Frequency of Mentions of Categories Calculated on a Set of Four Selected Domains Divided into Different Age Groups.

Note. In total 289 mentions for mathematics, physics, German and English were recorded – multiple responses were possible.

One can easily discern that no specific and separate descriptions of results were done, each stemming from Figure 5 and 6. But these same two figures show that percentage of the responses describing the properties of all the domains and the percentage of the responses describing the properties of the restricted domains are, at least structurally, the same. Thus the above description was made to encompass observations from both figures. It can be said additionally that the appendix (see Table 14) showing the exact numbers of the interview responses and how they are assigned to the categories for each of the four domains (mathematics, physics, German and English) and the two class levels, one can always refer to it for more details.

2.4 Discussion

The aim of this first interview study was to develop categories to describe the characteristics of school subjects. It should also be underlined that this investigation is an explorative study, a first approach aiming at gaining an overview on the subject in hand or on the characterization of domains. Its other aim was to prepare the following questionnaire study (Study II).

Getting down to the how this study was done, the study asked what characteristics specifically describe the school subjects; therefore the questions aimed particularly at assessing the image of the subjects, or the perception students have of school subjects.

Though it looks like the named categories are numerous, it is rather obvious that the described categories can be classified into a handful of aspect orientated “super-categories”, or namely into topics related, importance related, teaching method or lesson design related, and lastly accessibility related “super-categories”. Categories one, two, seven and thirteen, going to the first group, categories three, four and six in the second, and nine, eleven and twelve in the third group, and lastly categories five, eight and ten going into the accessibility related group.

However, the developed categorization make sense as parts of these “super-categories” are rarely mentioned, like categories related to “importance of the grade”, “topic relation”, “right solution” and “current topics”, whereas some are frequently mentioned.

Conversely all members of the last “super-category”, or the accessibility related categories all, score high in their mention rate. Undoubtedly, age difference plays a role here. It is quite obvious that accessibility related issues demand, maybe less maturity to grasp, compared to topic related issues. Since it stands to reason that school subject topic relation to current events for example demands a certain amount of maturity of the student, for him or her to be able to follow and understand his or her closer or farther environment. The same explanations can be advanced in explaining differences observed between the two classes of ages in issues related to “everyday usefulness” or to uniqueness of solution in questions or problems raised in a particular subject, and also in explaining why the continuity of the successive topics studied in each subject was not taken into account. It is indeed easy to observe that for these particular categories most, if not all contributions were made by 11th graders. More difficult maybe to explain is that some categories were favored more by the 8th graders and more mentioned by them than by their older counterparts, like the “level of difficulty”, the “illustration” and most curiously the “variety in lesson content.” Beyond the simplistic conclusion that one can draw from these observations saying 8th graders are more interested in variety in lessons content in order to avoid boredom and the basic generalization that 8th graders are more concerned with day to day flow of lessons and less by practicality problematic than 11th graders; It is quite apparent that 8th graders are more interested by categories related to lessons content and requirement. And this because, in all probability, they are lacking the skills to, project themselves or their skills, acquired through school studies, into the future or into their environment. Thus incapacity to project oneself in space and time could be a start of an explanation.

It is astonishing that some classifications, which are examined in other studies, did not come up in the interviews and therefore have not been developed as a categories, as for example the popularity of the domains or the interest for the subject (compare Bachmair, 1969; Jenkins & Nelson, 2005). A possible explanation could be that the current study explicitly asked for characterizations or how the subjects are perceived, and that this in turn excluded as a consequence direct judgment of the domains as in popularity or personal interest. Indeed it is probable that the wording of the question influenced the students to be more objective, living them with less room to emit personal judgment, but

also probably broadening their scope of search of characteristics to what they see as more objective matters. The form of the question also excluded other possible interesting categories as the influence of teachers. Even though it is possible that a general image of teachers in a specific subject as being a certain way can be considered as an attribute for that subject, it is likely that students discarded it because they perceive it as emitting personal judgment. Or they perceived it as a particular characteristic of a particular course in a particular school subject, rather than a characteristic of the subject itself. Other aspects have not been mentioned in this interview study, but have been found important in other investigations, as for example the perceived gender connotation and how great the students assess the opportunity for self-realization in the subject (see Kessels & Hannover, 2006, 2007). These aspects may be suitable for a characterization of school subjects but have not developed out of the interviews of this study. One simple reason to this can be that students are not aware of this type of characterization and thus did not take them into account.

Though, looking at the rate of mentions of the named categories, and comparing the ones calculated over the selected four subjects to those calculated on the overall seven subjects, there are not many differences evidently between them. It is indeed important to appreciate the fact that when only the four selected subjects were taken into account the two most mentioned categories were “level of difficulty” and “illustration”, and that where the overall subjects are considered the category “variety in lesson content” can be added in the list of most mentioned categories. This hints at the fact that minor subjects, like history, biology, and music are more described by characteristics related to variety in lessons compared to the other subjects.

On the other hand, although these categories were developed out of a study on seven school subjects, the fact that there is very little difference observed in frequency of mention of the categories as calculated on both, the seven overall subjects or on the specific four selected school subjects, suggests that the categories listed are a fairly good tool to characterize domains as they are not specifically dependent on the considered school subjects. In addition, the strength of the categorization system is demonstrated by a high value of kappa of about .85, sending the interrater reliability through the roof. This means

that the classification of characteristics used to describe school subjects by the students into these thirteen categories has a high agreement among the raters. Another strong point of this study is that it has broadened the scope of investigation not only in increasing the number of school subjects taken into account, but also in widening the latitude giving to students in their search of characteristics. As Study I was an explorative study and given that previous studies bound themselves mostly at analyzing one or a limited number of school subjects or characteristics, the present work was bound to be innovative, and demanded an innovative approach. Most importantly maybe it gave an extensive overview of the subject in hand.

However, due to this very explorative character of the study also shown through the lack of exactly matching theoretical background, categories only treat a limited number of aspects of characteristics as describe above. Only two age groups have been taken from only one type of school (German Gymnasium). Thus, although it is rather preferable to broaden the scope of investigation to other age groups and to other school types, there are only few limitations to the present study.

Hence, this explorative research should be furthered in subsequent studies, for example by use of detailed interview studies or by electronic diary studies, to look at the origins of subjective beliefs about school subjects. A fundamental practical implication one can deduce from this study is that teachers should know what students think about their subjects and how they describe them. And teacher and parents should be aware of the fact that good teaching does not only imply to communicate knowledge but also to propagate a positive image of the subject to better the perception student have of the school subject.

3 Study II – Questionnaire Study: Domain Specificity of Achievement Emotions and Characteristics of Domains

3.1 Research Aims and Hypotheses

As mentioned above, the main objective of the first study was to establish characteristics or rather categories of characteristics by which students describe domains, or school subjects. Once these categories listed, this dissertation aims at discriminating between domains using the same list of categories. Since the mere mention of a particular category to describe a specific domain is not enough to show how strong the former characterizes the latter, a way of quantifying qualitative data had to be set up. In order to do this, the second study developed scales out of the categories and used them to measure the extent to which the mentioned categories describe specific school subjects. Study I, out of which categories were, and scales will be developed was based on a rather limited sample of students. The assessment of the scales conducted in the second study was in turn based on a wider sample. Additionally the emotional experience of students in different school subjects was collected. This study is in a way a replication of former studies but with the very original characteristic that it links achievement emotions and characterizations of school subjects.

Hence, it could be said that the first aim of this study is to confirm the conclusions of previous educational psychology researches, in which students' emotional experiences were shown to be largely domain-specific (e.g. Goetz et al., 2007) and to further analyze age-related differences of the domain specificity. But most fundamentally maybe the second goal was the further investigation of the relations of the categories, developed in Study I, and students' emotional experience in different school subjects. In fact, the motivation behind this particular study, throughout which a number of hypotheses listed below were examined, is to discover differences and similarities of domains. In parallel, differences and similarities in the emotional experience of the different school subjects are measured, after which the relation of the categories and the emotional experience are looked upon.

***Hypothesis 1**: Domain Specificity of Achievement Emotions*

It is assumed that the relations of single emotions – enjoyment, pride, anxiety, anger and boredom – across domains mathematics, physics, German and English would be rather weak. Further it is expected that the correlations between subjects, that are intuitively considered to be similar (like mathematics and physics, or German and English), are stronger than across subjects that are intuitively considered to be dissimilar, and that the strongest relationship should be observed between emotions and achievement scores related to the same domain.

***Hypothesis 2**: Domain Specificity Increases with Age*

It is also hypothesized that more domain specificity should be observed in the group of 11^{th} graders compared to the 8^{th} graders; this means that the between-domain relations of academic emotions should be stronger for younger students compared to older.

Referring to Marsh and Ayotte's (2003) differential distinctiveness hypothesis, it is also hypothesized that the between-domain relations should decrease with age for emotional experience in disparate domains and there should be less or no decrease in related subjects.

***Hypothesis 3**: Similarities in School Characteristics are related to Similarities in Emotional Experience*

Finally, it is also assumed that subjects intuitively perceived as more similar also share more similarities in their characterizations, thus implying that the relations of the characteristics between these similar subjects would be stronger than across dissimilar domains. In addition, it is expected that if the category profiles of two subjects are similar, then the emotional experience in these subjects is also similar.

3.2 Method

3.2.1 Participants and Data Collection

Now as announced earlier the aim of the first part of this study is to develop scales out of the first study, to measure the extent to which categories describe specific domains and

to further assess these developed scales. This will be done through a questionnaire study based on a wide sample of students.

Thus, data for the second study were collected through 11 high schools and 74 different classes, 35 of which were 8th grade classes, and the rest being 11th grade classes, in the area of Konstanz and Bayreuth, in Germany. In total 1709 students with an average age of 15.95 years (*SD* = 1.64) were probed, 49.2% of whom were female. This sample included 855 students from Grade 8 whose mean age is 14.42 years (*SD* = .55), 49.2% of whom were female, and 854 students from Grade 11 whose mean age is about 17.51 years (*SD* = .56), with a proportion of female as high as 49.2%. All participants took part to this survey on a voluntary basis, and were tested in their classroom's settings. The data were collected between the months of May and June 2009 and via self-report instruments, and all responses were anonymous. These polls were conducted by trained testing personnel, using fully standardized student questionnaires.

3.2.2 Measures

Characteristic of Domains

Hence, for each category developed out of the Interview Study scales were put up in the following manner. One specific question was formulated for each of the thirteen categories, with answer's format designed as a 5-point Likert scales for each four domains. The exact wordings of each of these questions were arranged accordingly to the category considered. Table 2 down below summarizes questions examined and the range of each possible answer for each of the thirteen categories.

Table 2: *Scales for All Categories*

	Category	Question Item	Range of Scale 1	5
1	amount of subject matter	Would you say that the amount of subject matter dealt with in your lessons is, rather little, just sufficient or rather a lot?	little	a lot
2	variety in lesson content	How would you say you perceive the variety in the content of your lessons, rather as little diverse or diverse?	low	high
3	importance of the grade	How important would you say is, for you to get a good grade in this subject?	not important	important
4	subject matter importance	How important would you say is this subject for you, independently of the grade?	not important	important
5	effort	How strenuous would you say lessons in this subject are, as you perceive them?	little	a lot
6	everyday usefulness	How related would you say, topics dealt with in your lessons, are to your needs in your everyday life?	little	a lot
7	topic relation/continuity	How would you say, successive topics treated in lessons are, unrelated or related to each other?	unrelated	related
8	level of difficulty	How would you say you perceive the difficulty in this subject, rather low or rather high?	low	high
9	class discussion	Would you say that there is little or lot of exchange of views in this subject?	little	much
10	perceived learning requirements	What would you say is necessary to obtain good results in this subject, diligence or rather natural talent?	learned behavior	natural talent
11	illustration/explanation	Would you say there are little or much illustrations in this subject?	low	high
12	right solution	Would you say right answers to specific questions in this subject are, rather numerous or unequivocally unique?	many	one
13	current topics	How would you say is the amount of current events discussed in this subject, little or much?	little	much

Single Emotions in Different Domains

Well established scales were used to measure the emotions enjoyment, anger, pride, anxiety and boredom in the following domains; mathematics, physics, German and English. The scales were adapted from the mathematic version of the Achievement Emotions Questionnaire [AEQ-M] by Pekrun, Goetz, and Frenzel (2005). The exact wording of the question associated to each of the emotion scales was of the type: "How strongly do you experience … *(the corresponding emotion*) … in the following subjects?" The response format consisted of a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*very strongly*). Questions were submitted for every school subject.

Academic Achievement

Achievement in mathematics, physics, German and English was assessed by students' midterm grades. German grades range from 1 (*very good*) to 6 (*not sufficient*). Here also intermediate grade, as for example "1-2" which describe grades between 1 and 2, were taken into account. However these grades were inverted, high numbers indicating better performance, so that values could be more easily interpreted.

3.2.3 Data Analysis

Thus, the nature of the data being multilevel, or more specifically it has a two-level structure in which students (Level 1; N = 1709) are nested within classes (Level 2; N = 74), a number of statistics were generated out of it, namely single and average multilevel correlations. It is to be noticed that the analyses presented down below were conducted via multilevel statistics using Mplus 5.2 software (Muthén & Muthén, 1998-2008).

Average Correlations

In order to create a reliable similarity index, which could describe the strength of between-domain relations across overall emotions, or conversely overall categories, the average (Ø-) correlations between any two domains were calculated. Therefore the absolute values of the between-domain relations were Fisher-z-transformed, and then the arithmetic means over every emotion, or respectively category, was calculated and afterwards retransformed and z-standardized.

Comparison of the Average Correlations

In order to compare correlations of the categories between any two subjects in one hand, and of the emotions between these two same subjects on the other – correlations in all cases being average correlations – a statistical test of significance was put in place.
In fact, this comparison comes down to a test of significance of the difference between two correlations drawn from two different samples. Indeed the samples considered are assumed to be mutually independent as categories and emotions are two different constructs. The test is a straightforward one, that studies if the arithmetic difference observed between the *z* Fischer transformations of the two correlations in absolute value, accounting of course for the variance of this same difference, or to be precise the standard deviation, follows a standard normal distribution. Or more mathematically, if we assume that z_1 and z_2 describe the *z* transformation of correlations of categories between any two subjects, and of emotions between the same two subjects respectively, noting n_1 and n_2 the size of the sample from which these two correlations are calculated respectively, let $D = z_1 - z_2$, in this case the variance of this difference noted

$$\text{Var (D)} = \frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}.$$

The previous statement observes that,

$$U = \frac{|D|}{\sqrt{\text{Var (D)}}} \sim \mathcal{N}(\mu=0, \sigma^2=1),$$

in which case it is concluded that there is no significant difference between the two correlations observed (see Chen & Popovich, 2002).

Homogeneity Index h(d)

A fully unconditional two-level model was created, for which information was provided, concerning the outcome variability on both of the levels (see Figure 5; Raudenbush & Bryk, 2002). It is postulated that the relative sizes of proportions of these variances allow inferences about the degree of homogeneity (see Goetz et al., 2007; Hox, 2002) and consequently, about the strength of between-domain relations of each emotion. Therefore, two-level models were created separately for the 8th graders and 11th graders samples, in order to compare between-domain relations across the two age groups.

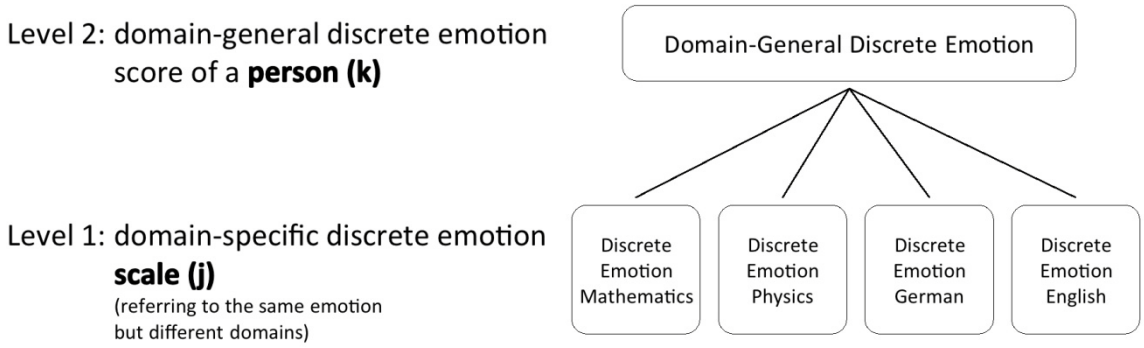


Figure 7: Multilevel Structure of Between-Domain Relations of Achievement Emotions.

The variance of the outcome variable Y_{jk} of this two-level hierarchical model is defined as variance of the domain-specific discrete emotion scale *j* (Level 1) nested within the domain-general discrete emotion score of person *k* (Level 3). That means, for each of the five discrete emotions, each person had one domain-general emotion score (Level 2 units) and four domain-specific emotion scores (Level 1 unit).

The variance of Y_{jk} can be partitioned into two parts:

$$Var(Y_{jk}) = Var(u_{ok} + r_{jk}) = \tau_{\beta} + \tau_{\pi} ,$$

with u_{ok} being the Level 2 residuum and r_{jk} being the Level 1 residuum. Now the index h(d) of homogeneity of the scales referring to the same emotion but different domains, can be calculated on the basis of these variance components:

$$h(d) = \tau_{\beta} / (\tau_{\beta} + \tau_{\pi}),$$

in which d stands for domain. That is, h(d) reflects the proportion of the Level 2 variance in the total variance. Hence, h(d) is an index for the homogeneity of Level 1 constructs. To be brief, h(d) indicates the strength of between-domain relations of the domain-specific single emotion scales. The higher the h(d) value is, the stronger the between-domain relations for the single emotions are.

3.3 Results

3.3.1 Descriptive Statistics

It has already been mentioned above that the two main goals of this study are, on the one hand to develop scales out of the categories and to use them to measure the extent to which mentioned categories describe specific school subjects, thus discriminating between domains. On the other hand, to look at if findings in this study actually support the hypotheses stated above. Hence data were collected for these specific reasons and the assessment of the described scales. It has also been observed, that the nature of the data collected being multilevel, multilevel statistics analysis of them was conducted and a number of statistics were generated out of them, which are going to be described below.

Means and Standard Deviations

The first groups of statistics described by this dissertation are the mean values of the range of scales, over all students, for each category and each school subject and their corresponding standard deviations. Statistics concerning 8^{th} and 11^{th} graders were calculated in a separate manner, following which they were summarized in Table 3 below.
A number of interesting facts can be observed in this table, but most important to note are:

- That for both age group's mean values for categories pertaining to the "amount of subject matter" are higher for mathematics and physics compared to English or German, though this contrast is particularly higher for older students.
- The mean values for categories related to the "variety of lesson" or more precisely to variety in lessons content are lowest for mathematics for both age groups, though these same mean values for physics are substantially lower for older students compared to younger students as they decrease from 2.92 to 2.5.
- In quite the same way, the mean values for categories related to the "importance of grades" are highest for mathematics and English with values of about 4.4, and lowest for physics with approximate values of 3.7, for both age groups.

- Similarly, mean values for categories associated to the “importance of subjects” independently from grades for physics are smallest and much higher in English compared to other domains, for both age groups.
- The mean values for categories pertaining to “effort” students perceive they must provide in order to succeed in particular subjects are remarkably higher for the quantitative domains, namely mathematics and physics, compared to verbal domains, namely English and German, and for both age groups.
- In opposition, mean values for categories related to the “everyday usefulness” of particular subjects are much lower in the quantitative domains, compared to the verbal domains, for both age groups.
- Also mean values for categories, associated with questions concerning the comprehensibility of relations between successive lessons in particular subjects are equally high for all domains, except German, for younger students compared to older students. In this respect, there is a clear contrast between mean values of quantitative domains, which are much higher and mean values of verbal domains which are quite low.
- The mean values for categories related to “level of difficulty”, and the abundance for “class discussions”, for quantitative domains are respectively higher and lower, compared to verbal domains, for both age groups, with the contrast being more marked among older students, compared to younger students.
- Also mean values for categories describing the “perceived learning requirements”, and conversely the perceive amount of “illustrations” in specific subjects have the same approximate value of 3.0, for all overall cases. Noticeable exception are, for English among younger students which is lower and for German among older students which is higher for the first category, as well as for physics which is much higher for both age groups for the second category.
- In addition mean values for categories pertaining to uniqueness of right solutions in specific school subjects, as described previously and dubbed “right solution”, are much higher for quantitative domains compared to verbal domains for 11^{th} graders. For 8^{th} graders mean values for all domains are equally high around 3.5, except for German for which it is as lower (3.0).

- The mean values for categories related to the possibilities of introducing “current topics” into lessons for quantitative domains are much lower compared to verbal domains for both age groups, though this contrast yet again is more marked for older students.

These results (Table 3) were also illustrated using graphics included further in the appendix in Figure 13 and 14, which will not be analyzed here to avoid redundancy.

Table 3: *Descriptive Statistics of the Categories*

		Grade 8 (n=855)		Grade 11 (n=854)	
		M	SD	M	SD
amount of subject matter	Mathematics	3.51	1.02	3.64	0.98
	Physics	3.53	0.97	3.70	0.92
	German	2.88	1.11	2.56	1.05
	English	3.14	0.84	2.66	0.84
variety in lesson content	Mathematics	2.29	1.36	2.01	1.01
	Physics	2.92	1.58	2.50	1.29
	German	2.97	1.45	2.93	1.37
	English	3.08	1.38	3.14	1.35
importance of the grade	Mathematics	4.40	0.76	4.23	0.93
	Physics	3.69	1.03	3.75	1.14
	German	4.10	0.93	3.92	1.09
	English	4.41	0.65	4.27	0.78
subject matter importance	Mathematics	3.43	1.55	3.41	1.81
	Physics	2.76	1.57	2.84	1.80
	German	3.40	1.36	3.35	1.59
	English	4.08	1.14	4.21	0.93
effort	Mathematics	3.67	1.42	3.76	1.45
	Physics	3.40	1.36	3.67	1.25
	German	3.00	1.12	2.84	1.30
	English	3.11	1.06	2.76	1.15
everyday usefulness	Mathematics	2.40	1.31	1.96	1.04
	Physics	2.55	1.43	2.80	1.39
	German	3.35	1.56	3.00	1.53
	English	3.59	1.35	3.75	1.06
topic relation	Mathematics	3.48	1.37	3.96	0.98
	Physics	3.06	1.17	3.47	1.11
	German	2.77	1.12	2.70	1.10
	English	3.30	1.18	2.90	1.16
level of difficulty	Mathematics	3.70	1.54	3.93	1.47
	Physics	3.54	1.45	3.81	1.26
	German	2.94	1.11	2.88	1.24
	English	3.06	1.13	2.81	1.20
class discussion	Mathematics	2.44	1.50	1.99	1.27
	Physics	2.70	1.38	2.21	1.29
	German	3.46	1.35	3.86	1.24
	English	3.02	1.29	3.45	1.07
perceived learning requirements	Mathematics	3.10	1.67	3.11	1.53
	Physics	2.95	1.59	3.06	1.42
	German	2.99	1.28	3.47	1.15
	English	2.52	1.19	2.83	1.19
illustration	Mathematics	2.82	1.46	2.93	1.42
	Physics	3.80	1.20	3.82	1.12
	German	2.72	1.07	2.58	1.01
	English	2.84	1.09	2.69	1.00
right solution	Mathematics	3.66	1.88	4.07	1.59
	Physics	3.48	1.52	4.05	1.41
	German	3.01	1.37	2.33	1.32
	English	3.54	1.08	3.11	1.11
current topics	Mathematics	2.03	1.19	1.58	0.80
	Physics	2.60	1.45	2.13	1.22
	German	3.18	1.40	3.09	1.47
	English	3.06	1.40	3.70	1.04

Note. The response format of the single scales consisted of a 5-point Likert scale; the ranges can be seen in Table 2; multi-level structure was considered.

The next groups of statistics described by this dissertation will be the mean values of the range of scales over all students for each emotion and each school subject and their corresponding standard deviations. Statistics concerning 8th graders and 11th graders were calculated in a separate manner, following which they were summarized in the Table 4 below.

Table 4: *Descriptive Statistics of Academic Emotion Scales and Achievement Scores (Questionnaire Study)*

		Grade 8 (n=855)		Grade 11 (n=854)	
		M	SD	M	SD
Enjoyment					
	Mathematics	2.44	1.71	2.52	1.70
	Physics	2.44	1.69	2.37	1.57
	German	2.84	1.46	2.86	1.53
	English	3.08	1.53	3.15	1.39
Pride					
	Mathematics	2.97	1.67	3.09	1.79
	Physics	2.84	1.59	2.85	1.66
	German	3.00	1.26	3.08	1.39
	English	3.23	1.22	3.24	1.34
Anxiety					
	Mathematics	2.76	2.16	2.95	2.18
	Physics	2.59	2.05	2.90	2.09
	German	2.20	1.53	2.17	1.39
	English	2.22	1.47	2.16	1.39
Anger					
	Mathematics	3.08	1.85	3.30	1.84
	Physics	2.89	1.76	3.19	1.71
	German	2.58	1.45	2.71	1.41
	English	2.52	1.43	2.56	1.37
Boredom					
	Mathematics	3.33	1.74	3.21	1.61
	Physics	3.46	1.75	3.32	1.66
	German	3.20	1.55	3.24	1.54
	English	2.93	1.57	2.92	1.45
Achievement					
	Mathematics	4.21	1.07	3.98	1.42
	Physics	4.18	1.16	4.09	1.27
	German	4.35	0.57	4.30	0.80
	English	4.36	0.92	4.33	0.80

Note. The response format of the single scales consisted of a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*very strongly*). Grades are depicted for achievement, therefore the grades between 1 (*very good*) and 6 (*insufficient*) were inverted. Thus high values on the inverted grades variables represent good achievement outcomes. The multi-level structure was considered.

If one concentrates on the emotional aspects of the questionnaire, he would find it remarkable that mean values for the following emotions, enjoyment, anxiety and anger, are similar within similar subjects. To this can be added the fact that, for mathematics and physics, mean values of student's anxiety and anger, are higher than in comparison to their values for German and English, and that the mean values of enjoyment for quantitative subjects are much lower compared to language subjects.
Pertaining to interesting differences observed, in students from different class levels, is the fact that older students report higher average levels of anger and anxiety in mathematics and physics, and also shows lower achievement mean values in mathematics.

To facilitate the observation of differences in mean values of the emotion variables across the four subjects, these values were illustrated using graphics and can be visualized in Figure 8 and Figure 9. Comparing these two figures, it is obvious that similar mean values for corresponding age groups are described by the present diagrams.

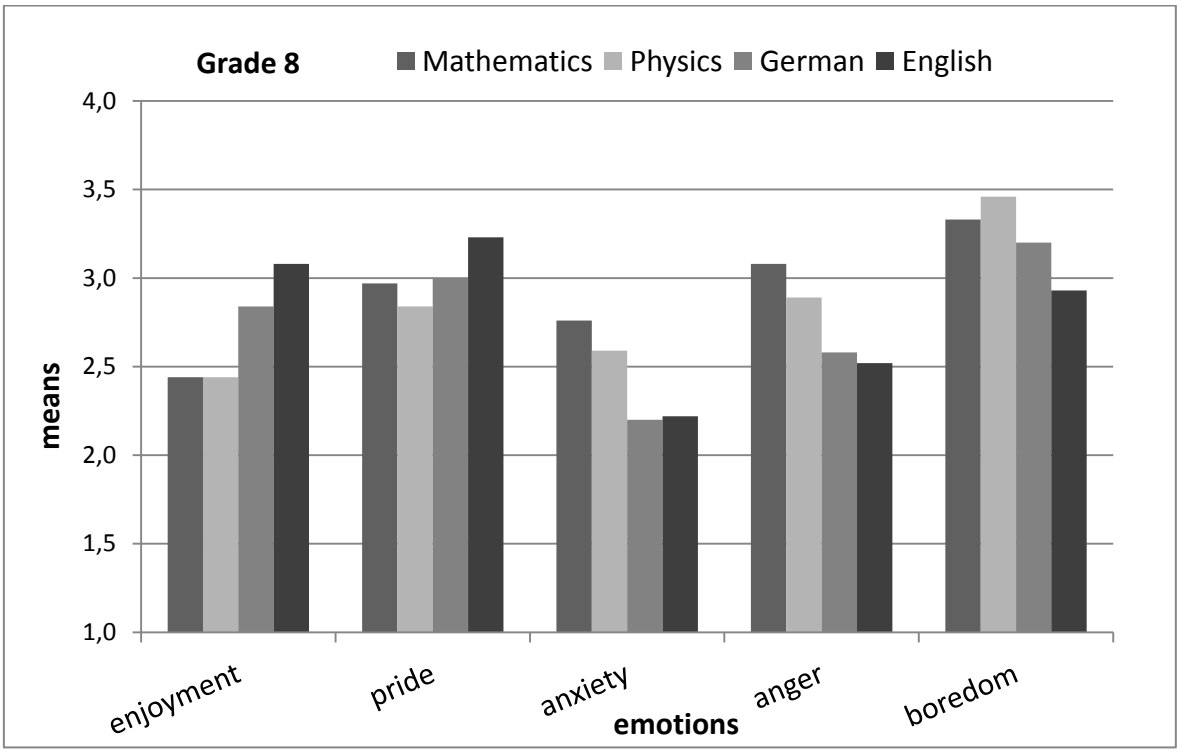


Figure 8: Means of the Achievement Emotions; Grade 8 (Questionnaire Study).

Note. Possible answers range from 1 (*not at all*) to 5 (*very strongly*).

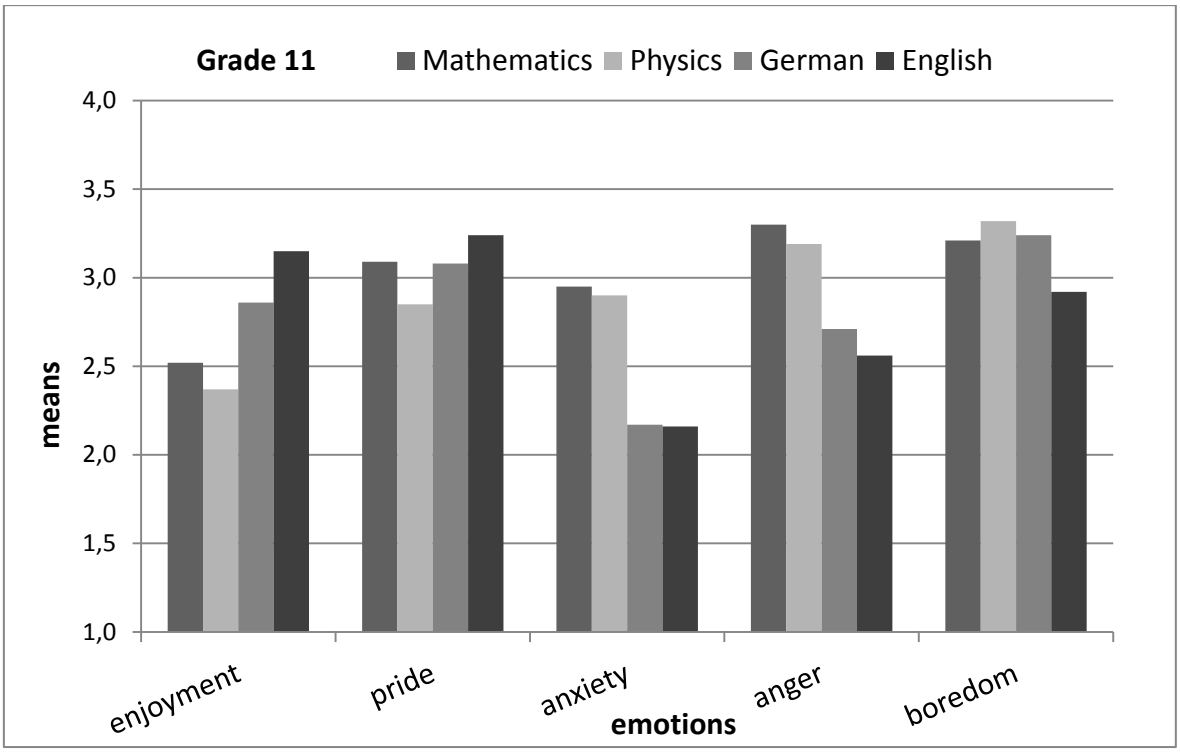


Figure 9: Means of the Achievement Emotions; Grade 11 (Questionnaire Study).

Note. Possible answers range from 1 (*not at all*) to 5 (*very strongly*).

Correlations between Categories and Emotions

In order to get a first impression of the relations between the categories and the emotions, correlations between these two variables were calculated for all domains and overall students but in a separate manner for the 8th graders and 11th graders. The results were summarized in Table 5 and 6. Even though the calculations were made for the two different age groups separately, descriptions will concentrate on explaining values by pairs, as values of correlations for the two different age groups do not differ by much in all cases. Table 5 and 6 together describing a large number of statistics, only the most interesting will be discussed below.

First, in order to have an overview, categories who correlate remarkably high, or remarkably low with the emotions will be singled out. Indeed, first it is interesting to note that over all subjects categories “variety in lesson content”, “subject matter importance”, “effort”, and lastly “level of difficulty” correlate high with emotions. Second, it is equally interesting to note that categories “topic relation”, “class discussion”, and “current topics” correlate low with emotions and category “perceived learning requirements” correlating

even lower with emotions. In addition, it is to be noted that in general, there are more high correlations between categories and emotions, for mathematics and physics compared to the language subjects. For the two quantitative domains, most of the correlations between the category “amount of subject matter” and the emotions are rather strong. Going down to details in a second part, one can observe that the category “amount of subject matter”, for the same two domains, correlates negative with emotions dubbed positive (enjoyment, pride), and correlates positive with emotions dubbed negative (boredom, anger and anxiety). Some correlation values are indeed more outstandingly higher than others, like positive correlations between the category “variety in lessons” and enjoyment, negative correlations between the same category and boredom, positive correlations of the category “subject matter importance” with enjoyment, or with pride, and negative correlations of the same category with boredom, the positive correlation between category “effort” and anxiety on one hand, and anger on the other, and finally positive correlations of category “level of difficulty” with enjoyment, as well as negative correlations of the same category, with anxiety on one side, and anger on the other. One last final noticeable strong outlier, specific to quantitative domains, is the correlation between category “effort” and enjoyment emotion, which correlates negatively.

Table 5: *Correlations between Categories and Achievement Emotions in Mathematics and Physics*

Grade 8/Grade 11		Mathematics				
		enjoyment	pride	anxiety	anger	boredom
Mathematics	amount of subject matters	-.41/-.45	-.29/-.32	.45/.45	.42/.45	.19/.19
	variety in lesson content	.45/.48	.30/.31	-.23/-.24	-.28/-.39	-.43/-.48
	importance of the grade	.23/.36	.25/.37	-.07/-.19	-.17/-.26	-.22/-.28
	subject matter importance	.50/.56	.44/.44	-.26/-.36	-.31/-.42	-.34/-.39
	effort	-.46/-.53	-.31/-.34	.51/-.51	.45/.54	.18/.30
	everyday usefulness	.29/.34	.26/.24	-.11/-.19	-.22/-.24	-.28/-.27
	topic relation	.18/.24	.16/.22	-.02/-.07	-.11/-.18	-.23/-.22
	level of difficulty	-.53/-.59	-.39/-.41	.57/-.59	.49/.60	.23/.31
	class discussion	.21/.09	.08/.09	.02/-.02	-.04/-.05	-.16/-.13
	perceived learning requirements	-.11/-.12	-.10/-.14	.15/-.14	.18/.15	.10/.14
	illustration	.27/.32	.23/.23	-.10/-.12	-.23/-.24	-.33/-.33
	right solution	.30/.31	.35/.31	-.29/-.25	-.29/-.29	-.24/-.27
	current topics	.15/.18	.09/.09	-.03/-.08	-.1/-.09	-.19/-.17
		Physics				
		enjoyment	pride	anxiety	anger	boredom
Physics	amount of subject matters	-.41/-.42	-.29/-.31	.41/.40	.41/.48	.29/.33
	variety in lesson content	.51/.58	.36/.43	-.24/-.31	-.34/-.44	-.52/-.57
	importance of the grade	.20/.41	.27/.38	.02/-.16	-.12/-.27	-.17/-.36
	subject matter importance	.53/.62	.47/.47	-.22/-.40	-.36/-.50	-.45/-.54
	effort	-.41/-.48	-.25/-.30	.47/.49	.43/.55	.24/.38
	everyday usefulness	.37/.41	.28/.35	-.19/-.25	-.22/-.35	-.33/-.36
	topic relation	.19/.32	.24/.27	-.08/-.14	-.17/-.29	-.24/-.26
	level of difficulty	-.51/-.55	-.35/-.38	.51/.52	.45/.58	.35/.41
	class discussion	.28/.26	.19/.14	-.11/-.12	-.17/-.16	-.28/-.31
	perceived learning requirements	-.10/-.11	-.12/-.09	.12/.15	.11/.11	.12/.02
	illustration	.26/.29	.23/.27	-.15/-.14	-.16/-.23	-.24/-.33
	right solution	.33/.29	.27/.23	-.24/-.26	-.27/-.31	-.27/-.27
	current topics	.29/.32	.18/.18	-.08/-.11	-.14/-.21	-.27/-.27

Table 6: *Correlations between Categories and Achievement Emotions in German and English*

Grade 8/Grade 11		**German**				
		enjoyment	pride	anxiety	anger	boredom
German	amount of subject matters	-.18/-.19	-.17/-.15	.30/.20	.24/.23	.09/.13
	variety in lesson content	.56/.60	.39/.37	-.15/-.13	-.29/-.33	-.49/-.54
	importance of the grade	.19/.39	.26/.40	.01/-.08	-.10/-.25	-.20/-.36
	subject matter importance	.37/.55	.36/.44	-.08/-.14	-.21/-.37	-.34/-.50
	effort	-.14/-.33	-.07/-.21	.33/.42	.25/.36	.08/.23
	everyday usefulness	.31/.42	.23/.32	-.08/-.02	-.15/-.25	-.28/-.40
	topic relation	.22/.37	.26/.18	-.06/-.05	-.07/-.24	-.22/-.34
	level of difficulty	-.21/-.36	-.17/-.30	.36/.42	.28/.35	.12/.25
	class discussion	.27/.31	.23/.21	-.06/-.07	-.09/-.21	-.17/-.25
	perceived learning requirements	-.07/-.07	.06/-.08	.02/0.8	.08/.12	.06/.11
	illustration	.28/.29	.24/.18	-.05/-.04	-.17/-.21	-.27/-.33
	right solution	.32/.19	.27/.11	-.14/-.08	-.24/-.18	-.32/-.24
	current topics	.34/.34	.29/.26	-.12/-.07	-.24/-.16	-.24/-.27
		English				
		enjoyment	pride	anxiety	anger	boredom
English	amount of subject matters	-.31/-.21	-.27/-.17	.32/.32	.26/.23	.12/.06
	variety in lesson content	.51/.54	.40/.34	-.20/-.10	-.29/-.36	-.53/-.56
	importance of the grade	.23/.28	.30/.33	-.03/-.09	-.17/-.19	-.23/-.16
	subject matter importance	.31/.42	.37/.33	-.08/-.18	-.24/-.30	-.28/-.28
	effort	-.30/-.32	-.21/-.22	.37/.42	.26/.34	.15/.11
	everyday usefulness	.24/.33	.30/.31	-.05/-.11	-.15/-.19	-.21/-.26
	topic relation	.25/.26	.27/.15	-.10/-.03	-.11/-.17	-.24/-.33
	level of difficulty	-.41/-.38	-.35/-.30	.45/.51	.38/.38	.19/.18
	class discussion	.29/.32	.19/.26	.03/-.11	-.03/-.24	-.21/-.31
	perceived learning requirements	-.01/.02	.04/-.01	.02/.02	.03/.00	.02/.06
	illustration	.35/.30	.30/.22	-.09/-.01	-.21/-21	-.34/-.30
	right solution	.27/.19	.26/.17	-.15/-.11	-.23/-.23	-.26/-.19
	current topics	.37/.33	.25/.25	-.11/-.12	-.17/-.21	-.31/-.32

3.3.2 Domain Specificity of Achievement Emotions

One object of this dissertation is the study of the existence of domains' specific organization of achievement emotions, and if there is one, the research of its possible root causes as described in the introduction. Thus succeeding parts of the analysis focused on the domain specificity of achievement emotions.

The first step into proving domain specificity across the data collected was the calculation of the emotion specific between-domain correlations. In a second step, the domain-specific achievement scores were correlated to the domain-specific emotions. A summary of the results of these calculations can be found in Table 7. Subsequently out of these set of correlations, in order to create as previously announced an easy and reliable similarity index which could describe the strength of between-domain relations across overall emotions, the average (Ø-) correlations between any two domains were calculated. The results were summarized in Table 8.

From the observations of these two tables one can conclude that in general the between-domain relations of emotions are relatively weak, in fact all coefficients were as low as .30 except for the greater correlations observed between emotions, in mathematics and physics on the one hand and in English and German on the other. Relations between any one emotion in mathematics and the same emotion in physics ranged from .28 to .71. These findings point out that even across two domains or subjects intuitively perceived as similar, such as mathematics and physics, the emotional experiences were quite dissimilar. But still the interrelation between same emotions experienced in two similar subjects is stronger, compared to correlations between this specific same emotions experienced in two dissimilar subjects.

Taking students' performance into account, the following results are interesting. First, it is interesting to see that correlations of achievement scores between any two subjects are in most cases stronger than the interrelation between same emotions experienced in the same two subjects. Equally remarkable was that any relation between emotions and achievement scores pertaining to the same domain was found to be the strongest relation observed. Altogether the relations between emotions dubbed as positive (enjoyment, pride) and achievement were positive, and in contrast, negative relations were found between emotions dubbed negative (anxiety, anger, boredom) and achievement.

Table 7: *Between-Domain Relations of Achievement Scores and Emotions (Questionnaire Study)*

	Achievement					Emotions																			
						Enjoyment				Pride				Anxiety				Anger				Boredom			
	M	P	G	E		M	P	G	E	M	P	G	E	M	P	G	E	M	P	G	E	M	P	G	E
													Grade 8												
					M	1				1				1				1				1			
					P	.35**	1			.49**	1			.51**	1			.42**	1			.28**	1		
					G	.03	.10**	1		.18**	.19**	1		.25**	.30**	1		.20**	.20**	1		.15**	.16**	1	
					E	.06	.00	.27**	1	.20**	.19**	.37**	1	.27**	.21**	.41**	1	.14**	.18**	.37**	1	.13**	.15**	.30**	1
M	1					.40**	.22**	–.03	–.07	.40**	.24**	.01	–.03	–.44**	–.23**	–.00	–.01	–.40**	–.23**	–.02	–.01	–.16**	–.14**	.00	.04
P	.61**	1				.23**	.40**	–.04	–.11*	.25**	.39**	–.01	–.05	–.25**	–.35**	.02	.04	–.24**	–.41**	–.01	.02	–.08	–.27**	.03	.06
G	.31**	.30**	1			.01	.02	.24**	.09*	.10**	.09*	.28**	.12**	–.04	–.02	–.25**	–.10*	–.05	–.04	–.21**	–.08*	–.05	.01	–.20**	–.05
E	.34**	.29**	.43**	1		.02	.00	.08	.30**	.09**	.07	.13**	.31**	–.05	–.03	–.04	–.31**	–.07*	–.05	–.06	–.24**	–.030	–.02	–.10**	–.19**
													Grade 11												
					M	1				1				1				1				1			
					P	.48**	1			.61**	1			.71**	1			.50**	1			.38**	1		
					G	–.08*	–.06	1		.16**	.09	1		.22**	.26**	1		–.03	–.02	1		–.02	–.04	1	
					E	–.06	–.06	.17**	1	.15**	.14**	.34**	1	.23**	.27**	.43**	1	–.02	–.02	.13*	1	.03	–.03	.12*	1
M	1					.51**	.31**	–.12**	–.17**	.41**	.30**	–.11**	–.13**	–.48**	–.35**	–.01	.02	–.47**	–.30**	.12**	.10*	–.26**	–.24**	.12**	.10*
P	.73**	1				.39**	.48**	–.12**	–.13*	.33**	.40**	–.16**	–.15**	–.39**	–.47**	.00	.00	–.37**	–.49**	.15**	.10*	–.19**	–.34**	.18**	.10*
G	.35**	.31*	1			.00	–.07	.30**	.06	.05	–.02	.30**	.08**	.01	.01	–.29**	–.10**	–.01	.04	–.26**	–.06	.01	.05	–.24**	–.05
E	.39**	.35**	.55**	1		–.02	–.07	.05	.33**	.03	–.02	.07**	.32**	–.01	–.04	–.10*	–.38**	.00	.04	.01	–.35**	.05	.06	–.02	–.18**

Note. *p < .01; **p < .001; N = 855/854; Mathematics (M), Physics (P), German (G) and English (E).

Table 8: *Average Between-Domain Relations of Achievement Emotions (Questionnaire Study)*

Ø-correlation (emotions)				
	M	P	G	E
	Grade 8			
M	1			
P	.41	1		
G	.16	.19	1	
E	.16	.15	.34	1
	Grade 11			
M	1			
P	.55	1		
G	.10	.09	1	
E	.10	.10	.24	1

Note. The absolute values of the between-domain relations were Fisher-z-transformed, then the arithmetic means over every emotion was calculated and afterwards retransformed and z-standardized; N = 855/854; Mathematics (M), Physics (P), German (G) and English (E).

3.3.3 Age-related Differences of Domain Specificity

Next furthering the study as announced in the introduction to questions related to existing differences between different age groups with regard to domain specificity of achievement emotions, a specific analysis was conducted. This task was greatly simplified by the fact that all calculations contained in previous sections of Study II were done in as separated manner for the two age groups.

Thus, Table 7 shows that nearly all relations between an emotion experienced in any one quantitative domain, and the same emotion experienced in any other verbal domain are weaker if observed for older students. The same observations can be made looking at relations between the two language subjects. In the cases described above only two values increase with age but only slightly: The correlations of anxiety between English and physics on the one hand and between English and German on the other hand. In fact for the older group, all average correlations of the emotions between any two subjects – except between mathematics and physics – are less substantial compared to the results obtained for the younger group. In reality, the spread of observed average correlations between the two age groups fluctuate from .05 to .1 (see Table 8).

By contrast, the already relative strong relations between mathematics and physics are even stronger if observed for the 8th graders compared to 11th graders as we can see it in

Table 7. The average correlation is actually about .41 for 8th graders and .55 for 11th graders as observed in Table 8.

Then, to evaluate if data supports Hypothesis 2 as described previously homogeneity index *h*(d) were calculated and summarized in Table 9. The observation of this table shows that anger and boredom have a higher *h*(d) index value for the sample composed of younger students. This implies that, the between-domain relations of anger and boredom are significantly stronger for the 8th graders sample, compared to the 11th graders sample. For all other emotions *h*(d) is of about the same value, for both observed samples.

Table 9: *Results of Multilevel Analysis of the Between-Domain Relations of Single Academic Emotions and Academic Achievement (Questionnaire Study)*

	Variance components			
Variable	Level 2: Domain-transcending scores of person	Level 1: Domain-specific score within person	Homogeneity of domains: Homogeneity index *h*(d)	95% confidence interval of *h*(d)
Grade 8				
Enjoyment	.191	1.479	.114	(.089-.139)
Pride	.381	1.074	.262	(.239-.285)
Anxiety	.563	1.295	.303	(.279-.327)
Anger	.394	1.279	.235	(.215-.255)
Boredom	.308	1.379	.183	(.161-.205)
Achievement	.350	.587	.374	(.337-.411)
Grade 11				
Enjoyment	.077	1.561	.047	(.030-.064)
Pride	.384	1.181	.245	(.219-.271)
Anxiety	.590	1.315	.310	(.293-.327)
Anger	.119	1.584	.070	(.049-.091)
Boredom	.113	1.475	.071	(.053-.089)
Achievement	.472	.624	.431	(.343-.519)

3.3.4 Relations of Domain Characteristics to Domain Specificity of Achievement Emotions

In a final stroke of pen, an investigation was lead in order to look at if findings support Hypothesis 3 as described previously, or rather if data support the idea that the relations of the characteristics between domains intuitively perceived as more similar would be stronger than across dissimilar domains. And if similarities in the category profiles of two subjects imply similarity in the emotional experience in these same subjects.

Therefore, the between-domain relations of the categories were calculated, and results organized and presented in Table 10. Additionally the average correlations of the categories were calculated and put up in Table 11. Both of these tables can be found below. These calculated statistics show that the relations of the categories of characteristics between mathematics and physics on one side, and German and English on the other side are stronger than across between one quantitative and one verbal domain. Analog observation was made for the relations of the achievement emotions from Table 7 and 8.

It is striking to see how much the average correlations of the categories described in Table 11 resemble the average correlations of the emotions described in Table 8. Thus a significance test which compared every single value of Table 8 with the corresponding correlation in Table 11 was conducted and showed that nearly all pairs of values are significantly similar or at least nearly significantly similar.

Likewise, it can be said with respect to age differences, that similar findings were made concerning the categories as well as the emotions. The correlations of the categories between the four different subjects as shown in Table 11, are noticeably weaker for the 11^{th} graders sample. Only the relations between mathematics and physics are unusually stronger for the 11^{th} graders sample. The same exact observation can be made concerning correlations of the emotions between the four different subjects.

Table 10: *Between-Domain Relations of Categories (Questionnaire Study)*

	M	P	G	E	M	P	G	E	M	P	G	E	M	P	G	E	M	P	G	E
	amount of subject matters				variety in lesson content				importance of the grade				subject matter importance				effort			
									Grade 8											
M	1				1				1				1				1			
P	.21**	1			.26**	1			.38**	1			.46**	1			.28**	1		
G	.15**	–.01	1		.08*	.08	1		.44**	.15**	1		.20**	.05	1		.13**	–.05	1	
E	.04	.07	.13*	1	.09	.09*	.25**	1	.44**	.25**	.52**	1	.26**	.11**	.34**	1	.16**	.04	.22**	1
									Grade 11											
M	1				1				1				1				1			
P	.45**	1			.31**	1			.53**	1			.52**	1			.49**	1		
G	.05	.00	1		–.02	.01	1		.32**	.13**	1		.02	–.09*	1		–.06	.00	1	
E	.08*	.06	.24**	1	.05	.04	.12	1	.30**	.18**	.46**	1	.12**	.04	.26**	1	.03	.06	.11*	1
	everyday usefulness				topic relation				level of difficulty				class discussion				perceived learning requirements			
									Grade 8											
M	1				1				1				1				1			
P	.26**	1			.22**	1			.40**	1			.31**	1			.28**	1		
G	.19**	.05	1		.12**	.22**	1		.04	–.01	1		.00	.07	1		.15**	.08*	1	
E	.17**	.16**	.29**	1	.12**	.13**	.19**	1	.07	.03	.25**	1	.27**	.24**	.19**	1	.11**	.17**	.23**	1
									Grade 11											
M	1				1				1				1				1			
P	.34**	1			.29**	1			.58**	1			.51**	1			.54**	1		
G	.01	–.01	1		–.05	.08	1		–.07*	–.07	1		– .07	–.06	1		.07	.02	1	
E	.08*	.13**	.25**	1	–.05	.05	.20**	1	.02	.07	.29**	1	.07	.02	.29**	1	.17**	.19**	.21**	1
	illustration				right solution				current topics											
					Grade 8															
M	1				1				1											
P	.09	1			.46**	1			.39**	1										
G	.12**	–.03	1		–.23**	–.20**	1		.07*	.07	1									
E	.16**	–.02	.31**	1	–.06	.02	.23**	1	.26**	.29**	.15**	1								
					Grade 11															
M	1				1				1											
P	.23**	1			.64**	1			.46**	1										
G	.07	–.08*	1		–.35**	–.30**	1		.03	.03	1									
E	–.04	–.06	.16**	1	–.21**	–.15**	.35**	1	.05	.06	.20**	1								

Note. *p < .01; **p < .001; N = 855/854; Mathematics (M), Physics (P), German (G) and English (E);

Table 11: *Average Between-Domain Relations of Categories (Questionnaire Study)*

Ø-correlation (categories)				
	M	P	G	E
	Grade 8			
M	1			
P	.31	1		
G	.15	.08	1	
E	.17	.12	.26	1
	Grade 11			
M	1			
P	.46	1		
G	.09	.07	1	
E	.10	.09	.24	1

Note. The absolute values of the between-domain relations were Fisher-z-transformed, then the arithmetic means over every category was calculated and afterwards retransformed and z-standardized; N = 855/854; Mathematics (M), Physics (P), German (G) and English (E);

3.4 Discussion

The present study had multiple goals. First, it had the mission to develop scales out of the categories of characteristics listed by Study I, to assess these scales with the help of data collected using a questionnaire survey, and to measure to which extent described categories characterized the considered domains. It was then used to confirm the conclusions of previous educational psychology researches in which students' emotional experiences were shown to be largely domain-specific, and to further analyze age-related difference of domain specificity. Another objective fixed was the investigation of relations of categories and students' emotional experience in different school subjects.

3.4.1 Domain Specificity of Achievement Emotions

Pushing through our announced objectives, data are analyzed below to confirm or to infirm the conclusions of previous educational psychology researches, in which students' emotional experiences were shown to be largely domain-specific.

In agreement with the assumptions made at the start of this study (Hypothesis 1) and previous findings (e.g. Goetz et al., 2006a; Goetz et al., 2007), the results indicate that students' emotional experience is largely organized in a domain-specific way. All between-domain relations of the emotions are relatively weak. However, between similar

subjects, namely mathematics and physics on the one hand, and German and English on the other hand, correlations were found to be stronger than between dissimilar subjects. It was also observed, quite understandably maybe, that relations between quantitative domains are stronger compared to relations between verbal domains. And this is probably because the study of German as a mother language, is lived quite differently than the study of English as a foreign language, whereas mathematics and physics are maybe more equally unfamiliar/familiar to students.

It is also to be noted that in addition, interrelations between emotions and achievement suggest the domain specificity of academic emotions. Indeed the relations between emotions and achievement scores confirm the assumption of weak between-domain relations, and in fact indicate high convergent and discriminate validity for the latent domain-specific emotion factors. In conclusion it can be said that these findings point toward the need to assess students' emotional experience not in a general-domain manner, but rather as a domain-specific constructs.

3.4.2 Age-related Differences of Domain Specificity

Now as stated previously, one aim of the present study being the analysis of hypotheses stated above, this part goes through the list, and tries to confirm or infirm Hypothesis 2.

In accordance with this hypothesis and consistent to findings of previous works on academic emotions (compare Goetz et al., 2007) a multilevel-analysis showed that the between-domain relations are in general relatively weak. But also that they are even weaker for the sample consisting of older students, compared to the sample consisting of younger students, implying that the domain specificity of achievement emotions gets stronger with age. Additionally, the relations between emotions pertaining to dissimilar subjects were weaker, for 11th graders compared to 8th graders, at times being even negative.

Likewise, relations between emotions in German and English which were weak for the 8th graders sample are even weaker for the 11th graders sample. However the relations of the emotions between mathematics and physics show quite the opposite, the between-domain correlations being relatively stronger for the younger age group, stronger or at least as strong for the older age group. These results are in line with the study by Goetz et

al. (2007), and Marsh and Ayotte's differential distinctiveness hypothesis (2003) which predicts that constructs with little correlations for samples composed of younger students get more differentiated for samples composed of older students. This hypothesis is transferable to academic emotions.

3.4.3 Relations of Domain Characteristics to Domain Specificity of Achievement Emotions

Going through the list of hypotheses, the study now focuses on whether the data support Hypothesis 3.

And as expected and as anticipated by Hypothesis 3, the correlations of the categories are much stronger between similar subjects, namely between mathematics and physics on one side and between German and English on the other, than between any two dissimilar subjects. In addition the same observations can be made for the correlations of the achievement emotions. The average correlations of the categories as summarized in Table 9 and of the emotions as summarized in Table 6 clearly demonstrate, average correlation in both cases have the same structure. It can be said in particular that the correlations, concerning the single categories as well as the single emotions between mathematics and physics are relatively strong. This is also noticeable in the average correlations. Though between German and English these same correlations are a bit less stronger they are still higher than all other combinations of different subjects, for which values are relatively weak. More importantly, the significance tests show that nearly all average correlations of the categories between two subjects are similar to the average correlations of the emotions between the same two subjects. That means that if two subjects are described more similar, their emotional experience is also more similar.

The above observations beg a question: Why is the emotional experience more similar in school subjects that are perceived as similar? The reason could be the similar description or characterization of the school subjects. In fact the categories used to make the characterization of the domains and memory-based reported emotions, are both strongly dependent on the beliefs about different school subjects. This is to say that the two have a

strong relationship with each other. That is probably why subjects that appear to be similar are also characterized similar and moreover are perceived emotionally similar.

One simple implication can be deduced out of these conclusions. For students emotions experienced during lessons are closely linked to the image that the students have about the subject. So that is why teachers should be aware of this fact in order to create conducive learning environment. Hence they should not only focus on imparting knowledge but also on conveying a positive image of their subject.

4 Study III – Experience Sampling Study: Domain Specificity of Achievement Emotions in Real Life

4.1 Research Aims and Hypotheses

One of the most pertinent observations made in the introduction is one pertaining to memory-based questionnaires. Indeed, it was already stated that most domain specificity studies are memory-based questionnaires which assess dispositions or the so-called trait emotions, rather than the current emotional situation, also called state emotions, which could be measured from online reports.

However since emotional experience being likely to be specifically dependent on precise moments and situations, it was considered to use an alternative emotion evaluation method, namely the experience sampling method which takes this problem into consideration.

As in Study II, the current study investigated, in general and separately for two age groups, the domain specificity of the emotional experiences, and later on analyzed the differences between the two age groups. Thus, when conducting this study two main hypotheses were taken into account, which will be explained in below. In fact, these hypotheses will be looked upon, in later sub-chapters in order to verify whether data support them or not.

Hypothesis 1*: Domain Specificity of Achievement Emotions*

It is assumed that the relations of single emotions – enjoyment, pride, anxiety, anger and boredom – across domains mathematics, physics, German and English would be rather weak. Further it is expected that the correlations between subjects, that are intuitively considered to be similar (like mathematics and physics, or German and English), are stronger than across subjects that are intuitively considered to be dissimilar, and that the strongest relationship should be observed between emotions and achievement scores related to the same domain.

***Hypothesis 2**: Domain Specificity Increases with Age*

It is also hypothesized that more domain specificity should be observed in the group of 11th graders compared to the 8th graders; this means that the between-domain relations of academic emotions should be stronger for younger students compared to older.

Referring to Marsh and Ayotte's (2003) differential distinctiveness hypothesis, it is also hypothesized that the between-domain relations should decrease with age for emotional experience in disparate domains and there should be less or no decrease in related subjects.

4.2 Method

4.2.1 Participants and Data Collection

As stated above, this study assessed emotions using data that were collected by the experience sampling method (Csikszentmihalyi & Larson, 1987; Hektner, et al., 2007). Indeed data were retrieved throughout 14 days, spread between May and July 2009.

For this purpose personal digital assistant (PDA) devices programmed with PMat software (Weiss, Beal, Lucy, & MacDermid, 2004) were utilized. Students were instructed to activate their device at the beginning of every, mathematics, physics, German and English class they had, and the device would signal the participant to answer a digital questionnaire at a randomly chosen moment, within the 45 minutes of lesson. All teachers were informed of the experimental procedure at the beginning of the study and they gave their agreement to proceed, as did the students. Each student completed at least one or at most 28 experience sampling questionnaires. In fact students completed an average of 12.38 questionnaires throughout the 14 days (*SD*=6.04), and looking at details, each individual student has been probed 1 to 10 times during mathematic classes, or 3.89 (*SD*=2.16) times in average over all students, 1 to 8 times during physics classes, or 3.09 (*SD*=1.80) times in average, 1 to 9 times during German classes, or 3.63 (*SD*=2.08) times in average, and 1 to 8 times during English classes, or 3.27 (*SD*=1.71) times in average. Immediately before being provided with the PDAs, students filled in a paper questionnaire which included a survey on the demographic data and on self-reported grades.

The number of participants was as high as 121, 50.4% of whom were female students, and drawn from 41 different classes, 21 of these being 8th grade and 20 of these 11th grade classes. They were recruited from seven different high schools in Munich, Germany. From each of these 41 classes, two to four students were randomly chosen for participation in the experience sampling study. The mean age of participants was 16.16 years (*SD* = 1.84), with 58 students being from Grade 8, and 63 students from Grade 11. The proportion of female students and mean age were 51.7% and 14.45 years (*SD* = .76), respectively for the 8th graders sample, and 49.2% of females and 17.73 years (*SD* = .88), respectively for the 11th graders sample. All participants took part in this survey on a voluntary basis and with the assurance that the data would be handled anonymously.

4.2.2 Measures

Most achievement emotion data – as in Study II – were collected by means of a one-time raised questionnaire. Even, if the goal of this study is a trait-oriented assessment (e.g. "usually I have a lot of joy") and not a state-orientated one (e.g. "I have a lot of joy at this moment"), by taking particular situations into account could probably increase the content validity. In fact as previous works remarked, even if one might think that experience sampling method assess state emotions in truth the averaging of state emotions gives an accurate description of trait emotions, as Goetz, Hall, Frenzel, & Pekrun (2006) put it "trait emotions may be conceptualized as cumulative state emotional experience" (p. 325). This implies that the average values of several state emotions, which could vary depending on situations, can be considered as more reliable trait emotions. Since the trait emotions display the cross-situational disposition or the typically experienced emotion. In the present experience sampling study (Study III) several state emotions were collected over a span of 14 days, and averaged as single data, for each student. This implies that these averages of the state emotions can be equated with the trait emotions. Though not as usually measured by memory-based report but rather as assessed with an online report, this method probably allows an even more precise trait assessment since it is improved by the introduction of situational context factors. In general, only a few studies take the situational aspect into account while dealing with such online reports.

Single Emotions in Different Domains

As in Study II, well established scales were used to measure the emotions: Enjoyment, anger, pride, anxiety and boredom pertaining to the following domains: Mathematics, physics, German and English. The scales were adapted from the mathematic version of the Achievement Emotions Questionnaire [AEQ-M] by Pekrun, Goetz and Frenzel (2005). The exact wording of the questions associated to each of the emotion scales was in the fashion of "How strongly do you experience … (*the corresponding emotion*) … in the following subjects?" The response format consisted of a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*very strongly*). Questions were submitted for every school subject.

Academic Achievement

Similarly to Study II, achievements in mathematics, physics, German and English were assessed by students' midterm grades, though grades were not assessed several times in each lesson but only once and immediately before the PDAs were distributed. German grades range from 1 (*very good*) to 6 (*not sufficient*), however these grades were inverted, so that high numbers indicate better performance in order for values to be interpreted more easily.

4.2.3 Data Analysis

The current study was analyzed in a way that is analog to Study II's. In particular, it should be noted that the multiple answers – observed from the measurement of several lessons during a span of two weeks – were averaged for every student. Subsequently every student was assigned a single value for each emotion in each of the subjects. Figure 10 below illustrates the exact procedure employed for the analysis of these data.

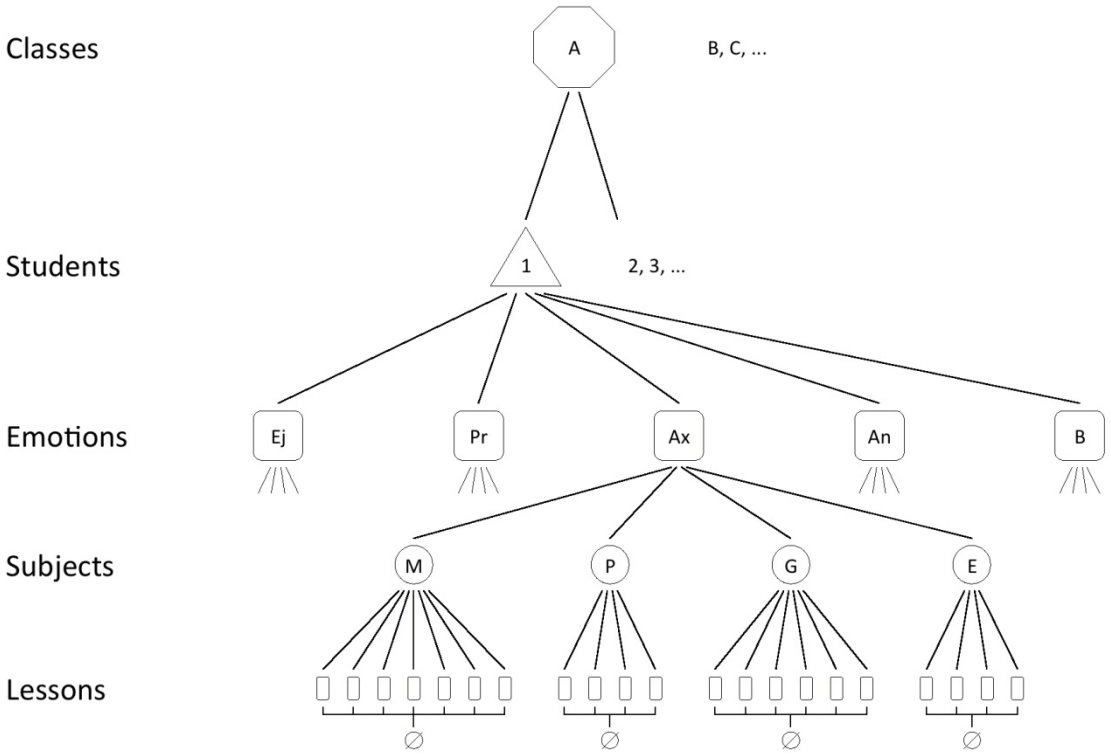


Figure 10: How to Average Multiple Answers per Student (Experience Sampling Study).

Note: Enjoyment (Ej), Pride (Pr), Anxiety (Ax), Anger (An) and Boredom (B); Mathematics (M), Physics (P), German (G) and English (E).

These values – single values for each student – represent the building block data for the calculations of single between-domain correlations and their average correlations as well as the homogeneity index, which were calculated in the following way.

The calculation of the between-domain relations takes into account the multi-level nature of our collected data. In fact, the data presents a two-level structure, in which students (Level 1; N = 121) are nested within classes (Level 2; N = 41). The analyses presented below were conducted via multilevel statistics by applying the Mplus 5.2 software (Muthén & Muthén, 1998-2008). In order to create a reliable similarity index, which could describe the strength of between-domain relations across overall emotions, average (Ø-) correlations between any two domains were calculated. Therefore, the absolute values of the between-domain relations were Fisher-z-transformed, and then the arithmetic average over every emotion, was calculated, and afterwards retransformed and z-standardized.

As in Study II the homogeneity index *h*(d) showing the strength of between-domain relations of the domain-specific single emotion scales was computed. Therefore once again a fully unconditional two-level model was created, in which information concerning the outcome variability on both of the levels was provided (see Figure 3; Raudenbush & Bryk, 2002). In order to compare the two age groups, two-level models were created separately for both grades. The higher the *h*(d) value is, the stronger are the between-domain relations, for the single emotions.

4.3 Results

4.3.1 Descriptive Statistics

It was already stated that the goal of this study was the alternative assessment of emotions. Hence data were accumulated for this specific reason. It has also been found that the nature of the collected data was multilevel, and therefore multilevel statistics analysis of these was conducted and a number of statistics were generated. These are going to be described in the following.

The mean values and the standard derivations as described in Table 10 led to the notice that, the measurements of the emotions give very similar results for all observed domains. These results, illustrated in figures 11 and 12, show that there are only three small outliers. The younger students experience in average more enjoyment in English, than in the other subjects, and 11th graders reported a higher level of pride in English, and less boredom, in mathematics, in average, compared to the three other corresponding domains. Moreover, it is striking that the average experience of pride, anxiety and anger are relatively low, in general.

Table 12: *Descriptive Statistics of Academic Emotion Scales and Achievement Scores (Experience Sampling Study)*

		Grade 8 (n=58)		Grade 11 (n=63)	
		M	SD	M	SD
Enjoyment					
	Mathematics	2.15	0.67	2.23	0.92
	Physics	2.21	0.67	2.24	1.03
	German	2.37	0.95	2.36	1.09
	English	2.77	0.95	2.46	1.19
Pride					
	Mathematics	1.63	0.55	1.71	0.76
	Physics	1.65	0.61	1.64	0.71
	German	1.55	0.44	1.59	0.63
	English	1.77	0.61	1.96	1.09
Anxiety					
	Mathematics	1.56	0.40	1.44	0.54
	Physics	1.58	0.51	1.32	0.47
	German	1.34	0.37	1.28	0.41
	English	1.38	0.35	1.30	0.27
Anger					
	Mathematics	1.80	0.76	1.79	1.03
	Physics	1.82	0.79	1.83	0.85
	German	1.59	0.48	1.81	1.23
	English	1.55	0.48	1.75	0.73
Boredom					
	Mathematics	3.10	1.10	2.76	1.27
	Physics	2.94	1.36	3.24	1.57
	German	2.97	1.00	3.20	1.46
	English	3.14	1.07	3.09	1.03
Achievement					
	Mathematics	3.91	1.24	3.76	0.94
	Physics	4.02	1.26	3.97	1.21
	German	4.16	0.75	4.29	0.59
	English	4.22	1.00	4.10	0.85

Note. The response format of the single scales consisted of a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*very strongly*). Grades are depicted for achievement, therefore the grades between 1 (*very good*) and 6 (*insufficient*) were inverted. Thus high values on the inverted grades variables represent good achievement outcomes. The multi-level structure was considered.

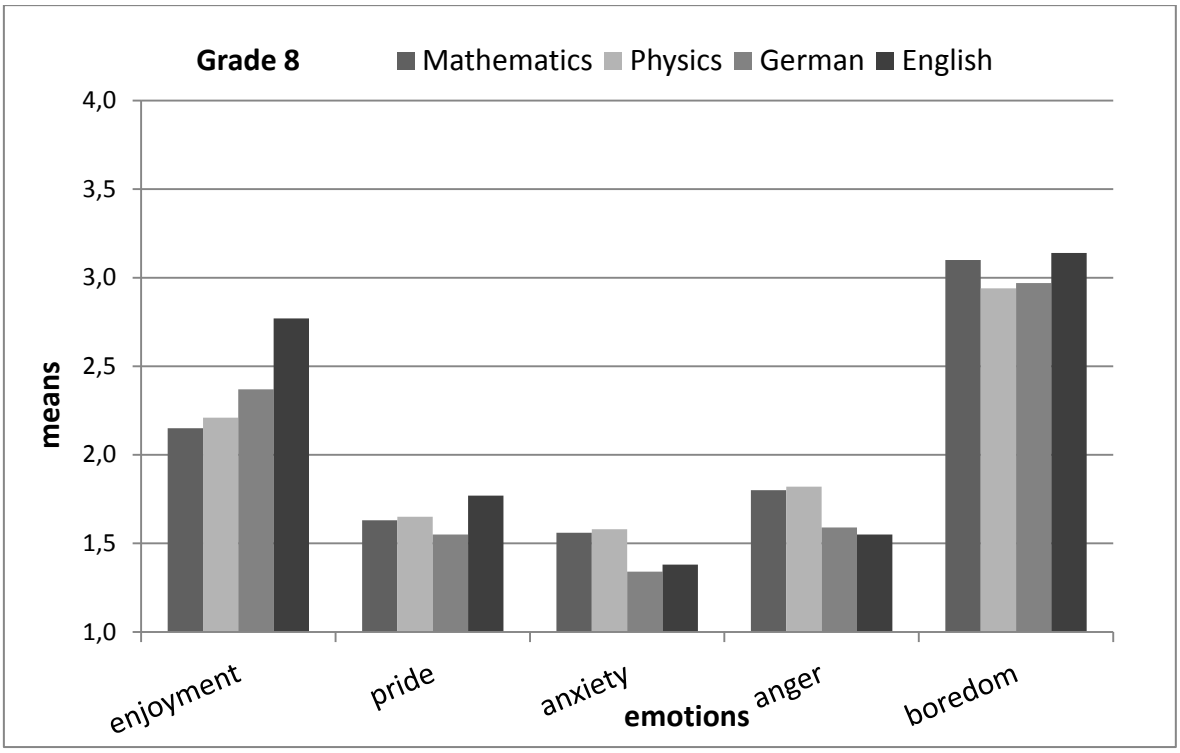


Figure 11: Means of the achievement emotions; Grade 8 (Experience Sampling Study).

Note. Possible answers range from 1 (*not at all*) to 5 (*very strongly*).

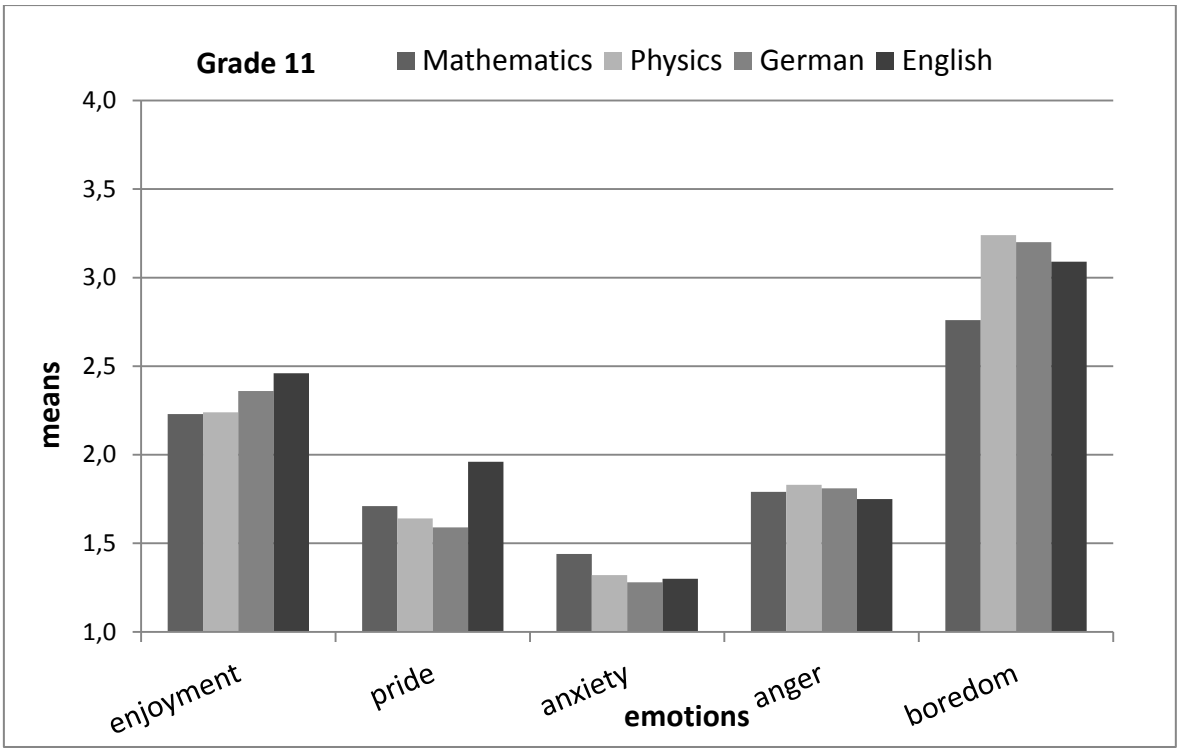


Figure 12: Means of the Achievement Emotions; Grade 11 (Experience Sampling Study).

Note. Possible answers range from 1 (*not at all*) to 5 (*very strongly*).

4.3.2 Domain Specificity of Achievement Emotions

The main subject of this dissertation is the study of the existence of domain-specific organization of achievement emotions. Thus, it will be the focus of the succeeding analysis.

The first step, in proving domain specificity across the data collected was the calculation of the emotion specific between-domain correlations. In a second step, the domain-specific achievement scores were correlated to the domain-specific emotions. A summary of the results of these calculations can be found in Table 11. Subsequently out of these sets of correlations the average (Ø-) correlations between any two domains were calculated, in order to create as previously announced an easy and reliable similarity index, which could describe the strength of between-domain relations across overall emotions. These results are summarized in Table 12.

Table 11 indicates that almost all between-domain relations of emotions are relatively weak. They hardly ever are higher than .35 in fact. In addition the between-domain relations in Table 11, and the average correlations of achievement emotions in Table 12 indicate that similar subjects do not have a particularly stronger relation. In fact, the highest average correlations of emotions were observed between physics and German, with .31 for 8^{th} graders.

Another observation we made is that, except only a few exceptions, there are only relatively weak correlations between achievement and emotions; hardly any value is greater than .25. Though it could be said that, in most cases the relations between achievement and positive emotions were mostly positive, and similarly that the relations between achievement and negative emotions were predominately negative.

Table 13: *Between-Domain Relations of Achievement Scores and Emotions (Experience Sampling Study)*

	Achievement					Emotions																			
						Enjoyment				Pride				Anxiety				Anger				Boredom			
	M	P	G	E		M	P	G	E	M	P	G	E	M	P	G	E	M	P	G	E	M	P	G	E
													Grade 8												
					M	1				1				1				1				1			
					P	.32	1			.23	1			.28	1			.28	1			.22	1		
					G	.21	.39**	1		.41**	.45**	1		.28*	.04	1		.09	.20	1		.28	.44**	1	
					E	.23*	.12	.18	1	.38**	.17	.42**	1	.33*	.16	.31*	1	.10	.26	.26	1	.16	.33*	.19	1
M	1					.17	.03	–.24	–.15	.23	.03	–.09	.09	–.12	–.21	.04	–.04	–.11	.16	.00	.00	–.17	–.22	.02	–.26*
P	.70**	1				.22*	.15	–.22	.02	.21*	.10	.05	.22*	–.26*	–.03	.24	.02	–.28**	.09	.08	–.06	–.25*	–.25	–.04	–.29*
G	.43**	.48**	1			.21*	.07	–.16	.16	.13	–.01	.14	.25*	.21	–.05	.18	.16	.14	.19	.19	–.03	–.05	–.07	.20	–.05
E	.34**	.35*	.34**	1		.07	.22	–.16	.18	.20	.15	.15	.25**	.12	.13	.11	.00	.11	.01	.20	–.02	.17*	–.25*	.25**	–.14
													Grade 11												
					M	1				1				1				1				1			
					P	.12	1			.45**	1			–.09	1			.15	1			.27	1		
					G	.22	–.12	1		.17	.08	1		.09	–.06	1		.14	.21	1		–.12	.02	1	
					E	–.02	.21*	.10	1	.27	–.06	–.07	1	.04	.03	.04	1	.17	.10	.21	1	.08	.12	.12	1
M	1					.35**	.08	–.08	–.26	.31**	.09	.03	–.10	–.19*	.01	–.16**	.00	–.26*	.00	–.13	.00	–.15	–.15	.10	.17
P	.65**	1				.08	.22	–.04	–.03	.24	.27*	.02	–.01	–.13	–.31**	–.16*	–.09	–.29*	–.14	–.11	–.21	–.22*	–.04	.12	–.03
G	.01	.22	1			–.11	.03	.05	.21*	–.10	–.05	.15	–.04	–.18	–.10	–.06	–.31**	–.02	.06	–.33*	–.24*	.10	.17	–.18	–.07
E	.19	.32**	.46**	1		–.23	.01	–.09	.13	–.07	–.10	.10	.28*	.09	–.09	–.17	–.22*	.15	–.09	–.19	–.02	–.05	.19	.13	.02

Note. *p < .01; **p < .001; N =58/63; Mathematics (M), Physics (P), German (G) and English (E).

Table 14: *Average Between-Domain Relations of Achievement Emotions (Experience Sampling Study)*

Ø-correlation (emotions)				
	M	P	G	E
	Grade 8			
M	1			
P	.27	1		
G	.26	.31	1	
E	.24	.21	.28	1
	Grade 11			
M	1			
P	.22	1		
G	.15	.10	1	
E	.12	.11	.11	1

Note. The absolute values of the between-domain relations were Fisher-z-transformed, then the arithmetic means over every category was calculated and afterwards retransformed and z-standardized; N =58/63; Mathematics (M), Physics (P), German (G) and English (E).

4.3.3 Age-related Differences of Domain Specificity

As in Study II, results obtained on domain specificity for the two age groups in a differentiated and separate manner were observed in order to elucidate the age-related differences.

The between-domain relations summarized in Table 11 are mostly smaller as calculated for the sample of 11th graders compared to the sample of 8th graders, with just a few exceptions. One of the most exceptional outliers is observed for the emotion pride between mathematics and physics – the younger students' value indicates .23, which is quite small compared to the .45 value obtained for older students. The average correlation values of between-domain relations, as illustrated in Table 11, are all within intervals of .10 and .31, which can only be labeled as relatively low. In addition, without an exception, all average values obtained for the 8th graders sample are higher than their corresponding values obtained for the 11th graders sample.

After checking, whether the collected data for Study III confirms Hypothesis 1, the dissertation proceeds now to test, if data supports or reject Hypothesis 2, based upon observation of single and averaged between-domain correlations. Another important analysis for the same hypothesis is to calculate the homogeneity index *h*(d) as done and summarized in Table 12. The latter indicates that all between-domain relations are stronger for the 8th

graders sample compared to the 11th graders sample in agreement with previous conclusions, except for the emotion anger.

Table 15: *Results of Multilevel Analysis of the Between-Domain Relations of Single Academic Emotions and Academic Achievement (Experience Sampling Study)*

	Variance components			
Variable	Level 2: Domain-transcending scores of person	Level 1: Domain-specific score within person	Homogeneity of domains: Homogeneity index *h*(d)	95% confidence interval of *h*(d)
		Grade 8		
Enjoyment	.158	.717	.181	(.097-.265)
Pride	.182	.371	.329	(.257-.401)
Anxiety	.085	.328	.206	(.143-.269)
Anger	.108	.528	.169	(.108-.230)
Boredom	.302	.839	.264	(.136-.392)
Achievement	.462	.613	.430	(.371-.489)
		Grade 11		
Enjoyment	.095	.974	.089	(.020-.158)
Pride	.089	.732	.108	(.033-.183)
Anxiety	.004	.418	.009	(.000-.064)
Anger	.163	.803	.169	(.105-.233)
Boredom	.081	1.275	.059	(.001-.117)
Achievement	.269	.663	.289	(.231-.347)

4.4 Discussion

As already mentioned, the goal of this study was the alternative assessment of emotions. The other main objective is of course to observe domain specificity of achievement emotion by looking at the validity of Hypothesis 1. The retrieved data were collected for these specific purposes. These data were multi-level in nature and Study III took this into account. It also observed the validity of Hypothesis 2 by discriminating between two age groups as described previously. The results will be discussed subsequently.

4.4.1 Domain Specificity of Achievement Emotions

First to observe Hypothesis 1 the between-domain relations of emotions were calculated. As stated almost all between-domain relations were found to be relatively weak. This in turn implies that the emotional experience is domain-specific. These results confirm the assumptions of the first hypothesis and are consistent with prior research (see Goetz et al., 2007; Goetz et al., 2006). If one takes data on achievement into account, results hint at the convergent and divergent validity of the data.

Surprisingly however, neither correlations between mathematics and physics on one hand, nor between German and English on the other, were particularly stronger compared to other pair matching correlations. These results contradict previous investigations like Goetz et al. (2007) as well as the results from Study II. Looking at the relations of emotions and achievement, it is remarkable that, likewise, there were no stronger relations between the subject mathematics and physic on one hand, and between German and English on the other. It might be the case that these unusual observations could be explained by the dependence of the perception on the beliefs of the school subjects. Explaining this idea, one could state that because emotions, as reported by online reports, depend less on beliefs that students' have about particular school subjects, there are no obvious underlining relations between reported emotions and student's perceptions. That is probably the reason why emotional experience in subjects perceived as similar does not have a particularly stronger relation compared to relations of emotional experience in subjects perceived as different, at least as measured with an online report. It is important to note that the present experience sampling study assess emotions in a way that is less dependent on the beliefs.

4.4.2 Age-related Differences of Domain Specificity

At last, it can be concluded that the results of the present experience sampling study confirms the second hypothesis, and they are also in agreement with former studies (e.g. Goetz et al., 2007). The homogeneity index *h*(d) as indicated in Table 1 show that between-domain relations – except for the emotion anger – are stronger in the 8th graders sample compared to the 11th graders sample. From this observation and from the fact that most between-domain correlations, as calculated for 11th graders are weaker than for 8th graders (see Table 10 and 11), one can conclude that domain specificity of achievement emotions gets even stronger with age. This conclusion is consistent with numerous previous works on domain specificity.

A slightly different version of Marsh & Ayotte's differential distinctiveness hypothesis (2003) was considered elsewhere in this work. In this version, this hypothesis is described as follows: Between-domain relations for emotional experience in between similar school subjects should, increase, stay the same, or only slightly decrease with age. In contrary,

between-domain relations for emotional experience, in between dissimilar school subjects, should only reduce with age. This version was considered as being an equivalent to the original hypothesis since another assumption made throughout this work was that correlations between subjects which are intuitively considered to be similar are stronger, than along subjects that are intuitively considered as dissimilar (see Hypothesis 2). Taking this particular version of the hypothesis it cannot be inferred that the hypothesis is transferable to the present data, as all between-domain correlations even between similar school subjects decreases with age.

But in fact, the original version of Marsh & Ayotte's hypothesis states that strong between-domain relations for self-concept should increase, stay the same or only slightly decrease with age but that weak between-domain relations should decline with age. Considering this version of Marsh & Ayotte's differential distinctiveness hypothesis, it can be said to be transferable to the present data, or transposed on emotions. And this, since all average between-domain correlations for 8th graders are relatively weak including correlations between similar school subjects, and that they get even weaker when observed for the 11th graders.

5 General Discussion and Conclusions

5.1 Discussion on Major Findings

There is a fine thread of logic that stretches through the three different studies binding them together that may be difficult to grasp for the untrained eye. The goal of this general discussion is making this thread thick enough it is easy to follow. Thus the main results stemming from all three studies, namely the interview study, the memory-based reported questionnaire study and the online reported experience sampling study will be discussed in a deeper analysis. First of all, findings on the characteristics of domains will be examined (see Chapter 2), followed by a debate to help decide if data collected actually support or reject the three aforementioned hypotheses (see Chapter 3.1 or 4.1). This will be done in light of the observations made in Study II and Study III (see Chapter 3 and 4). Later on, differences in these observations will be studied against the background of differences of assessment methods used to collect data for these two studies. In a conclusion, key findings of all results above will be weaved together and observed as a whole, in an attempt to find explanations and relations between these findings. This will help us understand the strength and limitations of this dissertation, as well as its theoretical and practical implications as will be described subsequently.

5.1.1 Characteristics of Domains

This dissertation focuses on the domain specificity of achievement emotions. Thus, it was central to describe the domains, and especially to know how to discriminate between them. Though the whys and the wherefores of such a move have been discussed already (see Chapter 1) coming back onto the characterization of domains was deemed essential. Indeed, in the first study – an interview study – students characterized school subjects, using an extensive list of characteristics that were later organized in categories of characteristics. This investigation, as an explorative study, was a first approach to the subject in hand and aimed at getting an overview of how students perceive school subjects, as well as to prepare the following questionnaire study. The study asked which set of characteris-

tics specifically describes one given domain, or school subject. As a consequence these questions assessed particularly the image associated to specific school subjects.

As already mentioned (see Chapter 2.4) this study showed that only a limited number of aspects related to school subjects were actually used by students to describe these. Coming back shortly on conclusions developed in this part, though the named categories were numerous, they could be classified into a handful of aspect orientated “super-categories”.

Categories one, two, seven and thirteen, pertaining to “amount of subject matter”, “variety in lesson content”, “topic relation” and “current topics”, could be summarized as the topics related “super-category”. It is obvious that all named categories are related in one way or the other to aspects of topics discussed in school subjects, their amount and their variety, how they are logically linked to one another or to the students’ environment. It is only sensible to put these categories together.

Similarly, categories three, four and six, pertaining to “importance of the grade”, “subject matter important” and “everyday usefulness”, all revolving around an aspect of subjects´ importance, can be regrouped into importance related “super-category”. In much the same way, categories five, eight and ten, pertaining to “effort”, “level of difficulty” and “perceived learning requirements”, were put in accessibility related “super-categories”. And this, since all these categories consider perceived requirements not only to access to a particular school subject, but also to succeed in it.

Likewise, categories nine, eleven and twelve, describing “class discussion”, “illustration” and “right solution”, were regrouped in a teaching method or lesson design related “super-category”. The amount of class discussion and illustration though perceived as more probable in certain type of school subjects than others are also related to how teachers design their lessons. Though students might have thought of the characteristic of “right solution” – if one unique solution or many acceptable solutions are usual in the subject – as nothing more than a characterization of the subject itself, it is in fact an attribute that is closely dependent to teaching methods or more precisely to examination methods. It could be perceived that in scientific subjects questions have unique possible answers whereas in more literary subjects solutions might be more debatable. But in fact, exams in mathematics for example can be designed to take into account numerous demonstra-

tion methods to explain a given equality relation, or a German test can be a multiple choice test with only one right answer possible. This in effect shows how uniqueness of solutions, or perceived uniqueness of solutions, can be related to how teachers design their lessons or examinations.

There are aspects which have not been assessed by the interview study but which have been used in other previous studies. These non-assessed attributes consist of, for example the popularity of the domains or the interest for the subject (compare Bachmair, 1969; Jenkins & Nelson, 2005). A probable reason why these attributes have not developed out of the interview study is associated with the wording of the survey-questions. Indeed, the study explicitly and specifically asked for characterizations of school subjects or how the school subjects are perceived. This had in turn a few number of consequences, one of which was, for example that direct judgments of the domains, as the popularity, were left out in the answers given by students. In fact, phrasing the questions also excluded other categories which could have been considered, like the influence of teachers. Another reason that could be alleged for missing categories, like the perceived gender connotation and how great the students assess the opportunity for self-realization in the subject, deemed important in other studies (see Kessels & Hannover, 2006, 2007), could be the lack of general awareness of these type of characterizations among students.

A high number of school subjects both main and minor subjects were observed in the interviews. In total seven school subjects were taken into account, with four of them, namely mathematics, physics, English, and German, being regarded separately. The categories developed were based on all subjects. This has broadened the scope of use of the categories not only for the description of the four selected school subjects but also for other subjects. Looking at how frequently categories were specifically mentioned, one can easily observe that frequencies calculated using answers for all seven subjects (mathematics, physics, German, English, biology, history and music) are practically the same compared to frequencies calculated using only the four subjects. This helps conclude that the categories listed are a solid tool to characterize domains since they are not specifically dependent on the considered school subjects. The strength of this categorization system

is also demonstrated by its high adequacy to the collected data as observed by a high Fleiss' kappa index value (see Chapter 2.2.2).

However in comparing these two classes of frequencies, one striking difference stands out. If one considers the four subjects only the two most mentioned categories are "level of difficulty" and "illustration", but if all seven subjects are observed the category "variety in lesson content" can be added to this list of most mentioned categories. What this finding implies is that compared to the four selected subjects, minor subjects are more characterized by their lesson's content variety.

5.1.2 Domain Specificity of Achievement Emotions

As announced earlier, after considering the characteristics of domains it was important to check whether data collected for Study II and III actually supported hypotheses made throughout this dissertation. This sub-section examines the validity of the domain-specific organization of achievement emotions hypothesis.

In truth, both quantitative studies, the memory-based questionnaire study (Study II) and the online reported experience sampling study (Study III) assumed, in accordance with the first hypothesis, the domain specificity of achievement emotions. Or more particularly, that the relations of the single emotions enjoyment, pride, anxiety, anger and boredom across the domains of mathematics, physics, German and English would be rather weak. The results of both investigations confirm this hypothesis and were found to be in agreement with conclusions of former studies (e. g. Goetz et al., 2006a, 2006b, 2007, and 2009).

It was further assumed that the correlations between similar subjects are stronger than across dissimilar subjects, and also that the strongest relationship could be found between emotions and achievement scores related to the same domain. These assumptions were confirmed by Study II but could not be verified by the experience sampling study. Indeed, it has been observed in the memory-based study that even across two domains intuitively perceived as similar, such as English and German, the emotional experiences were quite dissimilar. The interrelation between same emotions experienced in two similar subjects is stronger, compared to correlations between these specific same emotions

experienced in two dissimilar subjects. But also the relations between emotions and achievement scores pertaining to the same domain were found to be the strongest relations observed. Zooming out for a broader look, it could be said that the relations between positive emotions (enjoyment, pride) and achievement were positive and that negative relations were found between negative emotions (anxiety, anger, boredom) and achievement.

This last comment stands for both memory-based and online report studies, or Study II and III. In contrast, from the online reports it could not be concluded that there is a stronger similarity of the emotional experience of the quantitative subjects (mathematics and physics) and neither for the verbal subjects (German and English). This difference by its own account is an interesting remark that will further be analyzed and interpreted in the conclusions.

5.1.3 Age-related Differences of Domain Specificity

Continuing in our exercise of verification of validity of the initial hypotheses, this subsection considers the validity of Hypothesis 2. This hypothesis assumes that domain-specific organization of achievement emotions is much more pronounced in a sample composed of older students than compared to a sample composed of younger students. Or to put it simply that the domain specificity increases with age.

This second hypothesis was formulated for both studies (Study II and III) and can be explained as thus, that the between-domain relations of academic emotions are stronger in younger compared to older students. And, referring to Marsh and Ayotte's (2003) differential distinctiveness hypothesis, it was also hypothesized that the between-domain relations should decrease with age for emotional experience in disparate domains, and that there should be less, or even no decrease at all in related ones.

The results of both studies show that there is even more domain specificity in emotional experience as perceived by 11^{th} graders compared to 8^{th} graders (see Chapter 3.4.2 and 4.3.3). And the data of both studies clearly suggest that Marsh and Ayotte's differential distinctiveness hypothesis (2003) is transferrable to academic emotions. By which results

of both studies are in line with results of previous studies on the subject in hand, for example by Goetz et al. (2007).

5.1.4 Relations of Domain Characteristics to Domain Specificity of Achievement Emotions

Finally, investigation on the validity of the third hypothesis was conducted. This hypothesis comes as a positive answer to the question of whether school subjects that were described using similar categories were also similar in the student's emotional experience.

For this purpose, only the results of the second study were examined as it was the only quantitative study using categories in hand. Expressly speaking, the between-domain relations of the categories were computed as well as the between-domain relations of the emotions. In parallel, the average correlations over all the categories and subsequently over all emotions were calculated. The results were summarized and presented in tables 7, 8, 10 and 11 above. They show that the relations of the categories, between mathematics and physics on one hand, and German and English on the other, are stronger than across one quantitative and one verbal subject. An analog observation was made on the achievement emotions. These conclusions were made using the data of the memory-based reports of the second study.

It is quite striking to see how average correlations of the categories in Table 11 closely match average correlations of the emotions in Table 8. Also the significance test which compared every single value of Table 7 with the corresponding correlation in Table 10 showed that nearly every pair are significantly, or at least closely significantly, similar. Findings made on the age differences were similar for the categories as for the emotions, since the correlations of the categories between the different subjects are mostly weaker for the 11^{th} graders sample. Only the relations between mathematics and physics are stronger in the older sample compared to the younger. The exact same observation can be made concerning correlations of the emotions between the four different subjects.

5.2 Comparison of Assessment Methods

The study of domain specificity of achievement emotions was done, first in Study II with the help of memory-based data on emotions, and in Study III with the support of data collected by means of the experience sampling method. Results stemming from these studies were analyzed in previous sub-chapters separately. The present sub-chapter tries to compare different measurement methods and their conclusions with regard to domain specificity of achievement emotions.

In examining the differences of the two assessment methods – questionnaire and experience sampling method – the first step was the confrontation of the descriptive statistics resulting from data obtained by these two different methods. First of all, it is quite obvious that in general mean values of the emotional experience as calculated for the data collected in the experiment sampling study, are smaller than mean values as calculated from data collected with the help of a memory-based questionnaire. In particular, pride, anxiety and anger had remarkably small averages. Similar observations can be made for the values of standard deviation. Even more remarkable, mean values of one specific emotion in different domains are closer to each other in Study III as compared to Study II. These observations are clearly visible in figures 8, 9, 11 and 12.

In both studies, the between-domain relations were relatively weak in general (see Table 7 and 13), as discussed in previous sub-sections. In Study III (experience sampling method) the correlations of the emotions of the domains were weaker than in Study II (questionnaire study). This suggests that, even though it is proven that there is domain specificity of achievement emotions, in both studies; it is indeed higher for data collected during actual lessons. Interestingly, the proximity in the value of the single emotions between all four subjects projects first impression of a rather low or very low domain specificity. However, the correlations suggest that the online report study (Study III) shows an even stronger domain specificity of emotional experience, than Study II.

The between-domain correlations in Table 6, even more so their average correlations in Table 7 show that the data assessed with the questionnaire study suggest that similar subjects have stronger relations. These observations do not correspond with the results of the experience sampling study where the correlations of emotions between similar

subjects are not particularly stronger than between dissimilar subjects (see Table 13 and 14). It is quite interesting to look at the reasons of these findings which will be done in the conclusions of the present dissertation.

However, these differences between conclusions of Study II and III have a consequence on the study of the transferability of Marsh and Ayotte´s differential distinctiveness hypothesis. Though it has been observed in previous sub-chapters that data collected for studies I and II confirm the transferability of Marsh and Ayotte´s differential distinctiveness hypothesis (2003), they do not quite confirm the same version of this hypothesis. In fact, throughout this dissertation a slightly different version of Marsh and Ayotte´s hypothesis was adopted, stating that between-domain relations for emotional experience, in between similar school subjects, should, increase, stay the same or only slightly decrease with age. In contrary, between-domain relations for emotional experience, in between dissimilar school subjects, should only reduce with age. This version was considered acceptable as another hypothesis adopted throughout this dissertation was that similar subjects have stronger relations than dissimilar subjects. Whereas the original version of Marsh and Ayotte´s hypothesis can be read as follows: Strong between-domain relations should increase, stay the same or only slightly decrease with age but that weak between-domain relations should decline with age. Though Study II confirms the transferability of both versions to data on emotions, Study III only confirms the transferability of the original version to data on emotions.

By comparing findings about age-related differences, it can be also said that data generated from both assessment methods show that the between-domain relations of the emotional experience of younger students are, almost always, stronger than the ones of older students (see Figure 7, 8, 13 and 14). In Study III the disparities in results between the two different age groups are even stronger than in Study II. Though the same conclusions can be drawn by observing the homogeneity index *h*(d), whose analysis concludes that there is a stronger domain-specific organization of the emotions for 11^{th} graders, compared to 8^{th} graders.

In fact, the homogeneity index also indicates differences in domain specificity based on emotion type. Thus it can be said that in Study II, enjoyment showed the weakest, and

anxiety the strongest, between-domain relations for both samples – even if they were all relatively weak. These findings stand in line with former studies (Goetz et al., 2006a; Pekrun, Goetz, Titz, & Perry, 2002b) but diverge from the results of the study from Goetz et al. (2007). In line with the latter study, Study III also contradicts the findings of Study II, because its *h*(d) values differ much compared to the questionnaire study data. In the experience sampling study in contrary, anxiety and enjoyment showed the weakest between-domain relations.

A number of remarks which were made here are interesting enough to analyze further, as the conclusions explore the reasons for these particular results.

5.3 Conclusions

The first study, an interview study, assessed the image students have of different subjects using a number of characteristics for and from which later on categories were found. This turned out to be quite useful as school subjects which were intuitively perceived as more similar were actually described more similarly with the help of categories, and also the emotional experience of these was found to be more similar. The latter could be demonstrated in the second study, which assessed – as a paper questionnaire study – memory-based measurements of the emotions and the categories. The above observation can be said to be a first reason of domain specificity of achievement emotion. Indeed similarities and dissimilarities in the image projected by school subjects, translate into similarities and dissimilarities in the emotional experience of the students. In particular, this means that mathematics and physics on the one hand, as well as German and English on the other hand, are described more similarly and also generate more similar emotions than other pair wise combinations of these four subjects.

In addition Study II and Study III both confirm the hypothesized domain specificity of achievement emotions, as well as the fact that the domain-specific structure of the emotions gets stronger with age. However in the experience sampling study the correlations are even weaker, suggesting even stronger domain specificity. A start of explanation to this observation could be that current assessment of emotions is predominantly dependent on the context, whereas for non-current assessment students' feelings about specific

school subjects are amended by the same feelings for different school subjects. This, in turn, creates an underlining thread that links the different domains. But more interestingly, in contrast to data from the questionnaire study the measurement from the experience sampling study did not show stronger similarity in the emotional experience between the two quantitative subjects on the one hand, or between the two verbal subjects on the other hand.

This result is reasonable since on the one side categories are in a way strongly dependent of the students' beliefs about school subjects as they assess perceived images. And on the other side the memory-based report of the emotions reflecting more the image of how students think they feel, is even more dependent of students' beliefs. That creates a relation, a sort of link between categories, and memory-based report of emotions. This probably explains the fact that data in Study II confirms Hypothesis 3.

In contrast the third study measured current emotions during classes with the experience sampling method, an online report method which circumvents the aforementioned problem.

In their work, Robinson and Clore (2002) suggest that, "because experiential knowledge is more or less exclusive to online report of emotions, people often say different things about their emotions when actually experiencing them then when they are not actually experiencing them" (Robinson & Clore, 2002, p.935). Also the previous remarks could be read in light of their findings.

But if the image students have of domains and their emotional experience are dependent on beliefs, as observed above, beliefs themselves are dependent on a number of concepts. In fact, in general beliefs are formed not only on personal experience, but also on clichés and stereotypes in the environment. Hence it can be concluded that students' perceptions of school subjects are dependent in turn on clichés and stereotypes, but also their emotional experience. That is in fact why subjects perceived as similar are experienced in a similar manner for Study II.

To round off this dissertation, it could be said looking back at its introduction that in the definition of domain specificity of achievement emotions, the most important indicators

were relatively weak between-domain relations. An additional indicator was the fact that similar subjects have stronger relations than dissimilar subjects. The first indicator was found to be reliable for both Study II and III. The second definition of domain specificity however still fitting to memory-based reports should be reconsidered in the light of the online reported measurements. For the latter type of measurement, even though domain specificity was proved, the assumption that similar subjects have stronger relations than dissimilar subjects was found to be fallacious.

This has a consequence on the interpretation of Marsh and Ayotte´s differential distinctiveness hypothesis transposed on emotions. In fact, the slightly different version of Marsh and Ayotte´s hypothesis adopted throughout this dissertation seemed to be an equivalent to the original version as similar subjects were expected to have stronger relations than dissimilar subjects. The aforementioned fallacy proved this assumption wrong rendering the adopted version unusable on data of Study III. Hence, only the original version of Marsh and Ayotte´s differential distinctiveness hypothesis could be considered for Study III.

5.4 Strength and Limitations

As in any other scientific research, this dissertation has a number of strong points as it has unavoidable limitations.

Looking for strength, it could be said that first and outmost of such is the multi-method approach using interviews, memory-based questionnaires, as well as online-reporting experience sampling method, applied in this entire work. This is indeed one of the strength of this dissertation. As it tries to tackle the issue of domain specificity using a number of measurement methods instead of just one, it gives the reader a much more comprehensive picture of the subject at hand.

It has already been said that Study I is an explorative study in the research of categories of characteristics that could describe domains. This study has been innovative in its considering of a much wider range of domains and in the latitude it gave students to find characteristics that matched closely their specific description of a school subject.

Though Study II and III used two different samples to make their observations in both studies, the two same age groups and the same type of German school had been the focus of these investigations. This has the undeniable advantage of facilitating or even of giving a sense to the comparison of these two studies with each other. However, this strength can be regarded as a limitation as it considered only one type of school, and limited range in age. In addition, the questionnaire study (Study II) assessed 1709 students from 74 classes, dispatched in 11 different high schools in Germany, and the experience sampling study (Study III) 121 students in 41 classes from 7 high schools. This shows the high scope of broadness of the samples examined considering the corresponding measurement methods.

Despite all its strength, this study presents some limitations that are indeed worth mentioning. Beyond the shortage of specific scientific literature to compare with Study I, because of its very innovative nature, this study considered only a limited type of categories in a sense that they took into account only a narrow set of aspects. It is also true that most, if not all variables were assessed with single item questions; this being an undeniable limitation to the description of variables. Even though all three presented investigations were based on comparable samples as mentioned, all of them being based on the study of students in a single type of school in two similar age groups, these samples are different. Therefore conclusions drawn by direct comparisons of these three parts should be treated with caution.

These strengths and limitations have their own implications on the conclusions of this dissertation. But beyond that they also have deeper implications on past, but specifically on future investigations into domain specificity of achievement emotions which will be described down below.

5.5 Implications

Beyond mere dry theoretical observations, these conclusions have not only implications for further educational psychology research, but also concrete real life implication which will be discussed in detail down below.

5.5.1 Implications for Further Research

The implications for subsequent research in educational psychology are essentially logical consequences of the limitations described in the previous sub-chapter.

One part of these implications is tied to the method of data collection. Indeed as observed before, the different parts of this dissertation were based upon the analysis, and the conclusions, made on data collected on different samples of students. Thus it could be of interest to follow the logic that spreads throughout this dissertation to gather and analyze data only on a single permanent sample, revealing maybe a more direct logical connection between the conclusions of the different parts. In addition detailing the interview studies, by the use of electronic diaries for example, could also broaden the scope of research in the investigation on the origins of subjective beliefs about school subjects. This type of addition to research of the present type could bring a considerable amount of additional information that could be useful.

Another part of these implications is related to the investigation methods per say. While characteristics of domains taken into account in these studies are limited in their diversity, it could be interesting to enlarge the list of categories, and to consider a number of other interesting characteristics, like the popularity, interest, and gender connotations for example. It has also been noted, that in order to deepen the understanding about the interaction between subjective beliefs about domains, and the domain specificity of achievement emotions, subsequent studies should link these two themes and investigate their relations. One axis of research in this sense could be first to discriminate between groups of students who perceive stronger similarity between, mathematics and physics, and German and English, on one hand, and on the other hand those who do not. Then later to look in the first group if emotional experience in similar subjects are also more similar, compared to the second group. One way of establishing such differentiations between students would be to assess intensity of perception by utilizing self-reporting scales, after which students are grouped depending on results of these assessments.

Any study that in any way would link these two subjects, namely theories about subjective beliefs about school subjects and domain specificity of achievement emotions, would shed a better understanding on both subjects.

5.5.2 Implications for Practice

As implications for subsequent research were a logical consequence of the limitations, implications for practice are a consequence of observations made throughout this dissertation.

First of all, students' emotions should not be treated as domain-generalized constructs but rather as a domain-specific experience. In reality only few or even no conclusions at all could be drawn on students´ emotions in one school subject from the observations of their emotional experience in other subjects. This implies that, for parents and teachers especially, it is important to know that it is not possible to infer from the emotion of a student in one subject to the same emotion in another subject. Raising teachers´ awareness on domain specificity is central to improving school instruction, since this could help them avoid having the wrong opinion on students´ emotional experience in their subjects. This in turn could help teachers enhance the pertinence of their interventions.

As Goetz et al. (2006a) remark in their work, results also imply that emotion-related counseling and intervention programs should be designed in a domain-specific manner in order to be fully effective. Although this work does not exclude that domain-general programs may have an outcome on domain-specific emotions.

Second, teachers should be mindful of the fact that the beliefs about school subjects matter to students´ understanding of them and especially to students' beliefs of how they feel in the different domains. As there are differences between students´ real emotional experience during lessons and what they perceive of them, what they perceive being closely linked to their images of specific subjects. Any emotion-related counseling program should encompass these two aspects and so should educational instruction in general.

Hence teachers should be aware that in order to create conducive learning environment, they should not only focus on imparting knowledge but also on conveying a positive image of their subject. Indeed it has been pointed out that the perception students have of school subjects is linked to how they understand it and experience it emotionally. Any attempt at creating an emotionally sound instruction method should therefore also focus on building the image of specific school subjects on specific issues. Looking at the work of Kessel and Hannover (2007), this would mean for example working on undermining the negative connotations associated to science lessons to raise their attractiveness. Kessel and Hannover proposed solutions in changing the image of low self-realization and the gender connotation of physics lessons "by changing instructions, classroom activities, or social interactions during physics lessons" (p. 294), but more particularly by reinforcing the image of women contributors to the physics domain in lessons.

In general, perceived emotional experience in school subjects is connected to students´ perceived images of these school subjects, which in turn is connected to how students´ understand and describe them. Thus it is by knowing how students describe and view specific school subjects that one can hope to improve the image and the emotional experience of specific school subjects. This dissertation found out that aspects dealing with the diversity of subject matter treated in lessons and their perceived importance were central in the formation of students´ images of school subjects; so were aspects concerning effort needed to put into succeeding particular school subjects and their perceived level of difficulty in general. This would indicate that by diversifying subject matter discussed during class and raising awareness of their importance in real-life, teachers should be able to raise the attractiveness of their subjects and to create an environment favorable to learning.

Appendix

Table 16: *Frequency Table of the Interview Study*

		Grade 8					Grade 11				
		M	P	G	E	**Σ**	M	P	G	E	**Σ**
1	amount of subject matters	1	0	1	3	**5**	1	2	3	1	**7**
2	variety in lesson	1	4	10	4	**19**	5	1	3	1	**10**
3	importance of the grade	0	1	0	1	**2**	2	0	1	0	**3**
4	importance of the subject	1	0	1	2	**4**	4	1	4	6	**15**
5	effort	5	3	4	5	**17**	5	2	0	2	**9**
6	transferability on everyday life	1	2	0	1	**4**	5	5	2	9	**21**
7	interrelation between topics	1	0	0	0	**1**	3	1	0	0	**4**
8	level of difficulty	9	8	7	10	**34**	12	5	1	2	**20**
9	exchange of views	1	1	2	8	**12**	0	1	2	4	**7**
10	requirements	4	1	2	6	**13**	8	1	1	6	**16**
11	illustration	11	17	2	0	**30**	9	13	1	0	**23**
12	right solution	0	0	0	0	**0**	3	2	3	0	**8**
13	actual topics	0	0	0	0	**0**	0	1	1	3	**5**
	total	35	37	29	40	**141**	57	35	22	34	**148**

Figure 13: Mean Values of the Categories for Each of the Four Subjects; Grade 8.

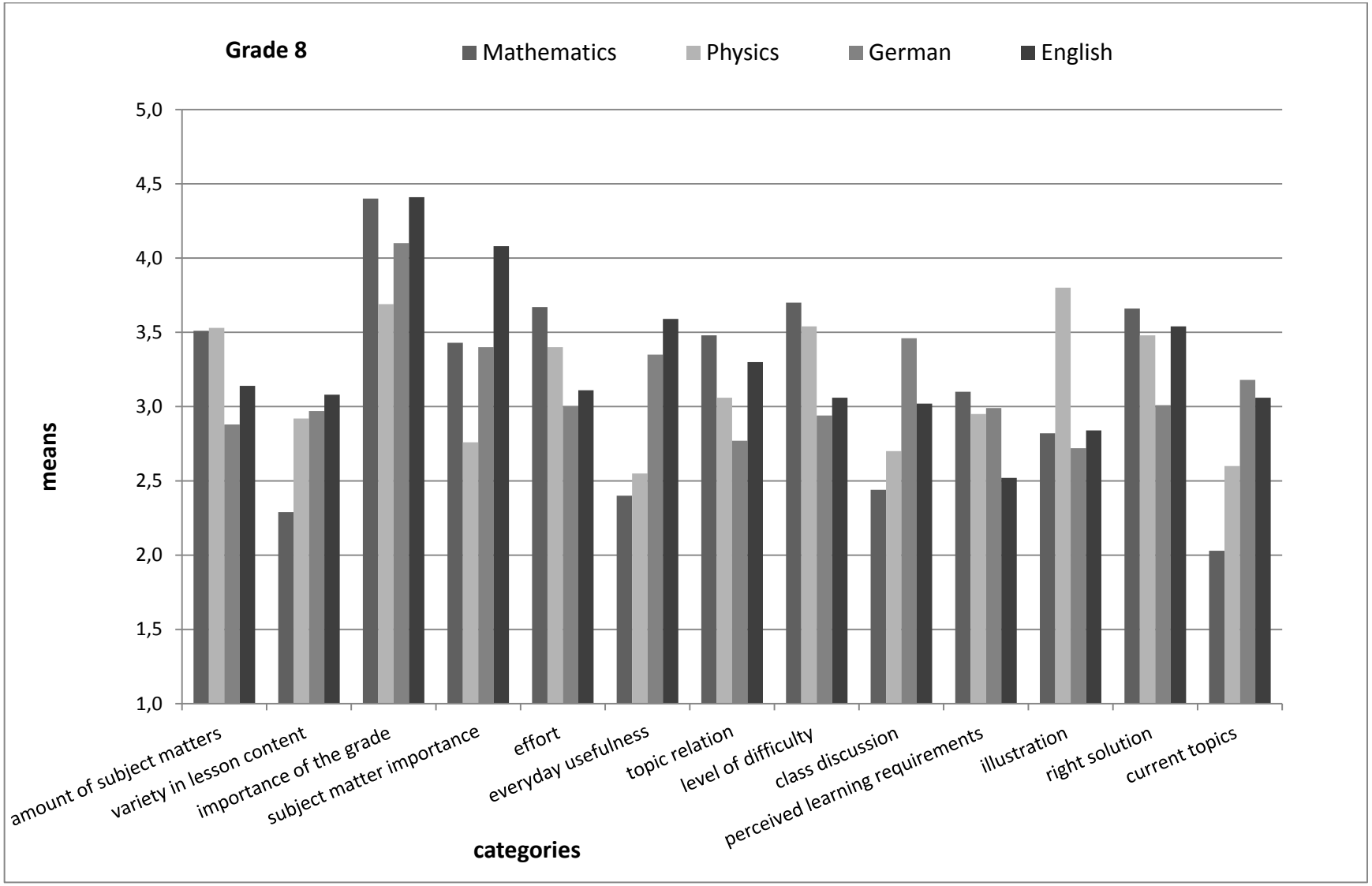
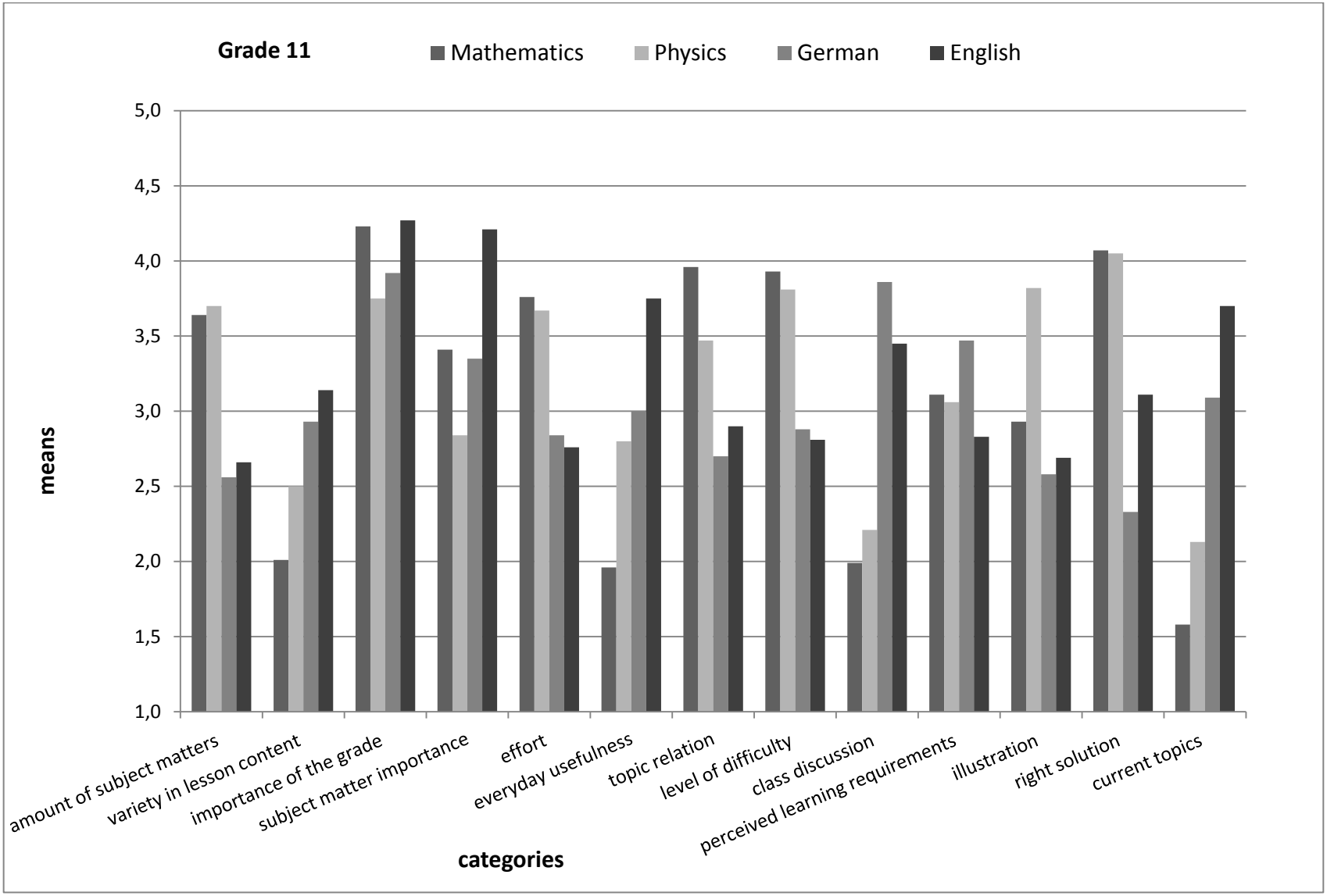


Figure 14: Mean Values of the Categories for Each of the Four Subjects; Grade 11.



Index of Tables

Index of Figures

References

Bachmair, G. (1969). *Einstellungen von Schülern zum Lehrer und zum Unterrichtsfach.* Dissertation, Erlangen-Nürnberg.

Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task value, and achievement goals. *Journal of Educational Psychology, 93*(1), 23-34.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1-40.

Cattell, R. B., & Scheier, I. H. (1961). *The meaning and measurement of neuroticism and anxiety*. New York: Ronald Press.

Chen, P. Y., & Popovich, P. M. (2002). *Correlation: Parametric and nonparametric measures*. (Sage Universtity Paper Series on Quantitative Applications in the Social Sciences, series no. 07-139). Thousand Oaks, CA: Sage.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *Journal of Nervous and Mental Diseases, 9*, 526-536.

Efklides, A., & Volet, S. (Eds.). (2005). *Feelings and emotions in the learning process. Special issue* (Vol. 15).

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*(5), 478-493.

Gieseke, W. (2007). *Lebenslanges Lernen und Emotionen. Wirkungen von Emotionen auf Bildungsprozesse aus beziehungstheoretischer Perspektive.* Bielefeld: Bertelsmann.

Goetz, T., Cronjäger, H., Frenzel, A. C., Lüdtke, O., & Hall, N. C. (2010). Academic self-concept and emotion relations: Domain specificity and age effects. *Contemporary Educational Psychology, 35*, 44-58.

Goetz, T., Frenzel, A. C., Pekrun, R., & Hall, N. C. (2006a). The domain specificity of academic emotional experiences. *Journal of Experimental Education, 75*(1), 5-29.

Goetz, T., Frenzel, A. C., Pekrun, R., Hall, N. C., & Lüdtke, O. (2007). Between- and within-domain relations of students' academic emotions. *Journal of Educational Psychology, 99*(4), 715-733.

Goetz, T., Hall, N. C., Frenzel, A. C., & Pekrun, R. (2006). A hierarchical conceptualization of enjoyment in students. *Learning and Instruction, 16*, 323-338.

Goetz, T., Pekrun, R., Hall, N. C., & Haag, L. (2006b). Academic emotions from a social-cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*(2), 289-308.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA: Sage Publications.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*(1), 33-46.

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Jäger-Flor, D., & Jäger, R. S. (2008). *Bildungsbarometer zum Thema „Mathematik" 1/2008. Ergebnisse, Bewertungen und Perspektiven.* Landau: Verlag Empirische Pädagogik.

Jenkins, E. W., & Nelson, N. W. (2005). Important but not for me: students' attitudes towards secondary school science in England. *Reseach in Science & Technological Education, 23*(1), 44-57.

Kessels, U., & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology, 74*(4), 761-780.

Kessels, U., & Hannover, B. (2007). How the image of math and science affects the development of academic interests. In M. Prenzel (Ed.), *Studies on the educational*

quality of schools. The final report on the DFG Priority Programme (pp. 283-297). Münster, Germany: Waxmann.

Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction, 12*(4), 383-409.

Landis, R. J., & Koch, G. K. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159–174.

Linnenbrink, E. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review, 18*(4), 307-314.

Marsh, H. W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology, 95*(4), 687-706.

Marsh, H. W., & Yeung, A. S. (1996). The distinctiveness of affects in specific school subjects: An application of confirmatory factor analysis with the National Educational Londitudinal Study of 1988. *American Educational Research Journal, 33*(3), 665-689.

Mayring, P. (2003). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (Vol. 8., first volume 1983). Weinheim: Beltz Verlag.

Merzyn, G. (2008). *Naturwissenschaften, Mathematik, Technik – immer unbeliebter?* Hohengehren: Schneider Verlag.

Möller, J., & Köller, O. (2001). Dimensional comparisons: An experimental approach to the internal/external frame of reference model. *Journal of Educational Psychology, 93*, 826-835.

Muthén, L. K., & Muthén, B. O. (1998-2008). *Mplus User's Guide. Fifth Edition.* Los Angeles, CA: Muthén & Muthén.

Nagy, G., Köller, O., & Heckhausen, J. (2005). Der Übergang von der Schule in die berufliche Erstausbildung. Wer die Sorgen scheut, wird von ihnen ereilt. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie, 37* (3), 156-167.

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development* (pp. 143–163). Oxford: Elsevier.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341.

Pekrun, R., Goetz, T., & Frenzel, A. C. (2005). *Academic Emotions Questionnaire - Mathematics (AEQ-M) - User´s manual*. München: University of Munich: Department of Psychology.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in student's self-regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist, 37*, 91 - 106.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2 ed.). Thousendoaks: Sage.

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: Psychometric data. *Journal of Counseling Psychology, 19*(6), 551-554.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*(6), 934-960.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion. Theory, methods, research*: Oxford University Press.

Schutz, P. A., & Pekrun, R. (2007). *Emotion in education*. San Diego: Academic Press.

Seipp, B., & Schwarzer, C. (1991). Angst und Leistung – Eine Meta-Analyse empirischer Befunde. *Zeitschrift für Pädagogische Psychologie, 5*(2), 85-97.

Spielberger, C. D. (1972). Anxiety as an emotional state. In C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research* (Vol. 1, pp. 23-49). New York: Academic Press.

Stevens, R., Wineburg, S., Herrenkohl, L. R., & Bell, P. (2005). Comparative Understanding of School Subject: Past, Present, and Future. *Review of Educational Research, 75*(2), 125-157.

Stodolsky, S. S., & Grossman, P. L. (1995). The impact of subject matter on curricular activity: An analysis of five academic subjects. *American Educational Research Journal, 32*(2), 227-249.

Weiss, H. M., Beal, D. J., Lucy, S. L., & MacDermid, S. M. (2004). Constructing EMA studies with PMAT: The Purdue Momentary Assessment Tool user's manual Retrieved from http://www.mfri.purdue.edu/pmat

Zeidner, M. (1998). *Test anxiety. The state of art*. New York: Plenum.