STUDENT WORKLOAD AND WELL-BEING IN THE INTERNATIONAL BACCALAUREATE DIPLOMA PROGRAMME

Final report of the 2016-2018 Mainstage study

Harry BG Ganzeboom, Department of Sociology, VU University Amsterdam

Ineke Nagel, Department of Sociology, VU University Amsterdam

Heike Schröder, IB Research, The Hague

Contributing research assistants:

Elske Stolte

Maaike Walravens

Sanne Bos

Final version, submitted May 12, 2020

Acknowledgements

International Baccalaureate
Baccalauréat International
Bachillerato Internacional

Table of Contents

EXECUTIVE SUMMARY

Contents

EXECUTIVE SUMMARY

The purpose of the study was to investigate students' complaints about workload and stress in the Diploma Programme worldwide in order to diagnose the causes and effects of these problems, with the ultimate goal of addressing them.

Overview of the study

- This report documents analyses and findings of the **Diploma Programme Student Workload and Well-being Study 2016-2018**, conducted by IB Research in collaboration with researchers from the Department of Sociology of VU University Amsterdam.

- **Research design:** The project sampled 258 schools in 50 countries in North America, South America, Europe, Africa and East-Asia. 98 schools from 38 countries agreed to cooperate and provided contact data for 4,854 DP students (including both full and course students). Students were invited to respond to online questionnaires three times during their two DP years, in November of Year 1, June of Year 1, and June of Year 2, after the exam period. In total, 3,565 students participated in the online survey at least once, 1,063 in all three waves. 2,699 students provided usable information on their workload in at least one wave. After concluding the surveys among students, the data set was complemented with data on participation in the DP exams. In the middle of the research period (May of Year 1), the project also collected information (via an online questionnaire) among DP coordinators worldwide. Coordinators in 1,393 DP schools provided information. For of the 98 schools in the student sample, 74 coordinators responded to the survey.

- **Questions asked to the students** dealt with:
  - Participation in the DP: how many and which subjects taken at higher (HL) and standard level (SL)
  - Participation in other programmes
  - Students' assessment of the difficulty and time burden of the different elements of the IB-DP (objective workload)
  - Time spent on school work (objective workload)
  - Time budget outside of school: extracurricular school and leisure time activities
  - Workload and level of difficulty as experienced (subjective workload)
  - Stress and well-being, using four standard scales from the psychological literature
  - Support and stressors from personal background, parents, pedagogues, peers and school policies ('P-factors')
  - Students' exam participation, exam results, aspirations and university admission (outcomes)
  - Students' evaluation of different ways of addressing workload concerns

- **Questions asked to the DP coordinators** mainly dealt with:
  - Their perception of the level of difficulty and time burden of the different elements of the DP
  - Their perception of the students' subjective workload and stress levels
  - School policies and practices regarding collaborative planning
  - Academic and social-emotional support services for students
  - School climate

- o   Suggestions for making the DP workload more manageable for students.

- **Core model and concepts**. Key to the design of the project and the organization of this report is the core model as displayed in **Figure 1**. We conceive of the process under observation as being separable into three layers. Our target variables are levels of stress[1] and student outcomes. Our basic research question is how these measures of stress are related to workload as experienced by the students on the one hand and to student outcomes, such as exam passes and grades on the other. Workload is separated into two dimensions: 'objective' workload and 'subjective' workload.
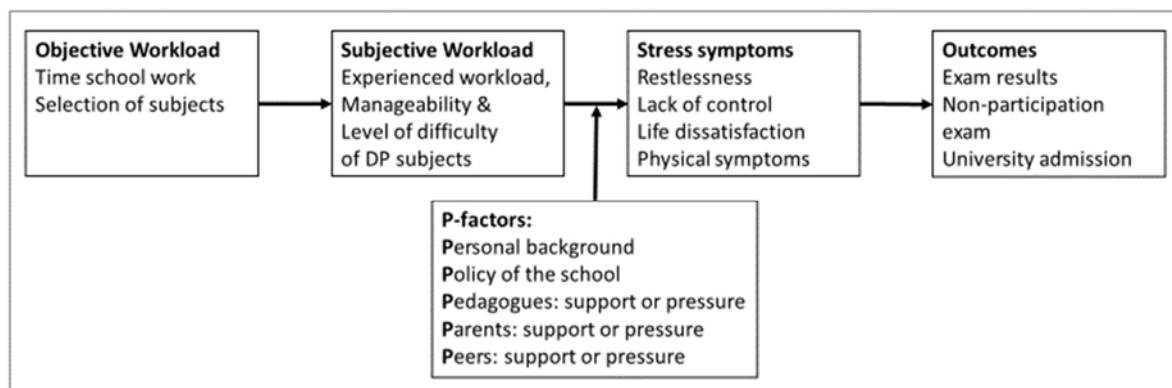
Figure 1: Causal model of the student objective and subjective workload and stress

- The two ingredients of **objective workload** are students' reports on participation in the programme (which subjects have they chosen?) weighted by an average evaluation of the burden of these subjects as obtained from the DP students and coordinators collectively, and the time students spend on them. While those three ingredients are still based on personal reports, they are 'objective' in the sense that no personal evaluation is part of an individual student's objective workload measure.

> The **objective workload** comprised the following measurements:
> - curriculum instruction hours spent on the DP
> - additional instruction time (homework private lessons) spent on the DP
> - the level of difficulty of students' selection of DP subjects

- **Subjective workload** refers to the degree the students report the Diploma Programme as '(un)manageable', and is measured dynamically, i.e. at each of the three waves. Subjective workload statements refer to cognitive pressures ('how difficult') and time pressures the programme presents to the students who have chosen a certain individual curriculum.

---

[1] We use "stress" to any report about lack of well-being by the students which do not directly refer to the programme or school. Stress is about how you feel; subjective workload is about how you feel about the programme.

> **The subjective workload** is an individual self-reported state and comprised these measurements:
> - experienced weight of DP workload
> - experienced level of challenge of the DP
> - experienced manageability of DP workload
> - academic capacity to cope with academic challenge of DP

- The **core research questions** of the study are:
    - **To what extent does the DP workload cause stress in DP students?**
    - **What are potential factors that impact students' development of stress due to the DP workload?**

  A pivotal aspect here is causality, ie, does workload *generate* stress? We address these questions via a **reciprocal causation model**, that controls for reversed and spurious causation.

- **Supplementary research questions**. Our supplementary questions focus on how workload and stress are related to school-specific and personal factors. We distinguish the following factors **(P-factors): P**ersonal (personal characteristics, ie, gender, academic ability, language ability, socio-economic background [SES]), **P**olicy (school well-being policies and services), **P**edagogues (teachers), **P**arents (support or pressure) and **P**eers (support or pressure). The theory behind this approach is that some of these factors, such as support from parents, teachers, peers, or professionals at school protects students from developing stress symptoms, while others, such as parents, teachers or peers increase their risk of developing stress. To answer these questions, we combine school characteristics and student characteristics in one model that is known in the literature as **multi-level model**. This model allows us to quantify and compare the impact of these individual factors. We also test for so-called moderator effects.

Findings

Objective workload

- DP subjects vary greatly in the workload in terms of time burden and difficulty. DP students and DP coordinators largely agree that **subjects from the Mathematics and Sciences groups are heaviest**. Students with **higher academic abilities** and whose parents have a **higher socio-economic status** tend to select **heavier subjects**. By contrast, **girls** tend to select subjects with a somewhat lower workload than boys.

- Students spend on average **47 hours on DP school work** per week, about **26 hours on regular classes**, and about **20 hours on homework**, and, **1.5 hours on additional lessons.** Full DP students spend on average between six and 10 hours a day on the DP. According to the international PISA study (OECD, 2016), the average total study time for students in OECD countries (of 15 years of age) is 44.1 hours a week. 13% of students in OECD countries spend at least 60 hours in total studying; this is 40% in China and the United Arab Emirates and less than 5% in Finland and Germany. By comparison, For the hardest working **25% of DP students spend at least 58 hours a week on the DP**.

- Students who had better **academic achievements prior to the DP, are more proficient in the language of instruction or study a heavier selection of subjects spend more time on DP homework and additional lessons**. **Girls**, moreover, **spend more time** on their school work **than boys**.

Subjective workload and stress among DP students

- **Stress levels**: DP students report high levels of restlessness, lack of control, life dissatisfaction, and physical health problems, on average consistently beyond the midpoint of our scales, with 16% of the students reporting extreme values at wave 1, 19% at wave 2 and 35% at wave 3. These high levels of stress symptoms are confirmed by students' reactions to the open questions, that give vivid descriptions of students finding themselves struggling to meet the demands of the DP.

- **Stress development**: The results indicate a moderate increase of stress levels between wave 1 (beginning of Year 1) and wave 2 (beginning of Year 2), followed with a sharp increase at wave 3 (end of Year 2 after the exams) – so a continuous rise of stress during the DP.

- While students appear to struggle with their DP workload during their entire DP career, there seems to be little variation between waves. Reports on **subjective workload are stable across waves.** For both subjective workload and stress levels there is much consistency *within* students *between* waves. To a large extent **both stress and subjective workload appear to be stable personal characteristics**, which means that some students always report high levels while others consistently report low levels of stress and subjective workload.

- Our three-wave panel design allows us to separate these stable characteristics from the within-person changes, and to determine the direction of influences between workload and stress. Unsurprisingly, causality flows in both directions: increased levels of workload cause increasing levels of stress, while increased stress levels cause workload complaints. The effects are fairly strong and statistically significant in both directions. This is an important finding, primarily because it rules out the possible criticism that the workload-stress relationship is predominantly reversed, with stressed students blaming the DP workload.

Risk and support factors (P-factors)

- Some **personal characteristics** of the students are related to their stress levels. A first consistent finding in our data (as well as in the literature), is that **girls report higher stress levels than boys**. At the end of the DP girls also report a heavier subjective workload than boys. Students' who reported better **academic or language abilities** prior to IB, report **lower levels of subjective workload and stress**.

- **Parents** can impact students in two ways. While **parental pressure increases students' subjective workload and stress**, parents also have a protective influence when they show an interest in and are involved in their children's lives in general and at school. The results show strong effects for parental support. Students whose parents are involved experience lower levels of **subjective workload and stress**.

A factor that can hardly be overestimated are **peers**. Similar to parents, peers strongly impact students' subjective workload and stress. Students with **positive peer relations** show substantially lower levels of subjective workload and stress. The same again holds for **pedagogues** (teachers). **Teacher support, positive student teacher relations** as well as **high teaching quality** are clearly associated with students' experienced workload and stress.

- While teachers and peers can protect students from developing stress, they may also have the opposite effect when they put pressure on students. **School pressure**, consisting of **pressure from teachers, peers and in DP classes** in general, noticeably augments students' subjective workload and stress levels and thus poses a risk to their well-being. However, this relationship is less strong than the positive association with teacher and peer support.

- As for schools' well-being **policy**, effects are few and far between. One measure that appears to have a measurable, albeit small effect are **entrance requirements**. Such requirements, surprisingly, appear to lead to students' working longer hours.

- Furthermore, there appears to be an **impact of the number of internal assessments** (school averages of IAs) on the time investment, subjective workload and stress levels in Year 2. Time investment, subjective workload and stress levels all increase when the school's number of assessments increase in a given period.

- The stress development process differs between girls and boys. Girls are significantly **more vulnerable to workload pressures than boys,** ie, given the same level of subjective workload, they are more prone to developing stress symptoms. By contrast, school policies geared to support and protect students have not been found to have measurable effects.

- No other factors have been found to explain differences in students' vulnerability of workload pressures to become stress.

- **To sum up**, a clear picture seems to emerge. It appears to be the quality of social relationships, be it with parents, peers or teachers that is pivotal. In a school context, this may be summarized as **school social climate**. When students are under a lot of pressure from their parents, peers or teachers or lack support, they are more likely to struggle with their workload and develop stress symptoms. When students get on well with and are supported by their parents, peers and teachers, both at school and in their private lives, they can mostly handle the workload and are less prone to develop stress symptoms.

Student outcomes

- **Non-participation in DP exam**: According to the exam data, some 19.3% of the sampled students did not participate in the final exams, or at least did not receive any grade. The reason for this may be coincidental, such as students being ill on the day of the exams, but likely many if not most of the no-shows left the programme. Not taking the final exams, turned out to some extent to be predictable from its antecedents at wave 1 and wave 2. We found that **students who report the DP as "unmanageable" at wave 1 or 2 and/or severely stressed, are twice as likely not to take the exam.**

- **Exam passes:** Students who say to be struggling with their workload in wave 1 and 2, are also slightly less likely to pass their exams when they do participate. Previous workload and stress problems, however, have a much more spectacular effect on students' final grades. They decrease the **grade point average** to a comparable degree it is increased by students' academic ability.

- DP students have very high **ambitions.** When they start the programme, more than 80% say they strive for a master's degree or higher, with 74% wanting to attend a **very good or top-level university**. This percentage drops to 61% by the time they reach the finish line. As many as **18% of DP students**, nonetheless, **say they have been admitted to one of the top 50 universities from the worldwide Shanghai ranking**. There is a strong relationship between the grades obtained at the exams and the likelihood of being admitted to a higher-level university. Subjective workload and stress do not affect students' ambitions. However much they may have struggled along the way, DP students who persist, have very high odds of realizing their dreams.

School practice and evaluations of different proposals to address workload issues

- A first important way of addressing workload issues is prevention. Roughly two-third of DP coordinators believe that **guiding students in their programme and/or subject choice** as well as training them in **organizational skills** could be effective ways of averting problems. Another important aspect of prevention are the **planning and coordination of assessments and homework restriction of homework**. Schools hold **collaborative planning meetings** most frequently with subject groups. Planning meetings across all subject groups or across the two DP years are much less popular. While schools excel in **coordinating deadlines**, with roughly three quarters of the schools engaging in at least coordinating the deadlines for IAS and Core components, **homework restrictions** are much less popular and practiced by less than half the schools. When presented with different options of how to make the workload more manageable, **homework-free days or periods** do, however, **rank high on the students' list** right after having a clear schedule of assessment deadlines, which tops it.

- A second way of addressing workload problems is the **provision of study and psychological support for students**. Well over 90% of DP schools provide some form of academic counselling and study facilities, which tend to be well used by students. Almost all schools provide some kind of **professional health or well-being service**, such as a school nurse or counsellor to help students when they run into problems. These latter services are used by about 40% of the students.

- A third and last way of addressing students' workload is specific to the **exam workload**. While well over 70% of the schools **coordinate deadlines for IAs and Core components**, the coordination of school-based assessments is much less practiced by schools. In a minority of schools, **non-exam assessments are brought forward** and finished either in year 1 or by November of year 2 in order to ease the burden during the peak period in year 2. This measure is valued by about 40% of the students.

- While no clear preference emerges regarding the best way to regulate exams and other assessments, both **DP coordinators and students believe that having (more) IB-approved subject or online resources is the most important way of making the DP workload more manageable.** Interestingly, even students, on average, rank the provision of resources (slightly) higher even than a reduction of content in DP subjects, which comes third.

- Only roughly a quarter of schools practice some kind of **homework restriction**, be it in limiting the amount of homework or in homework free days or periods. Homework restrictions do, however, score highly among students as possible remedies for workload issues.

Summary of the findings

Like many students in the 21st century who do high stake, academically challenging programmes, DP students have high educational aspirations and work hard. Like their counter-parts in other programmes, for many DP students complying with programme demands comes with feeling of psychological discomfort, with a substantial minority (15-25%) experiencing extreme levels of stress. Despite the stress, the vast majority of DP students succeed in realizing their goals. They pass their final exams and proceed to the university of their choice. The health and well-being of students are, however, of utmost importance.

- **On average, students find the DP workload challenging and hard to manage; they work long hours to fulfil the requirements.** Students appear to have more problems with the amount of work than with the academic requirements of the programme.

- **The workload of students' individual subject selections may differ considerably.** DP subjects are uneven in the level of difficulty and time investment they require: the Sciences and Maths subject groups have the heaviest workloads, even for SL.

- **Stress and anxiety are only weakly related to the DP workload itself.** Levels of school-related stress and anxiety are high among DP students, but mainly indirectly related to the DP workload**.** Stress levels rise the more students *perceive* the DP workload as unmanageable.

- **Perceived workload and stress appear to mainly be subjective individual student realities.** Two main variations stand out. Girls work more hours than boys and report higher levels of subjective workload and stress. Students who were academically less successful prior to the DP or less proficient in the language of instruction struggle more with their workload and are more stressed.

- **Perceived workload and stress fluctuate over the course of the DP and are related to high concentrations of assessments.** Stress in particular builds up and peaks in the second trimester of year 2. Accordingly, students' perceived workload and stress peak in that period.

- **Long term and consistent stress may affect students' learning outcomes.** Overall, students who report high levels of workload and stress early in the DP are more likely not to finish the programme. Stress decreases students' grade point average and may lead to exam failure.

- **Parents, teachers and peers play a major role in preventing stress by being supportive or increasing it by adding pressure.** Students' relationships at home and even more so at school are pivotal in how they cope with the DP workload. Parents have a large impact on students' wellbeing. Teachers have an even larger impact on students' wellbeing, both at individual and aggregate level and peers have the largest impact.

- **DP schools are already doing a lot to address workload, stress and anxiety by means of various policies and practices.** While these measures may help individual students, on the whole, they appear to have little effect as yet at school level.

Policy considerations

While DP-students are not alone in working long hours, everything should be done to avoid unnecessary burdens and help ensure that all students can keep a healthy balance between school and other activities. Based on the extensive analyses undertaken for this study, below we propose policy considerations aimed at making the DP workload more manageable for students and reduce workload-related stress.

How the IB may contribute to making the DP workload more manageable

The IB may consider articulating a comprehensive wellbeing framework/policy and initiate associated actions to support school actors in shaping their priorities around workload challenges. The actions listed below are in part derived from suggestions provided by DP students, coordinators and heads of school who participated in the study.

- **Provide IB-approved (online) subject resources for students.** Both DP students and coordinators think that IB-approved (online) student subject resources (eg, text books or study guides) would help in making the workload more manageable.

- **Provide resources for teachers and school leaders for a better understanding of the impact of the workload on student wellbeing.** Research briefs and evidence-based recommendations may help school actors in better understanding the impact of the workload on students' lives. Policy and practice reviews may also inspire school actors to develop effective solutions in supporting their students being happy and academically successful. Providing appropriate wellbeing tools for teachers, students and parents could help them better understand school-related stress and manage it appropriately.

- **Provide Professional Development modules for teachers** to sensitize them to the role of teachers and peers in the workload-stress process.

- **Reduce content of DP subjects, in particular in Sciences and Mathematics.** Reduced content in DP subjects would help in making the DP workload more manageable.

- **Consider reducing the number of internal assessments.** Instead of having IAs for all DP subjects IB could decide to reduce them to HL subjects.

- **Consider reducing the number of exam papers and provide increased flexibility for assessments**, so that students can replace an IA or Core component with an exam or vice versa**. Flexibility and equivalence between options need to be ensured.

- **Provide support for schools that are interested in implementing social and emotional skills development programmes**.

How schools may support students in managing their workload

Each school may have particular characteristics and challenges concerning student workload; therefore, school responses to student workload concerns may need to be tailored and unique. However, the following evidence-based policy considerations and promising practices may be a source of inspiration and provide direction to IB schools:

- **Acknowledge the workload challenges in your school.** A first step in supporting students in dealing with their school work is to identify realities and perception of all school actors about school work and related stress and anxiety. Gathering as much data as possible at the school level regarding workload will inform the best interventions in your context. Measuring the impact of perceived workload on stress in the DP compared to other relevant data at the national and international level may also help in understanding the particularities of student workload in your school.

- **Make wellbeing a priority in your school.** Based on the data collected in your context determine how much time schools can realistically demand from their students, taking into account the necessity of "down time" and sufficient sleep in helping students maintain their mental and physical health and paying careful attention to the amount of time teachers need to maintain the rigor of the programme. Research has shown that an integrated, comprehensive school wellbeing policy is more likely to be effective than ad-hoc measures in response to individual students being overwhelmed with stress and anxiety. Likewise, embedding wellbeing into the curriculum, practices and school life seems to be more effective than implementing wellbeing as an "add-on feature".

- **Develop an evidence-based workload policy, as part of your integrative school wellbeing policy**. A clear and explicit workload policy will ensure that students lead full and balanced, healthy lives. A good practice is also to regularly review and adjust the workload policy to any changes that might occur within your school**.** There are various creative ways that schools may implement as a part of a sound homework policy:
  o **Homework-free days or periods could help students achieve a workload balance**. There are examples of education systems showing that heavy homework is crucial to obtain good learning outcomes, while experiencing the joy and meaningfulness of learning is equally important. There is no homework in Finland and yet Finnish students are among the highest performers in the world. According to the OECD, students in Finland have the least amount of outside work and homework of all students in the world.
  o **Provide social and emotional skills development programmes that address particular challenges in your school.** Many studies have shown that students' well-being encompasses a set of skills that can be intentionally and explicitly cultivated and developed through education and with appropriate support. Students with high levels of social-emotional skills may navigate more wisely in striking a balance between school workload and other activities.
  o **Implement metacognitive strategies for teaching and learning around difficult learning and assessment tasks.** Research in the area of metacognition has shown that formative feedback and assessment, as well as sharing and deconstructing fears and stress before a difficult task in a safe and trustful environment may considerably reduce the level of students' anxiety and lead to better learning outcomes.
  o **Develop peer support activities for stress-related issues in your school.** The current study has highlighted the major role of peers in overcoming stress and anxiety. Students may be more open to share their fears with their peers than with adult authorities. Preparing

volunteer students to become support buddies for their classmates is largely acknowledged as an evidence-based practice easy to implement as a part of a wider wellbeing policy.

- o **Explore practices supporting growth mindset and resilience in school** and build a culture of trust that all students can be successful, no matter their background, gender or language skills.
- o **Strategically involve parents as legitimate partners in a whole school wellbeing and workload policy**.

- **Identify and monitor students who are at risk, such as girls, students with lower grades, students with lower language abilities, at an early stage.** Develop an early warning system for school-related stress and anxiety and monitor the objective workload by using specific class planning tools. Early interventions and preventive measures are usually more effective than complex approaches needed when the stress and anxiety become severe.

- **Make sense of the individual differences between students in managing the workload in your school**. Promote a culturally-appropriate gender awareness conversation in your school and try to deconstruct possible social biases in your context, using evidence-based resources.

- **Ensure balanced implementation of DP assessments.** Schools could make the planning, spreading and coordination of non-exam assessments a priority, including the communication thereof to students and parents.

Chapter 1: INTRODUCTION

Contents

Chapter 1: INTRODUCTION

Chapter 1 describes the rationale behind and the design of the IB Diploma Programme Student Workload study. It starts with an explanation of the nature of the Diploma Programme and what has motivated the study. Subsequently, the main research questions are outlined, and the population of schools and students is described that the study aims to represent.

The guiding theories, causal model and research design are discussed in Chapter 2 and 3.

1.1 The IB and the Diploma Programme[1]

- The **International Baccalaureate (IB)** is a non-profit educational foundation, motivated by its mission "to develop inquiring, knowledgeable and caring young people who help create a better and more peaceful world through its programmes of education that promote intercultural understanding and respect". Founded in 1968, the IB currently works with more than 5,100 schools in 157 countries to develop and offer four programmes to over one million students aged 3 to 19 years: the Primary Years Programme (PYP), the Middle Years Programme (MYP), the Career-related Programme (CP) and the Diploma Programme (DP), which is the subject of the present study.

- The **DP is an academically challenging programme for students aged 16-19** and is as of September 2019 offered by 3,421 schools in 157 countries worldwide**.** The programme aims to develop students who:
  - have excellent breadth and depth of knowledge
  - flourish physically, intellectually, emotionally and ethically
  - study at least two languages
  - excel in traditional academic subjects
  - explore the nature of knowledge through the programme's unique theory of knowledge course

- The DP consists **of three core components and six subject groups**. The three DP core components are: theory of knowledge (TOK), the extended essay (EE), and creativity, activity, service (CAS). TOK resembles an epistemology course, CAS asks students to engage in experiences and projects that foster creativity, physical activity and service to their communities, and the EE is a 4,000-word essay on a research question related to a one or more DP subjects. The six subject groups are: studies in language and literature, language acquisition, individuals and societies, sciences, mathematics, and the arts, all of which cover academic subjects traditionally part of pre-university curricula worldwide. Within each of the six groups, students choose one subject, at either standard level (SL) or higher level (HL). SL courses require a minimum of 150 class hours; HL courses a minimum of 240 class hours. **Each student must complete at least three subjects at higher level**.

---

[1] The information in this paragraph is abridged from "What is an IB Education" (2017), at www.ibo.org.

- The DP can be conceived of as a **programme of upper secondary education that culminates in standardised final exams delivered worldwide and aims to give students credit recognised in the process of university admission worldwide**. The final exams of the DP take place at the end of Year 2 of the programme (May for schools in the northern hemisphere; November for southern hemisphere schools). Schools may choose to conduct up to two exams for SL courses in Year 1, which are then called **anticipated** subjects.

- **The DP is an internationally-oriented programme that is taught and assessed in English, Spanish or French** (a handful of individual courses are offered in other languages). English is the language of instruction for most DP schools. However, while in many countries the DP is primarily conceived of as international (ie, English-language) education for non-native speakers, a considerable amount of IB-education takes place in English speaking countries (in particular the US, the UK, Canada and Australia). Likewise, the programme is exclusively taught in Spanish in Spanish speaking countries, and in French in French speaking countries (primarily French-speaking Canada).

- While the DP is described by the IB as "academically rigorous and challenging" (IB, 2014), **there are no formal entry requirements** to the DP. IB implements no selection at entry and only charges a fee for participating schools[2]. However, many schools that have adopted the DP as (part of) their curriculum, are private schools and may impose entry requirements and/or academic ability testing (particularly for students about to begin the DP), as well as additional fees. Moreover, state-funded schools may use a screening or selection process to determine which students ultimately enrol in the DP or require students and their families to contribute to the cost of the final exams.

1.2 Motivation for the study: Concerns about workload and student well-being

- **There is widely shared concern in the IB community that the strenuous Diploma Programme may harm student well-being.** Anecdotal evidence, in particular from student blogs (see for instance; https://www.ibsurvival.com/topic/28186-preventing-and-managing-stress/), testifies that some students experience the programme as too demanding, at least at times. Many DP schools have expressed the same concern and have in fact implemented policies and services to detect stress and remedy problems among their students.

- **It is frequently speculated that the high demands of the IB Diploma Programme are likely to cause more stress than other less-demanding secondary education programmes.** Small scale (selective or small samples) cross- sectional studies have found that the workload in an academically more rigorous high school curriculum, such as the DP, may put excessive demands on students, exacerbating the pressure on student performance (Suldo, Shaunessy & Hardesty, 2008).

---

[2] This fee has recently been abolished.

- **At the same time much remains unknown about the influence of workload on well-being among DP students.** Knowledge about the prevalence and severity of student well-being problems is limited to mostly cross-sectional studies, from a limited number of schools in one US state (Suldo, Shaunessy-Dedrick; Ferron & Dedrick, 2018). Due to the lack of longitudinal studies we do not know whether and how students' 'stress' is causally related to the Diploma Programme or any of its features or components. Stress may occur among DP students just as frequently as among students in other challenging or not-so-challenging programmes. It may also be the case that stress among DP students is high, but not causally related to the DP or its implementation, because the DP is chosen by students with (too) high aspirations who are particularly vulnerable to stress. What is lacking is a systematic study covering a representative sample of the DP student population as well as a study that investigates the role of different determinants of student stress, in particular those in control of the IB or schools (eg, student workload).

- In 2014, IB's Diploma Review Committee (DRC; an internal advisory body comprising IB staff and experienced IB educators) requested **a large-scale study on the DP student workload to examine the extent and severity of the stress problems among DP students and explore their possible antecedents.** The research was to be aimed at DP world-wide and be representative of schools and students in different countries and language areas.

1.3 Aim of the study and research questions

- **The main aim of the study** is to find out to what extent DP students experience the workload of the programme as overly demanding and how this affects their well-being. The main research questions of the study are:
  - **How manageable is the DP workload and how is it related to DP students' experience of stress?**
  - **What are potential factors that can impact students' experiences of workload and stress in the DP?**

- More specifically, the study addresses the following **research questions** to elaborate on these two main questions:

  - Which **specific elements in the DP** (eg, type or combination of subjects), if any, **increase** students' subjective **workload**? (Chapter 4)
  - Is there a **causal relationship between workload and stress**, or can stress be explained by pre-existing conditions? (Chapter 6)
  - Which **risk and support factors** affect student's workload and well-being? (Chapter 7)
  - Which **factors help students cope** with their subjective workload? (Chapter 7)
  - To what extent do **workload and stress vary across students, schools and the course of the programme?** (Chapter 7)
  - How are the **DP workload and stress related to student outcomes** (eg, exam results, university admission)? (Chapter 8)
  - What do schools do to **help students cope** with potential stress (eg, policies, support services) and how effective are these measures? (Chapter 9)

1.4 Study design and study population

- The study aims to answer these questions by collecting empirical data on workload and stress in a three-wave panel design. This allows for a large and worldwide representative sample of DP students to report on their experiences in the DP in the beginning, middle and end of the two- year curriculum. The panel design (with repeated measures of experiences among the same students) makes it possible to map out changes in workload and stress development during the programme and allows for conclusive answers about the causal relationships between workload and stress. At the same time, we have surveyed DP coordinators in schools worldwide about their opinions on these issues. We have also collected information on school policies regarding workload and well-being and have asked both DP coordinators and students to rate possible ways to make the workload more manageable for students. The combined responses may inform possible practical solutions.

- The IB has no enrolment records of its students. Students are enrolled into the programme by their individual schools, and only enter IB-records when they are registered as exam candidates. Therefore, **sampling of students had to be done through DP coordinators identifying students in their schools pursuing the DP**. **In May 2016 a sample of schools** was selected from the IB school registration system (IB Information System; IBIS) which at the time contained a total of 2,981 DP schools in 146 countries. **The aim was to select approximately 10% of those schools** in order to obtain a sufficiently large sample for the study. **Criteria for inclusion of schools in the sample were that the schools held May exams and had English, Spanish or French as their language of instruction.**

- The initial **sample** consisted of 258 schools (very close to 10%) from 50 countries. Out of the 258 **DP coordinators, 98 (from 36 countries) collaborated** and sent sufficient contact details of altogether **4,854 students.** Due to the adopted sampling procedure, the effective sample of schools is highly **representative with respect to**:
  - Continent
  - Language of instruction
  - Number of DP schools in country (with May schedules)
  - School status (private/state)

  In other words, the sample is diverse on these characteristics which makes it possible to draw general conclusions and to compare results on different characteristics.

- The sampled students were invited to take part in the survey at three points ('waves') in their two-year curriculum. **Three online questionnaires** were sent to the students:
  - Dec 2016   - at the beginning of Year 1 of de DP
  - June 2017   - at the end of Year 1 of the DP
  - May 2018   - after completion of the exams

- In the beginning of Year 2 of the DP, all DP schools were asked to report on their well-being policies in the middle of the panel period via a **school survey**. 2,894 DP coordinators and heads of school

were approached and around **2,200 filled out the survey,** including 119 from 91 schools where the student surveys were distributed as well.

Chapter 2: LITERATURE REVIEW: WELL-BEING AND STRESS AMONG HIGH-SCHOOL STUDENTS

Contents

Chapter 2: LITERATURE REVIEW: WELL-BEING AND STRESS AMONG HIGH-SCHOOL STUDENTS

The aim of this chapter is to study existing literature in order to:
- Find out what is known in the research literature about the influence of student workload on stress development
- Describe relevant factors that influence development of stress symptoms in high school students
- Identify benchmark studies for results from the current study on DP students
- Formulate hypotheses about DP students' experience of workload and stress

To answer these questions, we have mainly focused on the following resources which include samples of high school students worldwide and/or the IB student population specifically:
- The Well-being study report from the Programme for International Student Assessment (PISA, OECD 2017)
- Literature review from Suldo and her collaborators at the University of Southern Florida as summarized in the book *Promoting Student Happiness; Positive Psychology Interventions in Schools* (Suldo, 2016).
- IB related research papers by Suldo and her collaborators.

2.1 Terminology, theory and definitions: Well-being, stress and workload

Well-being

- **Different definitions of well-being** exist in the research literature. Since 2011 the emphasis in research has shifted from a one-dimensional approach to well-being as a general feeling of happiness or life satisfaction to **multidimensional models of well-being** (Seligman, 2011; Suldo, 2016). A multidimensional model is also applied in the important well-being study from the Programme for International Student Assessment (PISA), which examines 15-year old students in 72 countries (OECD, 2017). In the PISA study students' well-being was defined as 'the psychological, cognitive, social and physical qualities that students need to live a happy and fulfilling life' (OECD, 2017, p. 19). **Figure 2.1 shows** the multidimensional model of student well-being , It illustrates how different contextual factors interact and influence the different domains of students well-being. **The PISA study included the psychological, social, cognitive and physical domains and examined four areas of students' life:**
    o   their performance in school
    o   their relationships with peers and teachers
    o   their home life
    o   how they spend their time outside of school
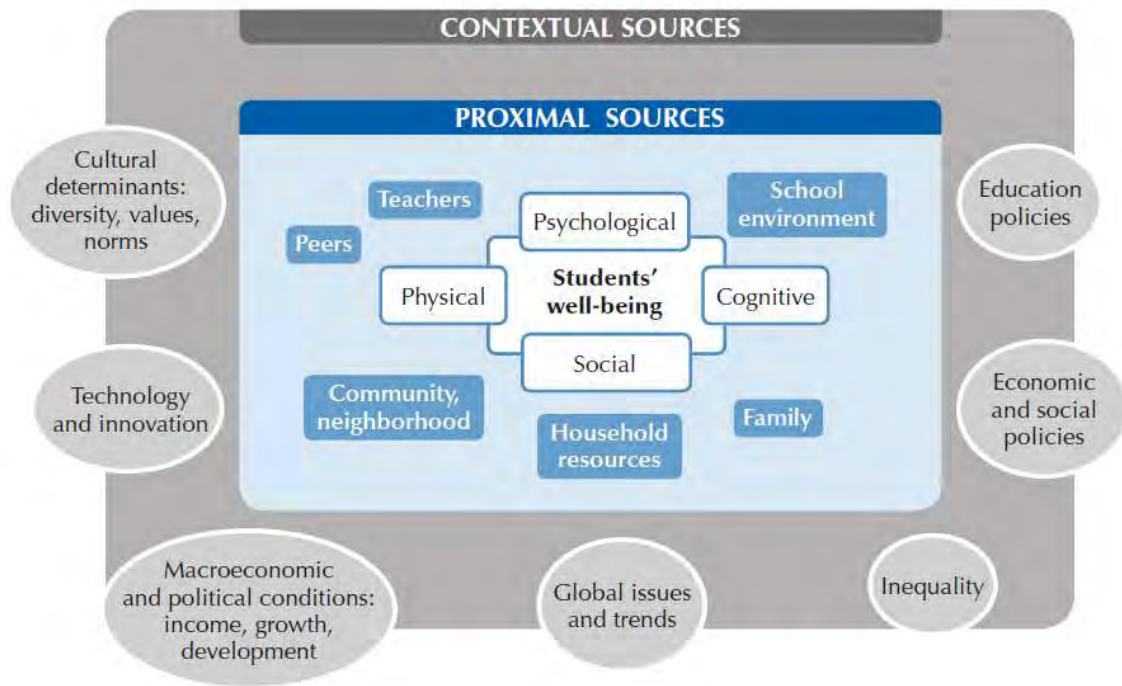
Figure 2.1 Multidimensional well-being framework from the PISA study (OECD, 2017)

Stress

- Stress too has been defined in different ways. A commonly accepted definition is that **"stress arises when individuals perceive that they cannot adequately cope with the demands being made on them or with threats to their well-being"** (Lazarus, 1966). A certain amount of stress is normal and can help individuals focus on what is important to them, such as having to meet a deadline or pass an exam. Prolonged and severe stress, however, can cause problems. In students stress can lead to ineffective coping and in turn to social and emotional problems such as depression, lower life satisfaction and school dropout (Suldo & Shaunessy-Dedrick, 2013a; Suldo, Storey et al., 2018). In other words, it can be expected that prolonged exposure to demands that exceed a student's resources has a negative effect on their well-being.

- **A useful model by McGrath (1970) distinguishes four stages of stress development (Figure 2.2)**
  - an event / situation or series of events that can create excessive demands called **stressors**
  - the perception (a personal **evaluation** of the situation) of this event as **manageable or not** by an individual based on their situation and resources
  - a **stress response** in the form of reporting **symptoms** of psychological or physical discomfort
  - **stress consequences** such as academic performance or other outcomes which can be negative or positive depending on the effect of the stress response on the person and their actions.

- The current study implements these stages in a conceptual model for DP students, see section 2.3.

Stressor	Evaluation of Stressor	Stress Symptoms	Consequences
Stage 1:	*Stage 2:*	*Stage 3:*	*Stage 4:*
Event(s) or situation(s) that can induce stress	How demanding is the stressor?	Physical or psychological symptoms	Performance and outcomes

Figure 2.2 Stress development stages based on McGrath (1970).

- Individuals can deal with stressors in different ways. Depending on the personal resources the same stressor may lead to little stress for one person while another person may be overwhelmed. In the psychology literature the personal efforts to deal with a stressor are called coping: **Coping is defined as "constantly changing cognitive and behavioral efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person, regardless whether the outcomes of such efforts are positive or negative"** (Suldo, Shaunessy & Hardesty, 2008, p. 274). Coping is part of the above-mentioned definition of stress; the evaluation of the extent to which a person feels able to cope with a stressor is part of their stress response.

Student workload

- The topic of this study is **DP students' workload** which is here regarded as a specific type of **stressor**. In the PISA study, an international study on students' school performance, workload is defined as the time students spend on their schoolwork and measured in hours of studying both in- and outside of formal classes.

- For students in the DP (as well as Advanced Placement (AP) and other IB Programmes) a recently developed **instrument** on academic requirements measures to what extent students experience stressors related to school demands and competing priorities in AP and IB schools (Student Rating of Environmental Stressors Scale [StRESS]; Suldo, Dedrick, Shaunessy-Dedrick, Roth, & Ferron, 2015), which is a comprehensive way of capturing various types of stressors high-school students may have to cope with, both at school and at home. While this measure has inspired the study at hand, we needed a more focused measurement of school- and programme related stressors, in particular students' workload. In Chapters 3, 4 and 5 we discuss how workload is measured in the current study.

2.2 Studies on student workload, stress and well-being

- An overview of pertinent studies is presented in **Table 2.1** (below), which is ordered by what we refer to as **P-factors** in this study **(all beginning with the letter p):**

  o **Policy** of the school regarding well-being;
  o **Pedagogues'** (teachers') support and pressure
  o **Personal** student characteristics, eg, gender, socio-economic status, academic ability, time-management skills)
  o **Parents'** support and pressure

o **Peers'** (classmates') support and pressure.

- The studies , however, did not include only DP students but 9 to 10th grade students in general and/or AP students.

Workload and academic performance

- According to the international PISA study (OECD, 2016), **the average total study time for students in OECD countries (of 15 years of age) is 44.1 hours a week.** 13% of students in OECD countries spend at least 60 hours in total studying; this is 40% in China and the United Arab Emirates and less than 5% in Finland and Germany. **The total number of study hours can be broken down into regular classes, homework and additional instruction. PISA students report an average of 26.9 hours in regular classes,** ranging from 24.2 in Finland to 31.9 hours in Chile. **On average 15-year old students spend 17.2 hours on after-school study time (homework and additional instruction)**, ranging from 11 in Germany to almost 30 in the United Arab Emirates.

- The PISA study also examined the relation between study hours and academic performance (OECD, 2017). **More time spent in regular science lessons was related to better performance while more time in additional study after school is related to worse performance** (OECD, 2016). Overall, students who spend an average of more than 60 hours on schoolwork achieve worse results than students who study a maximum of 40 hours a week, after accounting for students' social economic status. However, there are exceptions. For China, Korea and Taiwan studying 60 hours or more per week is related to large improvements in performance.

Workload and stress

- Pressure to excel in school is reported by students as being one of the top stressors in daily life (Feld & Shusterman, 2015). The PISA study also found that **anxiety about schoolwork is one of the sources of stress most often cited by high school students** (OECD, 2017). On average, 59% of students across OECD countries report that they often worry that taking a test will be difficult. 66% of students report that they worry about poor grades and 55% of students say they experience high levels of anxiety before a test even when they are well prepared. **Students who report anxiety about schoolwork, homework and tests perform slightly worse at school.** These results indicate that around the world many 15-year old students struggle with their workload.

- In general, **studying long hours was not related to lower life satisfaction in the PISA study.** However, **there are differences between countries.** In Austria, Greece, Japan, Korea and Peru, students who study longer hours reported life satisfaction at least 0.2 point above average. In the Czech Republic, Macao (China), and the United Kingdom studying long hours was related to below average life satisfaction.
- Students who experience **high levels of school-related anxiety** report lower levels of life satisfaction than students with lower levels of school-related anxiety.

- According to the PISA study, **academic performance has a weak positive influence on life satisfaction** in 15-year old students. Other studies on school performance and life satisfaction also find that students who score higher on school tests are slightly happier (Suldo, 2016).

DP students' workload, stress and well-being

- Is there previous evidence that school work causes stress in IB students? There are studies on stress in IB students which include DP students but are not limited to them. These studies are described first before discussing DP students specifically.

- Suldo, Shaunessy, Thalji, Michalowski & Shaffer (2009) found that the **primary source of stress** reported by **IB-students** from grade 9 to 12 are **academic demands**, while for students in general education, other stressors, such as problematic parent-child relations and peer relations appeared to be more important. Several studies have found higher stress levels in IB students compared to students following general education. IB students already reported higher stress levels than students in general education when entering high school in 9th grade (Suldo & Shaunessy-Dedrick, 2013a).

- **Stress-levels related to academic demands** were also **higher for** grade 9 to 12 **IB students** compared to students in general education (Suldo & Shaunessy-Dedrick, 2013b; Suldo, Shaunessy et al., 2008). While stress levels of IB and AP students were comparably high, their psychological functioning (life satisfaction, psychopathology and social anxiety) was similar or superior to that reported by their peers in general education (Suldo & Shaunessy-Dedrick, 2013b).

- A recent study with a large sample of 2,379 9-12 grade AP and IB students in 20 programmes (10 IB; DP, pre-DP and MYP and 10 AP) in one US state, found that **stress from academic demands was related to reduced life satisfaction, higher levels of school burnout and psychopathology** (Suldo, Shaunessy-Dedrick et al., 2018). This may mean that although on average, despite their high stress levels, the psychological functioning of IB and AP students is similar to students in general education, there might be subgroups of students experiencing psychological problems.

- In order to identify IB and AP students at risk, a study of 304 9th grade students measured emotional and academic risk factors (Suldo, Storey et al., 2018). The **risk factors** were **high levels of perceived stress and low levels of school satisfaction and academic performance**. More than a third (38.5%) of students scored higher than the cut-off score on one of the risk factors.

- Although the studies described above are not limited to DP students, they give good reason to hypothesize that the Diploma Programme is stressful for students as it is a rigorous high school curriculum, likely to put high demands on students (Suldo, Shaunessy-Dedrick, Ferron & Dedrick, 2018; Suldo, Shaunessy et al., 2008). They point out that **additional challenges in the DP** include **high performance expectations** in higher level classes and regarding college or university entrance exams as well as the required **service to the community (CAS)** (Suldo & Shaunessy-Dedrick, 2013b).

- In sum, these studies show that both general high school students and IB students experience workload-related stress, but IB students a bit more. The findings above illustrate how a rigorous curriculum is associated with higher levels of stress in students. However, it remains unclear whether the workload **causes** stress. It could be that students who are more worried about school work self-select into the IB and thus that the increased experienced stress is due to being more sensitive to stress from schoolwork. Moreover, it is important to know which factors increase or decrease stress respectively in order to provide the right support to the right students.

Factors influencing DP students' stress and well-being: The P-factors

- As stated above, for the purposes of this study, we have called the factors that influence students' stress and well-being the **P-factors: personal background, policy of the school, pedagogues, ie, teachers, parents and peers**. In the following, we will provide research results regarding the impact of each of these factors for IB students.

Personal background

- **Motivation for school** is positively related to life satisfaction but this motivation is also related to test anxiety (OECD, 2017). In AP and IB students, motivation for achievement at school, cognitive and emotional engagement and approach/problem-focused coping style with academic demands go together with better mental health (Suldo, Shaunessy-Dedrick et al., 2018).

- On average, students in OECD countries who expect to complete university education are 30% more likely to report high life satisfaction (a score of 9 or 10 out of 10; OECD, 2017).

- Life satisfaction decreases with **age** during adolescence (Suldo, 2016).

- **Gender** is related to adolescents' life satisfaction as 29% of girls and 39% of boys report they are very satisfied with their life (OECD, 2017). Girls report higher levels of schoolwork-related anxiety than boys (OECD, 2017).

- An **immigration background** is usually related to less sense of belonging at school although variations between countries exist (OECD, 2017).

- **Health behaviors** which are related positively to life satisfaction are sufficient sleep, healthy eating, no smoking and engaging in physical activity (Suldo, 2016). Students who reported taking part in some moderate or vigorous physical activity were less likely to report that they feel very anxious about schoolwork or that they feel like an outsider at school (OECD, 2017).

- In AP/IB students **mental health** is almost unrelated to students' gender, grade level, ethnicity or past educational history, but financial problems at home are related to worse mental health in AP/IB students (Suldo, Shaunessy-Dedrick et al., 2018).

- In AP/IB students' **stress from major life events** is related to worse mental health and academic outcomes (Suldo, Shaunessy-Dedrick, et al., 2018).

Policy of the school

- **School climate** and a **feeling of belonging** at school are important factors in the happiness of high school students (Suldo, 2016). AP and IB students who experience more stressors related to academic demands also report higher levels of school burnout and reduced life satisfaction (Suldo, Shaunessy-Dedrick et al., 2018).

Pedagogues

- School satisfaction is influenced more by student-teacher relationships than by peers and family (Jiang, Huebner, & Siddall, 2013). **Happier students** tend to **report positive relations with their teachers** and students report that they feel supported by their teachers in schools where students report higher levels of well-being (OECD, 2017).

Parents

- A **good relationship with parents** who have an **authoritative parenting style** (promoting psychological autonomy, behavioral supervision and monitoring and high level of responsiveness) and show more involvement in their children's live is related to **higher well-being in students** (Suldo, 2016). AP and IB students' parent-child conflict and financial problems at home predict worse mental health while an authoritative parenting style predicts better mental health (Suldo, Shaunessy-Dedrick et al., 2018).

Peers

- **Positive relationships with peers** are related to **better well-being** (Suldo, 2016). Students in IB were happier with their friendships and reported higher level of social support from classmates and less affiliations with delinquent peers than students in AP and general education (Suldo & Shaunessy-Dedrick, 2013b). In AP and IB students who experienced more support from classmates had higher scores on life satisfaction and lower scores on psychopathology but also lower AP/IB exam scores and lower scores on GPA (Suldo, Shaunessy-Dedrick et al., 2018).

2.3 Causal model

The literature reviewed in section 2.2 provides a good overview of the (P) factors that play a role in the relationship between students' workload, their evaluation of it and their well-being. Although there is no specific research on DP students, the studies on students in general education and IB students in grade 9 to 12 provide enough insights to inform a causal model for DP students' stress development.

P-factors previous research has shown to influence stress and well-being in students

Category	Examples
Personal background	• Gender, socio-economic status, age, immigration background • Cognitive ability, engagement at/motivation for school • Academic self-efficacy, study and planning skills, problem focused coping strategy with school demands • Health behaviors Major life events
Policy of the school	• School climate • Feeling of belonging at school
Pedagogues	• Teacher-student relationship • Teacher support
Parents	• Parent-child relationship, parenting style • Parent involvement, parent support • Conflict between parents, financial problems • Mother's education level
Peers	• Peer-relationships, feeling accepted by peers, social struggles with peers • Peer support and pressure in and outside of school

• The theory on stress development and factors that influence this process and its outcomes are summarized in a **causal model** in **Figure 2.3**, consisting of a **theoretical** and a **measurement** part. The model shows that in this study workload is the stressor of interest. In line with stress development theory, a distinction is made between **objective workload (the stressor)** and **subjective workload (the evaluation of the stressor)**. The current study investigates DP students' actual (objective) workload, how DP students evaluate that objective workload (subjective workload) and **how** this **subjective workload impacts** on their **well-being, academic performance** and other related **outcomes**. Stress symptoms are seen as a reflection of reduced well-being.[1]

• The four stages in the **stress development model** are:
  o **Stage 1.** The stressor is schoolwork related to the DP (**objective workload**).

---

[1] We take an empirical rather than an – a priori – theoretical approach to the dimensionality of stress and well-being. Whether stress and well-being are different poles of one dimension, or different dimensions, is in the end to be decided by statistical forms of dimensionality analysis

- o **Stage 2.** Students evaluate their objective workload as more or less demanding and manageable (**subjective workload**).
- o **Stage 3.** The subjective experience of the workload may affect students' well-being and lead to physical or psychological **stress symptoms** in students.
- o **Stage 4.** The stress symptoms may, finally, affect students' **outcomes** such as academic performance (consequences of stress).

- Both **individual student characteristics** and **environmental factors influence the stress development process**. The extent to which the objective workload is evaluated as manageable and the extent to which it causes stress symptoms and problematic outcomes is influenced by environmental factors (**Table 2.1** summarizes these P-factors). In the model these are so-called **moderators** which means that they interact with objective workload, subjective workload and stress symptoms. For example, students with more supportive parents might experience less stress symptoms from the same level of subjective workload as students who have less supportive parents. Parental support then moderates the relationship between subjective workload and stress symptoms.
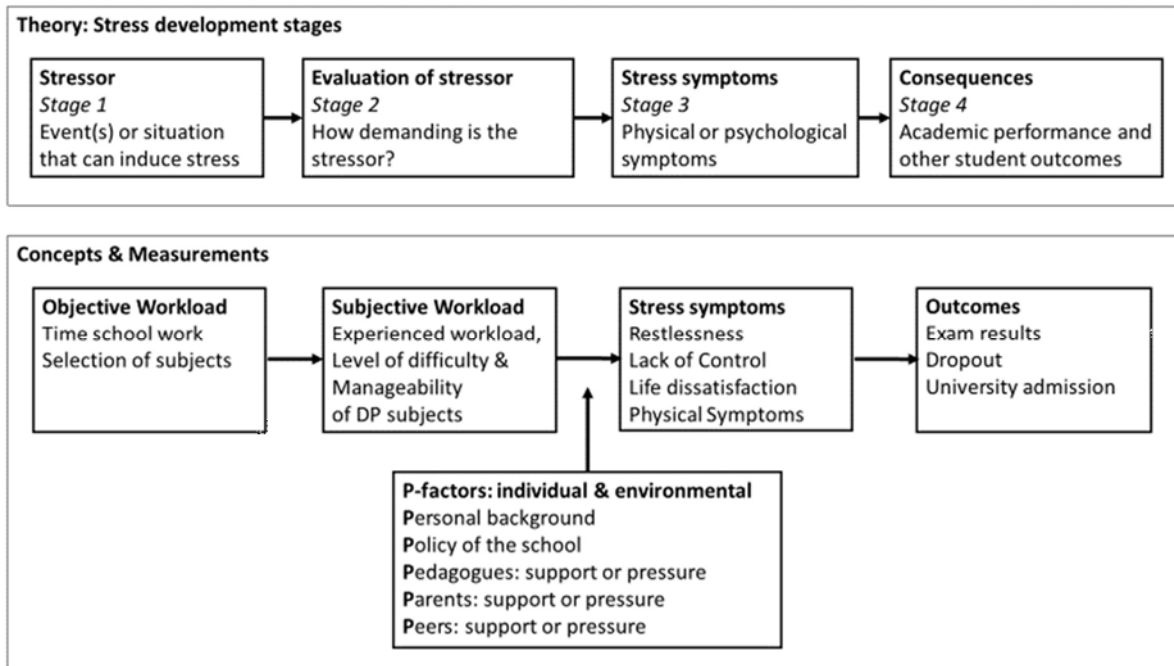

Figure 2.3 Causal model

Table 2.1 Overview of the findings from the literature on schoolwork, stress symptoms and well-being levels.

Study	Sample	Results
Policy of the school		
PISA 2015 study results (Volume II, OECD, 2016)	15-year old students in OECD countries	Mean study time in OECD countries: 44.1 hours a week, 26.9 hours in regular lessons and 17.2 hours of homework and additional instruction. Large variation across countries.
PISA 2015 study results (Volume III, OECD, 2017)	15-year old students in OECD countries	• On average across OECD countries more than half the students worry about school work. About 59% of students reported that they often worry that taking a test will be difficult; 66% reported that they worry about poor grades; 55% of students reported feeling very anxious for a test even if they are well prepared; 37% reported they get very tense when studying; 52% reported that they get nervous when they don't know how to solve a task at school. • Students with higher anxiety scores also report lower life satisfaction and worse performance at school. • On average the number of hours studying does not influence life satisfaction (except in some countries). • School performance has a weak relation to life satisfaction.
Suldo, Shaunessy-Dedrick, Ferron & Dedrick, 2018	2,379 students grade 9-12 enrolled in AP or IB in 20 school programs in one USA state	• Most strongly related to academic outcomes are academic skills in 8th grade. • Stress from academic requirements was related to reduced life satisfaction. • 66.5% of students scored in the positive range on global life satisfaction, while 71% of students reported some symptoms of burnout and 74.7% of DP seniors earned the diploma. • Academic support at school (i.e., preparation for entry to AP/IB, as well as ongoing guidance and tutoring services for AP/IB students) had almost no effect on student mental health or exam results once other factors were taken into account.
Suldo, 2016	Review of literature on young students	• School climate including relationships with people at school is most predictive of students' happiness and personal academic success. • Academic achievement hardly affects students' happiness.

Study	Sample	Results
		• Life satisfaction is linked to cognitive engagement at school and behavioral engagement in the classroom.
Suldo, Shaunessy & Hardesty, 2008	139 IB students and 168 non-IB students in grade 9 to 12, from a high school in a rural south eastern state of the US	Stress-levels related to academic demands are higher for IB students than students in general education.
Suldo & Shaunessy-Dedrick, 2013a	134 freshmen entering three high schools (IB and general education) in 9th grade	IB students reported more stress than students in general education in 9th grade.
Suldo & Shaunessy-Dedrick, 2013b	480 students in grade 9 to 12 from four high schools, comparing students from general education (n=113), to students enrolled in the AP (n=117) and IB programme (n=250).	Higher stress-levels experienced by AP and IB students, but their psychological functioning (life satisfaction, psychopathology and social anxiety) was similar or superior to the mental health indicators reported by their peers in general education.
Suldo, Storey, O'Brennan, Shaunessy-Dedrick, Ferron, Dedrick & Parker, 2018	304 9th grade IB and AP students from two public high schools in a southeastern state of the US	Measured risk factors were perceived stress, school satisfaction and academic performance. More than a third (38.5%) scored higher than the cut-off score on one of the risk factors; 15% scored higher than the cut-off score on stress, 16% on school satisfaction and 20% on academic performance.
Pedagogues		
Jiang, Huebner, & Siddall, 2013	547 middle school students at a US school in a south eastern state (half 7th grade, half 8th grade)	Student-teacher relationships are more important for school satisfaction than is support from peers or family.
PISA 2015 study results (Volume III, OECD, 2017)	15-year old students in OECD countries	• Happy students also report positive teacher-student relationships.

Study	Sample	Results
		• Negative teacher-student relationships are related to lower feeling of belonging at school. • Students in schools with higher than average life satisfaction in their country have much higher teacher support ratings. • Students were less likely to report anxiety if the science teacher provides individual help when they are struggling.
Suldo, Friedrich, White, Farmer, Minch & Michalowski, 2009	401 middle school students from one school in a south eastern state in the US	Most important aspects of teacher support for middle school students are emotional support including feeling like treated fairly and assistance during learning.
Personal background		
Suldo, 2016	Review of literature on young students	• Students who believe they can learn and achieve at school (academic self-efficacy) ten to be happier. • Life satisfaction decreases with age. • Behavioral factors which are related positively to life satisfaction are sufficient sleep, healthy eating, no smoking and engaging in physical activity. Life satisfaction is linked to engagement in the classroom.
Kalak, Lemola, Brand, Holsboer-Trachsler, & Grob, 2014	1,125 students between 10-16 years old from Switzerland and Norway	Getting more hours of sleep predicts greater well-being 6 months later. Subjective well-being did not predict later sleep trends.
PISA 2015 study results (Volume III, OECD, 2017)	15-year old students in OECD countries	• Students who feel they are outsiders at school were three times more likely to report that they are not satisfied with their life. • Children living in poverty have slightly lower well-being scores. Once the basic needs are met, little if any relationship with family income is found. In most countries, students reported less satisfaction with life if they perceive that they are not as wealthy as most of the other students in the school. • Students who expect to complete university education are 30% more likely to report high life satisfaction (a score of 9 or 10 out of 10)

Study	Sample	Results
		• 29% of girls but 39% of boys reported that they are very satisfied with their life. Girls reported more school related anxiety. • An immigration background is usually related to less sense of belonging at school although variations between countries exist. • Students who reported taking part in some moderate or vigorous physical activity were less likely to report that they feel very anxious about schoolwork and that they feel like an outsider at school. • Motivation to achieve at school and life satisfaction reinforce each other. The most motivated students score the equivalent of more than one school year higher in PISA than the least-motivated students. Countries where students are highly motivated to achieve also tend to be the countries where many students feel anxious about a test, even when they are well prepared for it.
Suldo, Shaunessy-Dedrick, Ferron & Dedrick, 2018	2,379 students grade 9-12 enrolled in AP or IB in 20 school programs in one USA state	• Stress from major life events is related to worse mental health and academic outcomes. • Financial problems at home are related to worse mental health in AP/IB students. • Approach/problem-focused coping style with academic demands is related to better mental health in AP/IB students • Academic motivation and cognitive and affective engagement promote mental health outcomes • Better academic outcomes are related to higher socio-economic status (SES). • Small to no association was found between educational background, SES, gender and racial background and mental health.
Suldo, Minch, & Hearon, 2015	624 high school students	Personality characteristics explain almost half of the differences in reported life satisfaction.
	Parents	
Suldo, 2016	Review of literature on young students	• Authoritative parenting style (promoting psychological autonomy, behavioral supervision and monitoring and high level of responsiveness) and good relationships

Study	Sample	Results
		with parents (securely attached, accepted by and open communication) are related to higher well-being. • Divorce has an influence on student's life satisfaction through diminished economic status and mostly through lower levels of harmony between parents. • Frequent and intense parents' arguments have a negative impact on life satisfaction.
Jiang, Huebner, & Siddall, 2013	547 middle school students at a US school in a south eastern state (half 7th grade, half 8th grade)	School related support from parents is the strongest predictor of global life satisfaction
PISA 2015 study results (Volume III, OECD, 2017)	15-year old students in OECD countries	• Students whose parents reported "spending time just talking to my child", "eating the main meal with my child around a table" or "discussing how well my child is doing at school" regularly were between 22% and 39% more likely to report high levels of life satisfaction. • **"Spending time just talking" is the parental activity most frequently and most strongly associated with students' life satisfaction.** • Students whose parents reported "spending time just talking" were one-thirds of a school year ahead in science learning after accounting for socio-economic status. • Girls who perceive that their parents encourage them to be confident in their abilities were less likely to report that they feel tense when they study.
Crede, Wirthwein, McElvany, & Steinmayr, 2015	411 German high school students following the most academically prestigious secondary school track ("Gymnasium").	Students who reported better academic achievement also reported a little bit higher life satisfaction. The relation is stronger for students with a mother with high educational attainment. The educational background of the father did not play a role.
		Peers
PISA 2015 study results (Volume III, OECD, 2017)	15-year old students in OECD countries	On average across OECD countries, students who talk with or meet friends after school reported a level of life satisfaction around 0.3 point higher on the life satisfaction scale

Study	Sample	Results
		(which ranges from 0 to 10). The average score on life satisfaction is 7.3 on a scale from 1 to 10, with large variations between countries.
Suldo, Shaunessy-Dedrick, Ferron & Dedrick, 2018	2,379 students grade 9-12 enrolled in AP or IB in 20 school programs in one USA state	• AP and IB students who experienced more classmate support had higher scores on life satisfaction and lower psychopathology but also lower AP/IB exam scores and lower scores on GPA. • Stress related to social struggles with peers (quarrel with friends or bullying) is related to worse mental health.
Suldo, 2016	Review of literature on young students	• Feeling accepted by peers is related to greater life satisfaction, particularly in cultures with less emphasis on family values. • More positive social acts from peers at school is related to higher life satisfaction and positive affect. • Having at least one friend protects against negative effects of peer rejection and is measurable years later in greater life satisfaction.
Suldo, Gelley, Roth & Bateman, 2015	500 general high school students grade 9 to 11 in the US	Students with positive peer relations are more likely to report better well-being (life satisfaction and positive moods), while students experiencing peer victimization and aggression are more likely to experience more psychopathology (internalizing stress and aggression).

References

Crede, J., Wirthwein, L., McElvany, N., & Steinmayr, R. (2015). Adolescents' academic achievement and life satisfaction: The role of parents' education. *Frontiers in Psychology, 6*, 1–8. http://doi.org/10.3389/fpsyg.2015.00052

Feld, L. D., & Shusterman, A. (2015). Into the pressure cooker: Student stress in college preparatory high schools. *Journal of Adolescence*, *41*, 31–42. http://doi.org/10.1016/j.adolescence.2015.02.003

Jiang, X., Huebner, E. S., & Siddall, J. (2013). A Short-Term Longitudinal Study of Differential Sources of School-Related Social Support and Adolescents' School Satisfaction. *Social Indicators Research*, *114*(3), 1073–1086. http://doi.org/10.1007/s11205-012-0190-x

Kalak, N., Lemola, S., Brand, S., Holsboer-Trachsler, E., & Grob, A. (2014). Sleep duration and subjective psychological well-being in adolescence: A longitudinal study in Switzerland and Norway. *Neuropsychiatric Disease and Treatment, 10*, 1199–1207. http://doi.org/10.2147/NDT.S62533

Lazarus, R. S. (1966). *Psychological stress and the coping process.* New York, NY: McGraw-Hill.

McGrath, J. E. (1970). *Social and psychological factors in stress*. New York, NY: Holt, Rinehart & Winston.

OECD. (2016). *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*. PISA, OECD Publishing, Paris. https://doi.org/10.1787/9789264267510-en

OECD. (2017). *PISA 2015 Results (Volume III): Students' Well-Being*. PISA, OECD Publishing, Paris. https://doi.org/10.1787/9789264273856-en

Seligman, M. E. P. (2011). *Flourish: A Visionary New Understanding of Happiness and Well-being*. New York, NY: Free Press.

Suldo, S. M. (2016). *Promoting Student Happiness: Positive Psychology Interventions in Schools*. New York, NY: Guilford Publications.

Suldo, S. M., Dedrick, R. F., Shaunessy-Dedrick, E., Roth, R. A., & Ferron, J. M. (2015). Development and Initial Validation of the Student Rating of Environmental Stressors Scale: Stressors Faced by Students in Accelerated High School Curricula. *Journal of Psychoeducational Assessment*, *33*(4), 339–356. http://doi.org/10.1177/0734282914552164

Suldo, S. M., Friedrich, A. A., White, T., Farmer, J., Minch, D., & Michalowski, J. (2009). Teacher Support and Adolescents' Subjective Well-Being: A Mixed-Methods Investigation. *School Psychology Review*, *38*(1), 67–85.

Suldo, S. M., Gelley, C.D., Roth, R.A., & Bateman, L.P. (2015). Influence of Peer Social Experiences on Positive And Negative Indicators of Mental Health Among High School Students. *Psychology in the Schools*, 52(5), 431-446. https://doi.org/10.1002/pits.21834

Suldo, S. M., R Minch, D., & Hearon, B. V. (2015). Adolescent Life Satisfaction and Personality Characteristics: Investigating Relationships Using a Five Factor Model. *Journal of Happiness Studies*, *16*(4), 965–983. http://doi.org/10.1007/s10902-014-9544-1

Suldo, S. M., Shaunessy, E., & Hardesty, R. B. (2008). Relationships among stress, coping, and mental health in high-achieving high school students. *Psychology in the Schools*, *45*(4), 273–290. http://doi.org/10.1002/pits.20300

Suldo, S. M., Shaunessy, E., Thalji, A., Michalowski, J., & Shaffer, E. (2009). Sources of stress for students in high school college preparatory and general education programs: Group differences and associations with adjustment. *Adolescence, 176,* 925-948.

Suldo, S. M., & Shaunessy-Dedrick, E. (2013a). Changes in Stress and Psychological Adjustment During the Transition to High School Among Freshmen in an Accelerated Curriculum. *Journal of Advanced Academics*, *24*(3), 195–218. http://doi.org/10.1177/1932202X13496090

Suldo, S. M., & Shaunessy-Dedrick, E. (2013b). The Psychosocial Functioning of High School Students in Academically Rigorous Programs. *Psychology in the Schools*, *50*(8), 823–843. http://doi.org/10.1002/pits.21708

Suldo, S. M., Shaunessy-Dedrick, E., Ferron, J., & Dedrick, R. F. (2018). Predictors of Success Among High School Students in Advanced Placement and International Baccalaureate Programs. *Gifted Child Quarterly*, *62*(4), 350–373. http://doi.org/10.1177/0016986218758443

Suldo, S. M., Storey, E. D., O'Brennan, L. M., Shaunessy-Dedrick, E., Ferron, J. M., Dedrick, R. F, Parker, J. S. (2018). Identifying High School Freshmen with Signs of Emotional or Academic Risk: Screening Methods Appropriate for Students in Accelerated Courses. *School Mental Health*, *62*(4), 350–373. http://doi.org/10.1007/s12310-018-9297-9

Chapter 3: METHODOLOGY

Contents

Chapter 3: METHODOLOGY

The DP Student Workload study is designed as a three-wave panel survey, in which DP students are surveyed online at three points in their two-year DP career. A subsidiary data collection was held among DP coordinators in all DP schools in the beginning of Year 2 of the DP. In this chapter the design and methodology of the study are discussed; we describe the advantages and challenges of this longitudinal, multi-level research design. Also, a description of the sampling process of DP coordinators and characteristics of the schools and students in the sample is provided. Furthermore, methods of data collection, questionnaire design and data preparation are discussed.

3.1 Study design and chapter overview

- The DP Students Workload Study is designed as a longitudinal survey study, in which DP students are asked to respond online at three points in their two-year DP career ('waves'): beginning of Year 1, end of the Year 1 and end of Year 2.

- Additionally, DP coordinators were surveyed in order to gather information about school policies and services to support students. In the following paragraphs, we describe the advantages and challenges of this longitudinal, multi-level research design, and the characteristics of the sample.
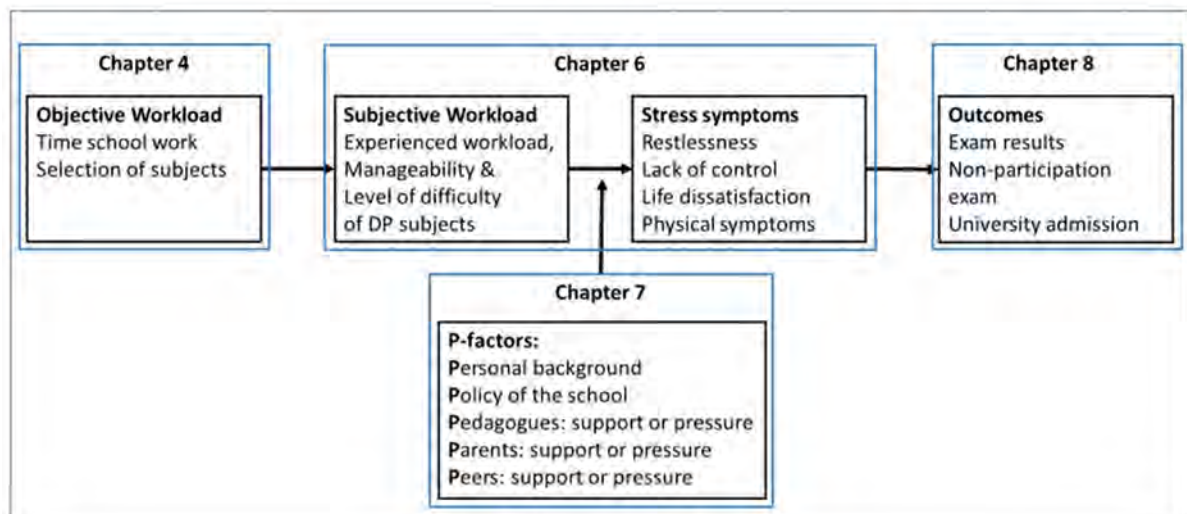
Figure 3.1 Causal model with chapter overview

- As explained earlier we conceive of stress development as a four-stage process in response to a stressor. These stages are depicted in the left column in **Figure 3.1**. The topics of the chapters are depicted by the blue outlines. **Chapters 4 and 6** describe the development of respectively objective and subjective workload over time. **Chapter 5** (not part of the causal model) provides some additional qualitative data on students' subjective workload and stress.

- **Our first main research question is to what extent subjective workload leads to stress symptoms.** This is subject of **chapter 6**. Here we focus in particular on the question of causal order (the arrows point both ways). Unlike objective workload, subjective workload reports may to a large extent be caused by existing stress symptoms, as stress develops during the programme (*reverse causation*) or as it existed already in students before they entered the programme (*spurious causation*). In other words, the extent to which students perceive the DP workload as manageable may also be influenced by their already existing symptoms of stress. The power of the three-wave panel design we use is its ability to separate effects of stress as a pre-existing condition from stress as developed during and due to programme participation.

- **Our second main research question is whether there are objective programme characteristics that influence how students experience and deal with complex and intense school work**. This is subject of **chapter 7**. Students report on their personal experiences of this **objective workload** via our **subjective workload** instruments. Here, we are particularly interested in the degree to which DP coordinators and individual students vary in the degree to which they judge the programme to be too demanding.

- **Our third main research question is which environmental factors (P-factors; Personal** (students' individual and social background), **Parents** (parental pressure and support), **Pedagogues** (teacher pressure & support), **Policy** (school well-being policy), **Peers** (peer pressure and support) **influence the stress development process**. This question too is answered in **chapter 7**. We expect students' academic abilities to affect the experience of subjective workload (i.e. being able to manage the workload).. Furthermore, schools may differ in the planning and coordination of exams and assignments which may influence how students experience pressure and the manageability of the DP.

- **Chapter 8** describes the outcomes of DP students (e.g. academic achievement, university admission, non-participation in the exam) and how they are impacted by stress symptoms.

- In **Chapter 9**, finally, we summarize DP students' and coordinators' respective evaluations of possible measures to reduce the DP workload or to make it more manageable for students.

A three-wave panel design

- The study is designed as a **longitudinal, three-wave panel survey:** We follow students during their entire two-year DP career, with measurements taking place at three time points, waves. The three-wave panel design has decisive **advantages** over two-wave panel designs as it **allows for the assessment of change *within* the measurement period**. A comparison between the first two waves brings out the effect of Year 1 of the DP. A comparison between the last two waves brings out the effect of Year 2, which – due to the exam period –is structured rather differently than Year 1.

- We can compare the **development of subject workload and stress in Year 1 to the developments in Year 2.** We can also compare **subjective workload and stress within students** in this way. Of

concern here is how students perceive their workloads at different stages of the two-year DP cycle as well as the stress levels that they report in response to these workloads. With these three waves, we can trace students' development over the two years. In addition, as part of the third wave, we added retrospective questions to trace the development of workload and stress during the second year.

- Another use of the panel model is to resolve issues of **causality** in non-experimental designs**.** In cross-sectional (one-wave) designs, associations between individual attributes cannot be interpreted as cause and effect. If we consider the relationship between subjective workload and stress, there are three alternatives to account for this correlation:
  - o *Causal*: workload causes stress. It is plausible that the relationship is positive (more workload causes more stress).
  - o *Reverse*: exogenously generated stress symptoms cause students to report their workload as being manageable.
  - o *Spurious*: the association between workload and stress is due to pre-existing conditions, such as gender or cognitive ability, which could influence both the amount of workload and the level of stress reported. The influence of such so-called confounders may be positive or negative.[1]

- **Multi-wave panel data** such as applied in this study **can resolve the issues of reversed and spurious causation** because of two related features. First, the multi-wave design allows one to identify the direction of causality, which may still be of a reciprocal nature. Traditionally, this is done by a so-called cross-lagged design in which characteristics in a given wave t are caused by characteristics in wave t-1 (the previous wave). A more direct approach, however, is the simultaneous reciprocal causal effects model, in which causality is examined within one wave, and the identification of the two arrows is obtained by excluding cross-lagged effects. The basic structure of this model is shown in **Figure 3.2**. It can be estimated with a structural equations model, for which we used the statistical programme Stata 14.

---

[1] A major problem here is that confounders may be – and likely are – unobserved (not measured), and unknown factors affect the relationship in unknown directions. This problem is known in econometrics as "unobserved heterogeneity" and is regarded as one of the defining delimiters between experimental and non-experimental research designs.
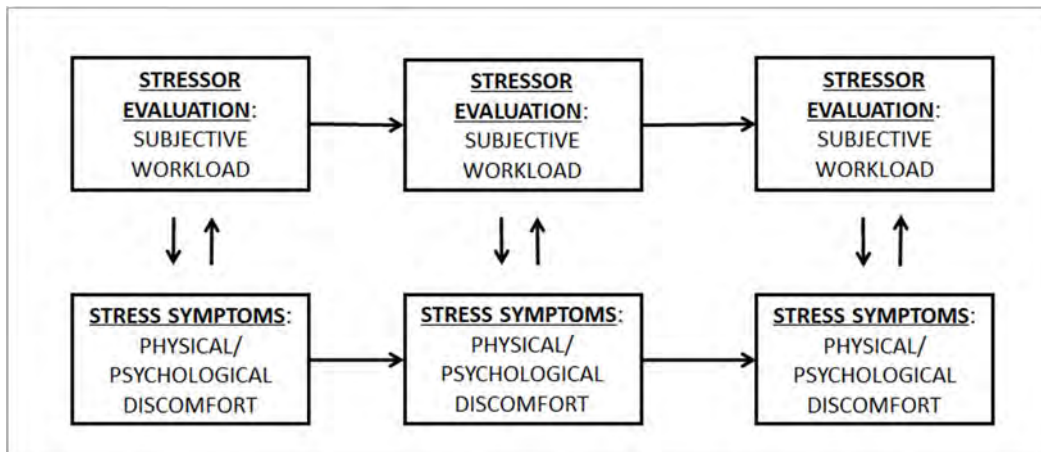
Figure 3.2: Reciprocal causal effect between Students' Subjective Workload and Stress symptoms

Methodological assumptions and limitations

- Another concern is the **lack of comparison data** in the study on non-IB students and schools. As this study includes only DP students and they are enrolled in the Diploma Programme from the start of the study, we do not know how scores on for instance stress and workload compare to regular students and whether scores over time are stable or changeable due to the programme characteristics. It may be possible, for instance, that increases in stress are normal among students who are working towards their finals in this age group. What we can do is compare scores to other studies and base conclusions on these comparisons.

- The current study has a large sample of students around the world and standard measurements were used which is useful for comparison across samples from other studies. A drawback of this method is that we cannot directly compare backgrounds of IB and non-IB students, nor can we compare students doing their exams in exactly the same time period. **Thus, we cannot exclude external influences on the findings and we cannot fully exclude spurious causation**. This means that the differences between DP students in this study and students in other studies might be related to characteristics of DP students rather than of the DP itself. This needs to be taken into account when drawing conclusions based on this study.

- The possibility of **selective response** during the survey is also a concern. Possibly stressed students responded to the survey because the topic was in their interest. However, it is also possible that students did not respond to the survey because they were too busy with school assignments. To address this issue, we will look at the stress levels of students at the different waves.

- A result of the design of the study is that in **wave 3** students were no longer DP students (the third survey was held after the May exam session). All **measurements** therefore were in **retrospect** and were contingent on students' memory. Moreover, in the third wave not all independent variables

were available, and some were specifically about the exams at the end of Year 2. This makes it more difficult to compare the third wave to wave 1 and 2, at least for some variables.

3.2 The sample

- The **initial sample** of the **student survey** consisted of **258 schools** (very close to 10%) from **50 countries**. Due to the adapted sampling procedure, the selected sample of schools is highly representative with respect to:
    - o Continent
    - o Language of instruction
    - o Number of DP schools in country (with May exam schedules)
    - o School status (private/state)

- In other words, the sample is diverse on these characteristics which makes it possible to draw general conclusions and to compare results on different characteristics. Out of the 258 school **coordinators 98 from 36 countries collaborated** and sent sufficient contact details on **4,854 students.**

- Schools were asked to report on their well-being policies in the middle of the panel period via **a school survey**. **2,894 DP coordinators**[2] were approached and around **2,200 filled out the survey,** including 119 DP coordinators from 91 schools where the student surveys were also distributed**. Appendix B** provides a detailed description of the sampling process.

3.3 Questionnaire design and measurement instruments

- **Table 3.1** provides an overview of the concepts and measurement instruments in the student and school surveys respectively.

- For **workload measures** a distinction is made between **objective and subjective workload** as described in the causal model. Both are measured using items from the student surveys. To further validate these measures, they are compared to what DP coordinators think about the objective and subjective workload of DP students at their school.

- **Objective workload** is measured by the specific curriculum that students follow in the DP, ie, the specific subjects they choose from each DP subject group and whether they take it at standard or higher level. Each **subject** is weighed by its difficulty and time burden. Both students and coordinators were asked to rank DP subjects along the two dimensions of difficulty and time

---

[2] Due to miscommunication, the survey link was sent to all DP coordinators, rather than only those from the participating schools. Where possible, eg, in the description of school policies and services, we used all responses. Also, the survey was sent to DP coordinators and Heads of school, which is why for some schools, we have information from two respondents.

burden. Furthermore, the **time students spend on school work, homework and additional lessons** are taken into account.

- For **subjective workload** students report as how heavy or light they experience their DP workload and how manageable they find it, using four standard statements, repeated at each wave of the survey[3].

- **Stress symptoms** are measured using four standard, often-used batteries of statements, adapted to the context of the DP.

- Student outcomes are measured partly with survey items, comparing aspired and achieved **university admission** and partly by IB exam records to estimate **non-participation in exams** and measure **academic achievement**.

- Possible moderators that could weaken or strengthen the relationship between workload and stress are the **P-factors**: **Personal** background (gender, academic ability, language proficiency, home situation, socio-economic status (SES) of parents), **Policy**, **Pedagogues** (teacher support or pressure), **Parent** and **Peer** support or pressure. These are described in chapter 7.

3.4 Data Analysis

- Descriptive statistics are used to depict overall scores on the important variables (student workload, stress) over time and to compare these with other studies.

- Factor analysis is used for the stress scales and the well-being measures to determine the **sub-dimensionality of the indicators** of stress not a priori, but empirically; we have conducted several factor analyses, both within and between waves.

- Multi-level analysis is used to disentangle the **respective role of individual, programme-related and school-level predictors of student workload and stress levels**, using the statistical programme MPLUS-MLWIN. Multi-level modeling has the advantage of allowing for the **simultaneous analysis of the total impact of multiple contexts** (ie, individual, school and country). This analysis technique assumes a nested data structure. It yields estimates for variables of different levels, in particular, the impact of programme-related factors on student stress levels and as part of the school impact. Individual background variables will be used both as dependent and control variables.

- Structural equation modeling in Stata 14 is used to simultaneously model change in **stress and subjective workload over time** in a cross-lagged model.

---

[3] See Appendix for relevant questionnaire items.

- The next two sections describe the most important features of the participating schools, the differences between schools as well as DP students' background information. Tables are usually shown in **Appendix A**.

3.5 DP schools' characteristics

We start this section by reporting background information about the schools as reported by DP coordinators. We begin by describing general features of the schools; the region of schools in the sample, the school status (private or public) and the size of the DP enrolment.
We have around 1870 schools in the sample of DP coordinators.

Region and Language choice

- **Figure 3.3** shows that the sample contains **schools from all over the world**. Most schools are in North America (39.2%), followed by Western Europe (14.1%) and Asia (14.1%), South America (9.6%).
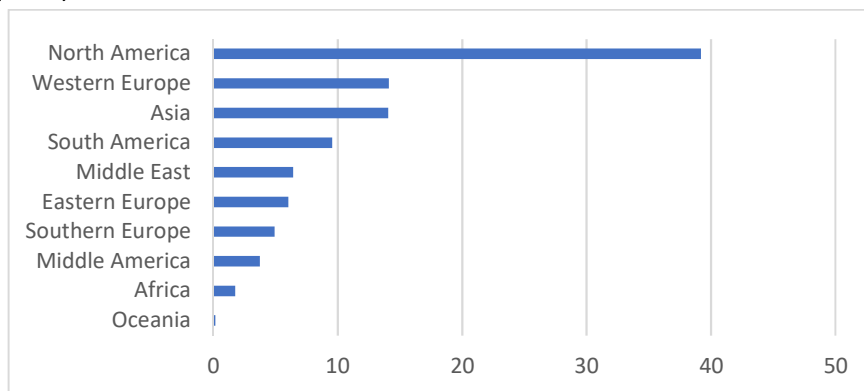
Figure 3.3: Schools by region

- **Figure 3.4** shows which **language** the coordinators chose for filling out the questionnaire. This gives an indication of the language background of coordinators; we do not know, however, if it corresponds with the language of instruction of the schools. **English** was chosen by the majority of coordinators (77.5%), **Spanish** is chosen second most (20.6%) and **French** is rarely chosen (1.9%). Of all the schools in the sample where English is the first choice about 44.7% is located in North America, 16.3% in Western Europe and 16.6% in Asia. Most schools where coordinators choose Spanish are in South America (58.5%), Southern Europe (22.5%), and Middle America (19,0%). Schools (36) where French is the first choice are mainly in North America (26) and Western Europe (7) (see also **Appendix 3.1**).
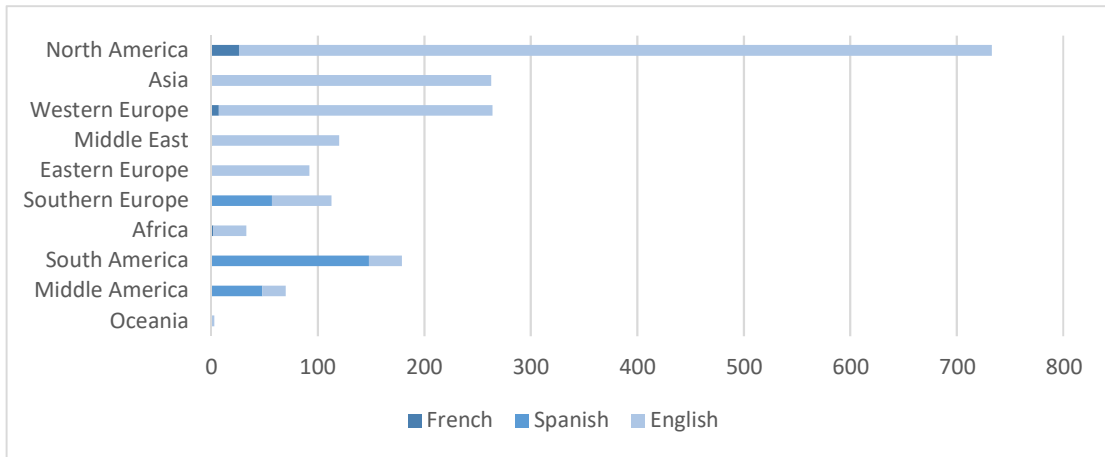
Figure 3.4 Language choice by region

School status by region

- Just over half of the schools in our sample are **private** (50.6% versus 46.0% public schools). **Figure 3.5** shows that in Oceania, Africa, the Middle East and Asia, a large majority of the schools are **private**. In Europe the number of state schools is much larger, but still a minority. In South America about half of the schools are private, and the other half state. In North America state schools are strongly overrepresented (81.3%) in comparison to other regions (see also **Appendix 3.2**).
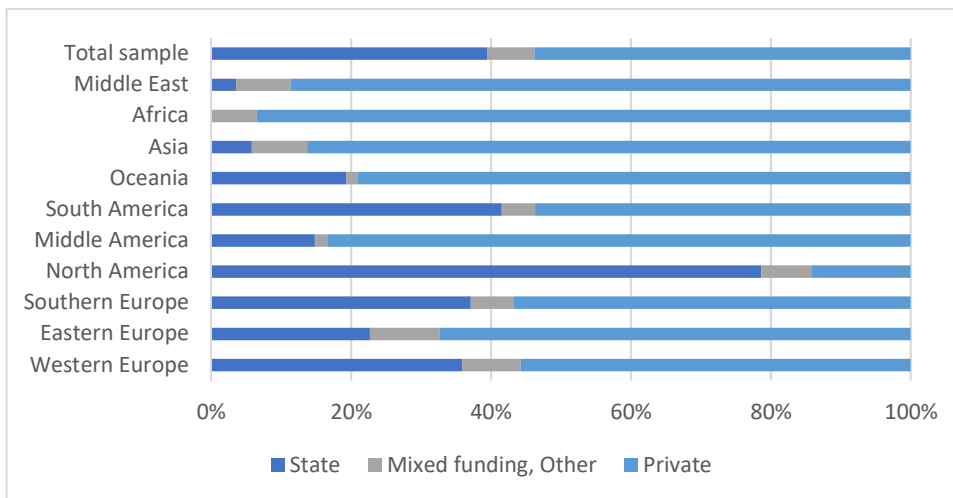
Figure 3.5 School status by region

DP enrolment

- To see what the size of the Diploma Programme is at the schools in our sample, coordinators were asked how many **students were enrolled** in the DP, as of October 2017 (see **Appendix 3.3**). **Most students enrolled for the first year** or the fulltime programme. The average number of fulltime DP students is 45 at each school for the first year (median = 30, maximum = 1,750) and 37 (median = 26, maximum = 439) for the second year. The number of students that enrolled part-time is much

lower. An average of 19 students (median = 0, maximum = 700) is enrolled part-time for the first year at the schools in the sample, 18 (median = 0, maximum = 440) for the second year.

- **Figures 3.6ab** (and **Appendix 3.4)** describe in more detail how the enrolment is divided amongst the schools in the sample. Most of the coordinators have between 1 and 50 students in the first year (full-time: 67.6%; part-time 29.9%) and second year (full-time: 73.1%; part-time: 36.7%). Somewhat more than half of the schools report to have no part-time DP students (59.6% in year 1; 53.1% in year 2). At hardly any of the schools are there more than 200 DP students (< 10%).
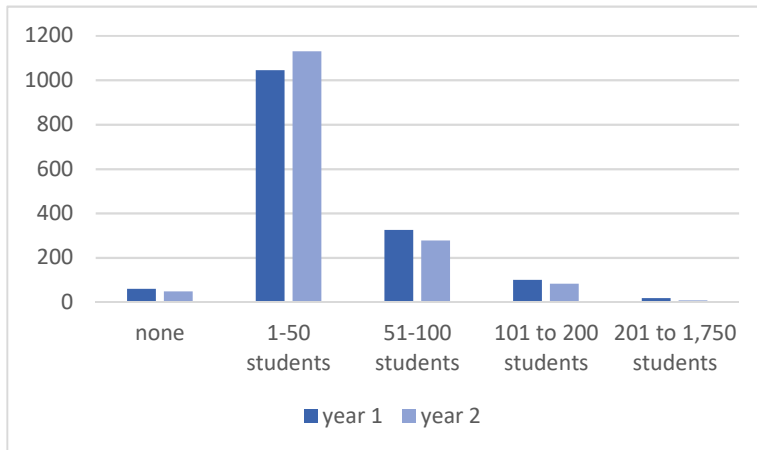
Figure 3.6a Number of full-time DP students in year 1 and year 2)
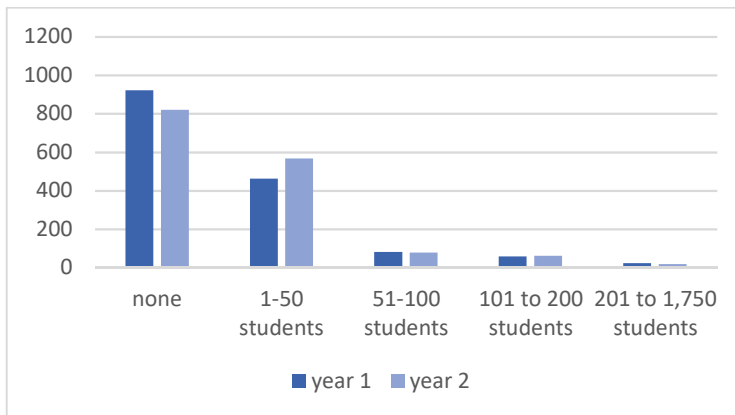
Figure 3.6b Number of part-time DP students in year 1 and year 2)

Percentage of DP students for whom language of instruction is not first language

- **Appendix 3.5** presents the percentage of students **for whom the language of instruction at school is not their first language**. About 40% of the schools say this applies to less than 10% of the students. According to a quarter of the schools this is the case for more than 60% of the students.

- When we look at the **language division** we see in **Figure 3.7** (and **Appendix 3.5**) that the smallest percentage of less than 10% applies mainly to the Spanish speaking schools. At 88.0% of these

Spanish schools for less than 10% of the students Spanish is not the first language. This accounts for 30.2% of the English and also 31.0% of the French schools. At another 32.3% of the English schools for more than 60% of the students English is not their first language. At a quarter of the French schools this counts for between 40 and 60% of the DP students.
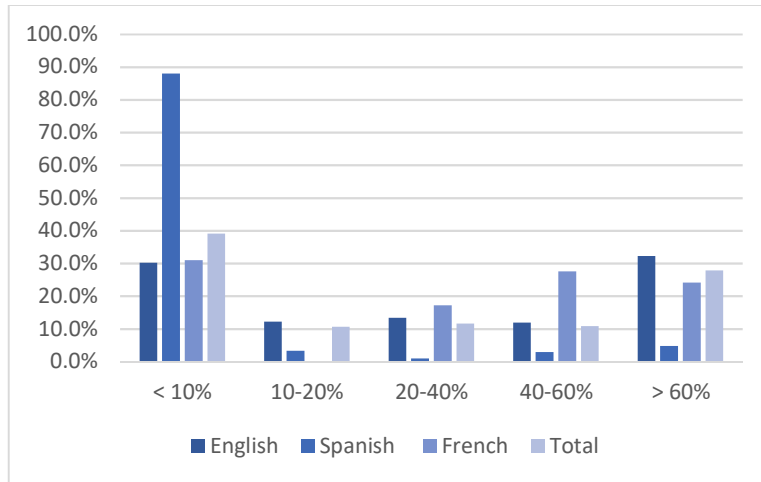
Figure 3.7: Language choice of school and percentage of students whose first language is not language of instruction

3.6 DP students' personal backgrounds

In this section we describe the backgrounds of the sampled DP students and how these differ between countries and between schools. As these characteristics already exist before students start the DP and may also be related to experiences of workload and stress, they are important control variables in the analyses. We describe: gender, academic ability, language proficiency, home situation, and parents' socio-economic status.

Gender

- The data on the gender of the DP students in the sample is based on school registers. **Girls** are slightly overrepresented in the Diploma Programme: 54.6%[4]. There are no differences in this percentage between countries or between schools[5].

Academic ability and language proficiency

- To estimate academic and language proficiency prior to the DP students were asked to report on their **academic ability in** general as well as on their abilities in English, mathematics and science. Students were also asked how their academic ability compared with their former classmates.

---

[4] This is similar to what is reported by the IB on DP students who did exams in May 2017 (56.6%) and May 2018 (56.7%) (IBO, 2017/2018).
[5] For analysis we use the numerical variable girl (range 0-1).

Finally, students were asked to rate their **abilities in the language of instruction** (understanding, reading, speaking and writing).

- **Figure 3.8** (and **Appendix 3.6**) shows that DP students estimate their **academic abilities** as high – 72.4 % says these are very good or excellent. About half of the DP students think their abilities in the language of instruction[6] (48.5%), Mathematics (50.3%) and Science (51.0%), and their overall academic abilities (52.4%) exceed those of their former classmates to a high degree (very much or even extremely above average).
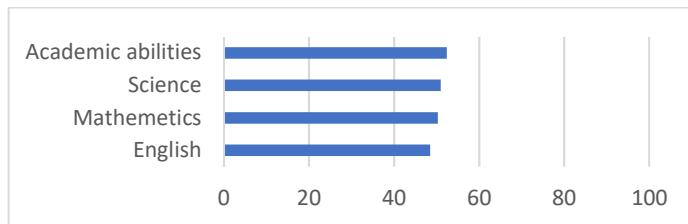
Figure 3.8 Percentage of students reporting academic abilities very much or extremely above average

- **Figure 3.9** (and **Appendix 3.6**) shows that most DP students estimate their **language abilities** as very good or even as good as a native speaker (between 65.3% and 78.1%).
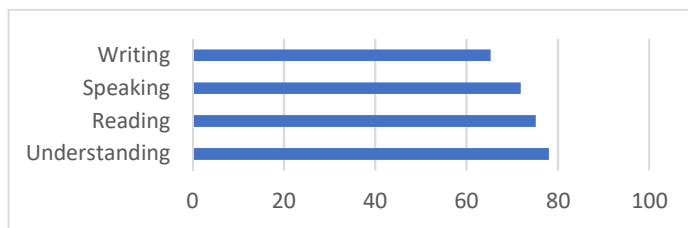
Figure 3.9 Percentage of students reporting language proficiency as very good or native speaker

- Factor analyses showed that both sets of questions were two separate dimensions: **general academic ability** and **language proficiency**. This means the questions can be added up into these two categories for further analyses. Both scales have been measured with high reliability, Cronbach's alphas are .83 and. 93 respectively[7].

- **Variations in academic ability** and **language proficiency** are mostly between students of the same school, but to some extent (9% - 14% respectively) schools and countries also differ in the average abilities of their students.

Home situation

---

[6] English, French or Spanish respectively.
[7] For analysis we constructed two scales in a 0-1 range, 0 representing students with lowest, 1 student with the highest academic and language abilities.

- Almost all students (98.4%) report living with at least one parent, 83.5% with two parents. However, when answering this question students apparently did not always refer to their daily situation, as 10% say they are **internal in a boarding school**. The number of students living internal in a boarding school differs rather strongly across countries (see also **Appendix 3.7**).

Parents' socio-economic status (SES)

- As **Figure 3.10** (and **Appendix 3.8**) shows, DP students' parents are rather highly educated: 69.3% of the mothers and 74.0 % of the fathers have at least a bachelor's degree. There are some school differences in the **education level of the parents**, meaning that schools differ in the degree they attract students from higher or lower educated parents[8].
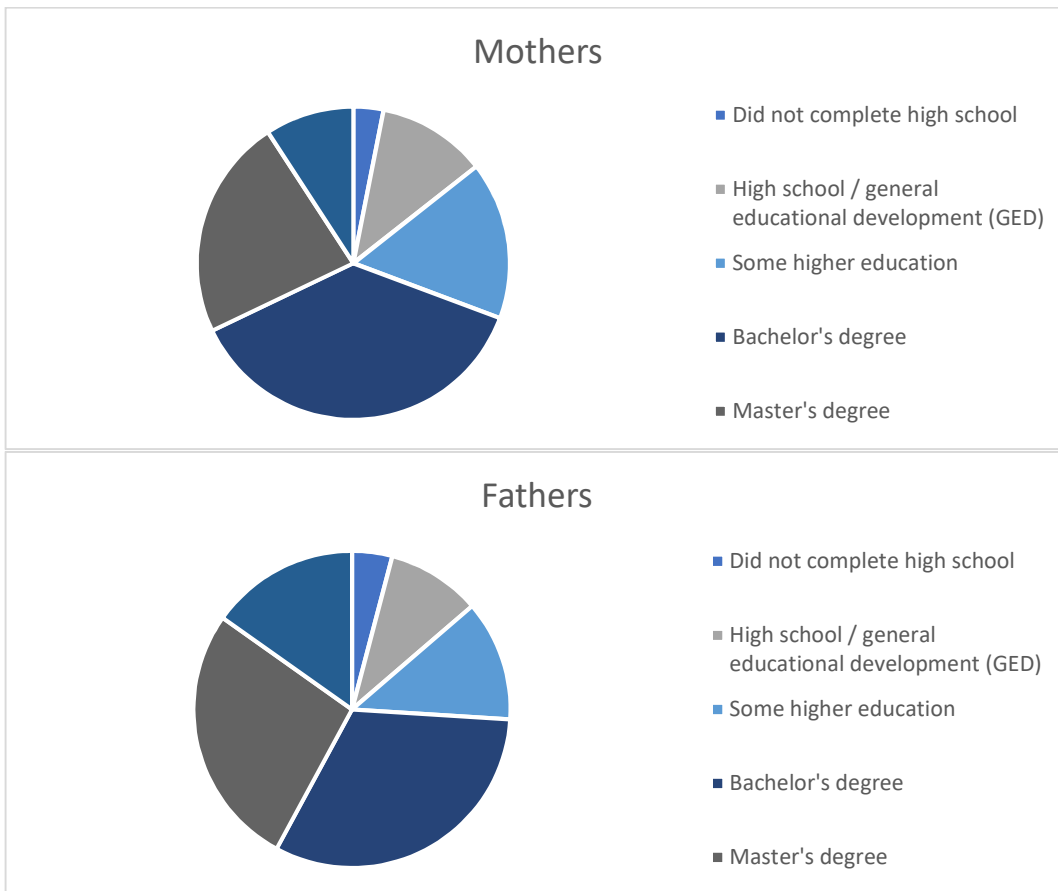
Figure 3.10 Education level of DP students' parents by gender

- **Figure 3.11** (and **Appendix 3.8**) shows that two thirds of the mothers (66.4 %) and 92.1 % of the fathers have a paid job. Some country differences emerge in the percentage of mothers having a paid job.

---

[8] For the analyses we used mothers' and fathers' education recoded into a 0-1 range. For the mother's labour market position, we used a dichotomous (0/1) variable, indicating whether the mother has a paid (full time or part-time) job.
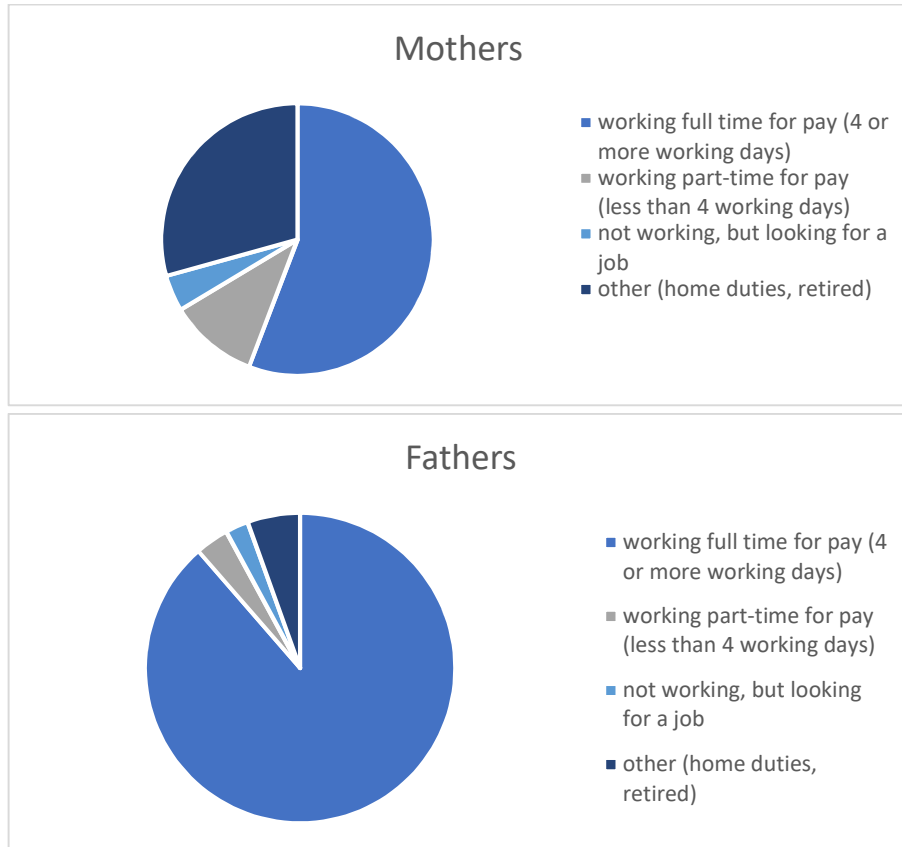
Figure 3.11 DP students' parents' labour market position by gender

Students' motives for pursuing the DP

- **Figure 3.12** (and **Appendix 3.9**) lists the factors students mention that made them decide to pursue the DP, in order of importance. **Most important are the opportunities the DP offers to enter good universities (92.2% answered important or very important), a good job (83.3% answered important or very important) and personal development (77.1% answered important or very important)**. The international orientation and being with like-minded students are second most important (68.1% and 57.5% respectively answered important or very important). Pragmatic factors, such as geographical vicinity, are much less important to DP students (between 21.8% and 28.3% answered important or very important).
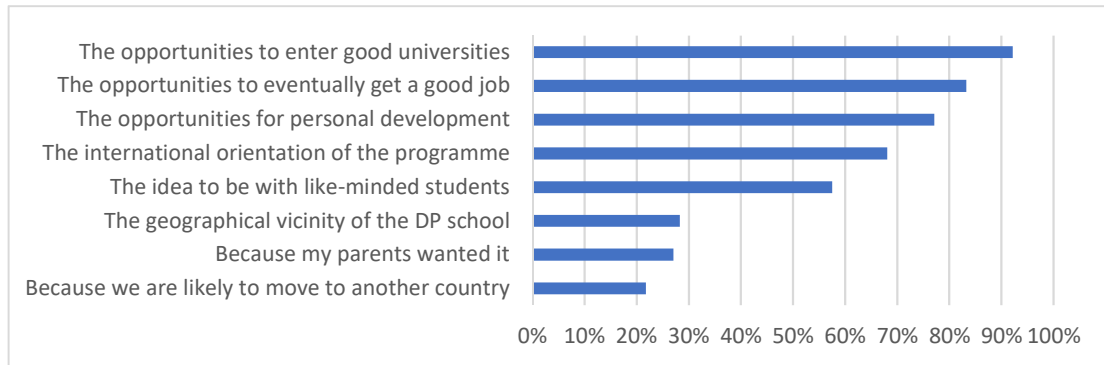
Figure 3.12 Students' motives for pursuing the DP

- Factor analysis (**Appendix 3.10**)shows that three latent dimensions underlie these items[9] which could be named: **personal development, career opportunities, and practical concerns**. We constructed composite variables for the three items that correspond to the factors 'career opportunities' and 'personal development' and 'practical concerns' respectively.

- Multilevel analyses (see **Appendix 3.11**) show that there are hardly any **country or school differences in the career motive** (1.9%). This means that the career motive to opt for DP does not differ across countries or schools and that differences are mainly due to individual differences between students that occur within all schools. School differences in the personal motive are a bit larger (4.7%) but still small. We conclude that differences in career and personal motive to pursue the DP are due to individual characteristics, rather than schools attracting students with certain motives to enrol in the programme.

- What are these individual characteristics? Multilevel regression models (**Appendix 3.12**) show that the **motives of career opportunity and personal development are more often mentioned by students with higher academic abilities**. The personal development motive is also more often mentioned by girls than by boys. There are no differences in these two motives according to language proficiency, being an internal student or father's education. Differences in the **motive 'practical concerns' only occur according to students being internal or not**. For internal students the geographical vicinity, the likelihood to move to another country and their parents' preferences are stronger motives to pursue the DP.

---

[9] Both Kaiser's criterion of eigenvalue > 1, and Cattell's point of inflection

Chapter 4: OBJECTIVE WORKLOAD IN THE DIPLOMA PROGRAMME

Contents

Chapter 4: OBJECTIVE WORKLOAD IN THE DIPLOMA PROGRAMME

In this chapter we study the objective workload the DP generates for students. Our first aim is to identify the more difficult and time-consuming parts of the curriculum. While in principle all students follow the same programme, the DP leaves students some freedom to adjust the curriculum to their own needs and capabilities. Consequently, we can identify variation in how students customize the programme at the individual level. There are variations in students' individual selection of subjects and the time they spend on their homework and out-of-school lessons. **We describe differences in difficulty and time demands of DP subjects and answer the question which students have the highest objective workload in terms of their individual DP curriculum. Figure 4.1** illustrates which part of the causal model is addressed in the current chapter.
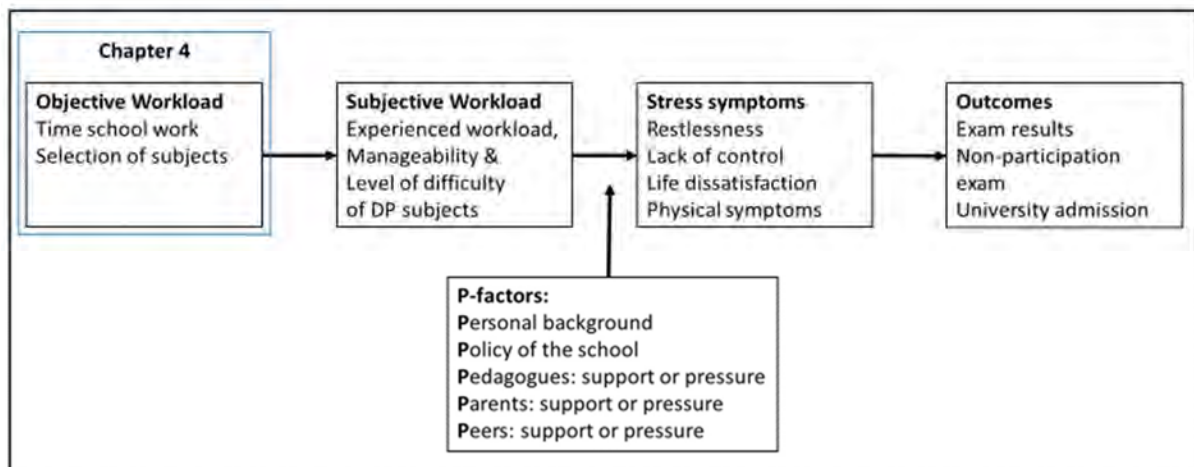
Figure 4.1: Causal model of the student objective and subjective workload and stress

DP subjects vary greatly in the workload in terms of time burden and difficulty. DP students and DP coordinators agree on the **subjects that are heaviest, mainly Mathematics and Sciences subjects**.

Students' subject selections vary by **academic ability (prior to DP) and by language proficiency**. Students with higher academic abilities select heavier subjects. Similarly, students with a lower language proficiency choose heavier subjects than students with higher language abilities, likely due to a preference of subjects that rely less on language. Moreover, the higher the socio-economic status of the parents, the heavier the workload of the selected DP subjects. **Girls** select subjects that on average have a somewhat lower workload than those selected by boys.

Students spend on average **47 hours on DP school work** per week, about **26 hours on regular classes**, and about **20 hours on homework**, and, **1.5 hours on additional lessons.** Students with higher academic abilities spend more time on their DP homework and additional lessons; students with higher language

abilities spend less time. Students with heavier subject selections spend more time on their school work than students with the lighter subject selections. Girls spend more time on their school work than boys.

Next to their school work, students spend time on sports and arts (one and a half hours daily), and on relaxing, internet and (social) media (between three and four hours a day). They sleep between six and seven hours a night, on average.

4.1 Structure of the Diploma Programme

- The DP is broadly structured in the following way: **DP students choose subjects from each of the six subject groups**: (1) Language and literature, (2) Language acquisition, (3) Individuals and societies, (4) Sciences, (5) Mathematics and (6) The arts. Students may opt to study an additional subject from the first four groups instead of a course in the arts, and or take one or two interdisciplinary subjects, which count simultaneously across two groups. These students still need to take a total of six subjects.

- Within each subject group, students choose to do subjects at standard level (SL) or higher level (HL). **Each student must complete at least three subjects at higher level**.

- **In addition to these six subjects, students must complete course work in three core subjects that are specific to the DP: Theory of knowledge (TOK), Creativity, activity, service (CAS) and The extended Essay (EE).** TOK can be regarded as an epistemology course, CAS asks students to engage in experiences and projects that foster creativity, purposeful activity and community, and the EE is a 4,000-word essay on a topic related to an academic subject of the student's choice.

4.2 Enrolment in DP subjects in wave 1

- The sample includes both 'diploma' students who follow the entire DP and 'course' students who follow some DP subjects but not the entire programme. The large **majority** of the students in the sample reports, at the start of the programme in year 1, to do the **full DP** (93.8%). This percentage is higher than that of diploma students who registered for the final exams in year 2 (82.0%), see **Appendix 4.1**. The difference is the result of the higher response among diploma students. Among DP students who participated in the survey at waves 1 and 2, 93.2% characterize themselves as full DP students at both waves, 2.8% as course students. 2.7% switched from full DP student in wave 1 to course student in wave 2, 1.2 % the other way around. Unexpectedly, given that we aimed at a sample of year 1 DP students only, 4.4% of the students says to be in year 2. A likely explanation for this is that some schools allow students to take DP subjects over the course of three years.

Construction of full/part-time enrolment in DP

In the analyses on objective workload we include whether students follow **the DP as diploma or course student**. To construct this variable, we used students' reports in the survey and when there are missing,

---

complemented with information from the exam file. The variable is labeled 'course DP', a zero indicating full DP, a 1 indicating not full DP.

- When asked about the **number of DP subjects** they are taking, students report an average of six subjects (full DP students 6.1), which is well below the required nine subjects (by comparison: course students report an average of 3.8). If instead we count students' reported selection of subjects (an alternative measure of their number of subjects) the average number of subjects is 8.0 DP subjects for full DP students. The exam registration, finally, counts 8.5 DP subjects for full DP students. About one third of the students (30.8%) reports to also follow subjects from **other international (14.6%) or national or regional (26.1%) programmes that qualify for university, such as Advanced Placement (AP) in the US.** Among those who do, the average number is estimated as 3.3 subjects from international and 2.7 from national programmes (see also **Appendix 4.1**).

4.3 Difficulty and time-burden by subject in relation to students' subject selection

- Within the requirements of the DP, **students' individual subject selections** may lead to variations between DP students in the difficulty and time burden of their individual curriculum. These variations in subject choice constitute a first measure of students' objective DP workload.

  To begin with, we estimated the workload of each **DP subject**. Both DP students and coordinators were asked to name the subjects that in their opinion are the heaviest in terms of the difficulty of the subject matter and the time burden of homework and exam preparation. By compiling all students' and coordinators' individual ratings, all subjects could be ranked from the lowest through the highest workload. **Table 4.1** presents the ten 'heaviest' DP subjects according to students and coordinators respectively, corrected for number of enrolled students. **Subjects in the Sciences and Mathematics groups are rated as most difficult and also as having the heaviest time-burden, by both students and coordinators.** The only (noteworthy) exception is that coordinators also name two arts subjects - Music HL and Visual Arts HL - as having exceptionally high workloads. Overall, DP students and coordinators largely agree, not only on the 10 highest ranked subjects, but generally on all subjects' workloads (correlation .83). An overview of the mean workload ratings of all DP subjects and correlations between several measures are presented in **Appendix 4.2** and **4.3**. The detailed procedure to construct this measure of objective workload, which combines both students' and DP coordinators' views is presented below.

- Subsequently, having estimates of the workload of all DP subjects, we could estimate the workload of each student's individual DP curriculum. **Table 4.2** presents some examples of more and less heavy subject selections. The individual curricula with the highest workload tend to contain more Sciences subjects, which are rated among the heaviest. The individual curricula with higher workloads typically include Mathematics HL, whereas Mathematics SL is found much more in the lighter ones. On the other hand, group 1 and 2 subjects in the heavier individual curricula tend to be SL, whereas they are mostly HL in the lighter curricula.

Construction of Workload of subject selection

Difficulty and time-burden by subject, students' ratings

In wave 2 and 3 students were asked to rate DP subjects in terms of **difficulty and time-burden**. Students in wave 2 reported the **easiest** and second easiest, the **most difficult** and second most difficult DP subject, and also the subjects for which they spent the most and second most amount of time doing homework, and the least and second least amount of **time doing homework respectively**. Moreover, in wave 3, students reported the DP subjects they spent the most and second most, and the least and the second least amount of **time to prepare for the exam papers**. DP coordinators were asked to rate the DP subjects in a similar way, by the same workload indicators.

We took into account only the evaluations of students who were enrolled in the subject, according to their report in wave 1 or the exam registration, which is the majority (all ratings > 80%). Students who were not enrolled in a given subject provided less valid answers (positive correlations between low and heavy workload indicators). Furthermore, we included the opinions of all DP coordinators, also from schools not in the sample.

We used the ratings of the highest difficulty and time burden to derive the rankings (we disregarded the easiest and least time requiring). Correlation analysis (**Appendix 4.3**) showed that these ratings did not constitute one single dimension, but a two-dimensional clustering of the heavy workload items (most difficult, high time-burden) on the one hand and low workload ratings (easiest, low time-burden) on the other. This structure of a low and a heavy workload dimension is confirmed by factor analysis, revealing two latent factors (moderately correlated r = -.355). Also, reliability analyses (not in table) show that separating the two dimensions gives the highest consistency (Cronbach's alpha high workload = .83 (students) and .89 (DP coordinators), Cronbach's alpha (low workload = .74 (students) and .97 (DP coordinators), Cronbach's alpha whole scale = .78 (students) and .85 (DP coordinators)). As the items on the heaviest workload are more closely aligned to the research objective, we chose to use those.

The number of times a subject was named as most or second most difficult, requiring the most or second most time spent on homework or time needed for exam preparation, is a first rough measure of workload. As subjects with large numbers of participants have a higher probability to be named in all questions, the number of ratings was corrected for the estimated number of enrolled students, according to the exam registration. The calculation was as follows:

$$\text{Workload score difficult} =$$

$$\frac{\text{mean(nr times subject most difficult, nr times subject second most difficult)}}{\text{(nr enrolled students} * \text{(nr valid answers on most} + \text{second most difficult)}/2 * \text{sample size)}}$$

For students, this measure reflects the percentage of enrolled students who rate the subject as most difficult or having the highest time burden and ranges (after winsorizing outliers) from (almost) zero to (theoretically) 1 (100%), but in reality, to .54. (difficulty), .57 (most time homework), and .54 (time exam preparation). Subjects that were not named were given a score of 0.

Next, we took the average of z-standardized scores of both students' and coordinators' ratings, which we again z-standardized to end up with a variable in which zero refers to a subject with average workload. The correlation between the two ratings is .83, implying that students and coordinators largely agree on the workload of the DP subjects.

Difficulty and time-burden by students' individual DP curriculum

The workload of each student's individual selection of subjects was constructed by taking the average of the workload of the selected DP subjects, according to both students and DP coordinators. The selection of subjects was measured in wave 1 of the survey and is also available in the exam registration. The two measures do not fully overlap and have their own advantages and disadvantages. Wave 1 excludes the students who did not respond at that time. Moreover, many students changed their subject selection afterwards and would not have reported the subjects that had not yet started, in particular the core components. The exam registration excludes students who did not take the exam, but also is only an indication of the final curriculum, including possible adaptations in terms of the workload. The two measures of the selection of subjects lead to four measures of workload (students and coordinators: correlations between the two measures among students .68, among coordinators .71[1]). These four measures were z-standardized before we took the average, which we standardized and subsequently winsorized to a range from -2.5 to 2.5. On average, the workload of the selection of subjects in November year 1 to the exam in June in year 2, is similar (not statistically significant) and is .76 correlated (average students' and DP coordinators' ratings). For the analyses, we take the average of the subject selection as reported in November year 1 and as registered in June year 2, both by students' and DP coordinators' ratings. Finally, we recoded this variable **'workload of the subject selection'** into a range of -.5 to .5. Multilevel analyses show that there are considerable differences in the workload due to students' subject selection between schools (24.3% of the variance), but no country differences.

4.4 Differences in objective workload due to subject selection

- Which students select **a heavier DP curriculum**? **Figure 4.2** shows how differences between students' objective workloads, as measured by their subject selections, related to their personal background (see also **Appendix 4.5**).

---

[1] These correlations increase to .79 and .84 among respondents who have about the same number of subjects in November year 1 and in June year 2 (maximum difference one subject).
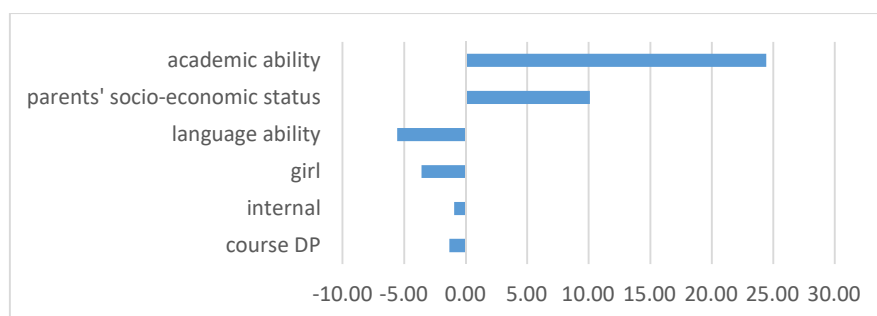
Figure 4.2: Workload due to the chosen set of courses, variations with social and personal background

- The most important factor is **academic ability** (prior to the DP). Students with higher academic abilities select subjects that are heavier in terms of difficulty and time burden. Surprisingly, students with a higher **language proficiency** choose lighter subjects, and, vice versa, students with lower language abilities choose heavier subjects. This could be caused by students avoiding language subjects at the higher level (HL), which are among the lower workload subjects and prefer different kinds of higher level subjects (eg, among the heavier Science and Mathematics subjects) that require less language skills. Also, **parents' socio-economic status** has a positive effect: the higher the socio-economic status of their parents, the heavier the workload of a student's subject selection. Finally, **girls** select subjects that on average have a somewhat lower workload than those selected by boys.

4.5 The time DP students spend on school work, homework and additional lessons

- A second measure of objective workload is the time DP students spend on school work, both for the DP and other programmes. **Students spend an average of almost 47 hours on school work for the DP**, both in wave 1 (beginning of November) and in wave 2 (end of year 1 June) (see for a detailed overview **Appendix 4.4**). In year 2, they spend about the same amount of time on schoolwork, with a slight peak in December-February. Across the whole sample, DP students additionally spend an average of seven hours a week on other programmes. This number is, however, much higher for the students who are actually following these programmes. On average, they spend about 16 hours per week on their non-DP subjects[2].

- Of their 47 hours of school work, **DP students spend 26 hours on regular lessons, 18 to 20 hours on homework and 1.5 to 2 hours on extra lessons**. This pattern remains rather stable throughout the DP. At the end of year 2 there is a small increase in the time spent on homework, and a slight decrease in

---

[2] The percentages students who report to spend time on school work of other programmes are 46.1% in wave 1 and 42.3% in wave 2, higher than the estimate based on the number of subjects (30.8%). Unlike the survey questions on the number of non-DP subjects, in the questions on time spent on school work there is no further specification of international or national/ regional programme, and no addition that programmes are meant that qualify for university.

the time spent in regular lessons, likely caused by the exam preparation period, which has fewer regular lessons. **Figure 4.3** depicts the developments in hours students spend on school work over time.
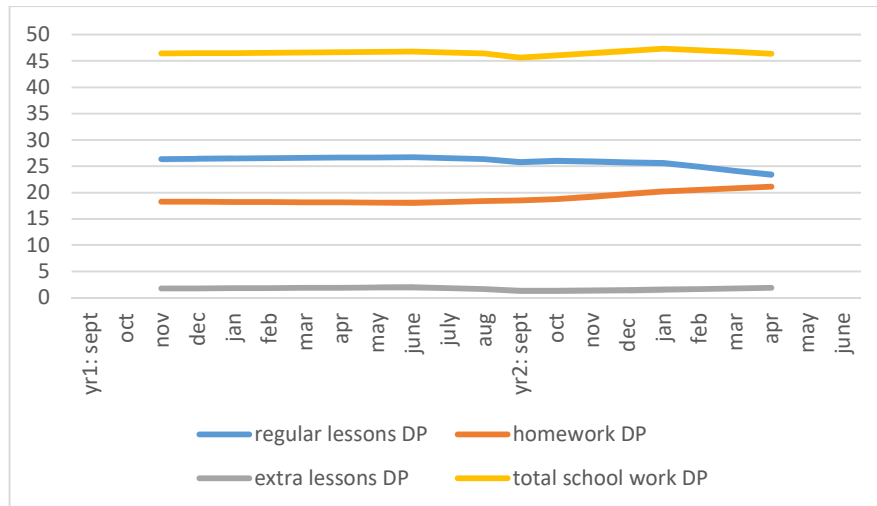
Figure 4.3: Hours per week time spent on school work

- The 47 hours spent on DP school work is somewhat more than the 44.1 hours estimated in the PISA 2015 study (OECD, 2016; Volume II), see Chapter 2, for 15-year old students in OECD countries. If the hours spent on non-DP subjects are taken into account, the number of hours DP students spend on school work surges to almost 54 hours. The 26 hours DP students spend on regular lessons are comparable to the 26.9 hours reported by students in the PISA 2015 study (OECD, 2016; Volume II), see also Chapter 2. The 20 hours they spend on homework and additional instruction are somewhat higher than the 17.2 hours found in the PISA study. Note, however, that DP students (aged 16-18) are not only older than PISA students, but they also all follow a programme that qualifies for university, whereas the PISA student population is much more diverse in level.

Construction of time spent on school work

Students in wave 1 (November) and wave 2 (June) in year 1 of the DP were asked to estimate how many hours they typically spend per week on school work, both in DP and in other programmes. Time spent on regular classes and homework (outside regular classes) was asked in ten categories varying from no time to more than 40 hours a week. Time spent on additional lessons (with tutor or similar, outside regular classes) was asked in ten categories, varying from no time to more than ten hours a week. To arrive at a measure of hours per week of each category, generally the class center was taken, for instance 'between 5 and 10 hours a week' was coded as 7.5, the lowest 'no time' category was coded as 0, the upper categories were coded as the bottom value named in that category, for instance 'more than 40 hours' was coded as 40. Students who gave partial information by answering some but not all these (six) questions were assumed to spend no time on that part of the school work (missing values coded as zero), mostly for the questions on the time spent on non-DP school work.

Students in wave 3, at the end of year 2 of DP, were asked to look back on year 2. They estimated their time spent on DP for the period of the beginning of the school year until November, from December

until February, and from March until the end of the school year. They did not report on the time spent on non-DP subjects anymore. Again, the hours they spent weekly were asked in the same categories as in wave 1 and 2 and were coded similarly. Again, missing values were substituted by 0 if one of the other (of nine) questions referring to the same period was answered. An overall estimate for the number of hours spent in year 2 was constructed as the average number of hours spent in the three separate periods.

4.6 Differences in objective workload due to the time students spend on school work

- Which students spend **most time on their homework and extra lessons**? **Figure 4.4 (a** and **b)** shows that in the beginning of year one students with **higher academic abilities** spend more time on their homework and extra lessons (see also **Appendix 4.6**). The difference is not due to the fact that students with higher academic abilities also choose heavier subjects (as this is controlled in the analysis). Overall, the difference is not so large (about six hours a week between students with the lowest and highest abilities). A **heavy subject selection** also leads to a higher time burden in terms of hours homework and additional lessons (three and a half hours per week difference between the lightest and heaviest curriculum). Students with higher academic abilities spend more time on their school work, partly because of their heavier curriculum, but also in general, irrespective of the subjects they choose. Students with lower **language abilities** spend more time on school work than their schoolmates who have better language skills, possibly because they need more time to process the subject matter. **Girls** spend more time on their homework and extra lessons than boys with similar personal backgrounds, on average an extra 2.82 hours per week. Finally, for obvious reasons, course students spend less time on their DP school work.
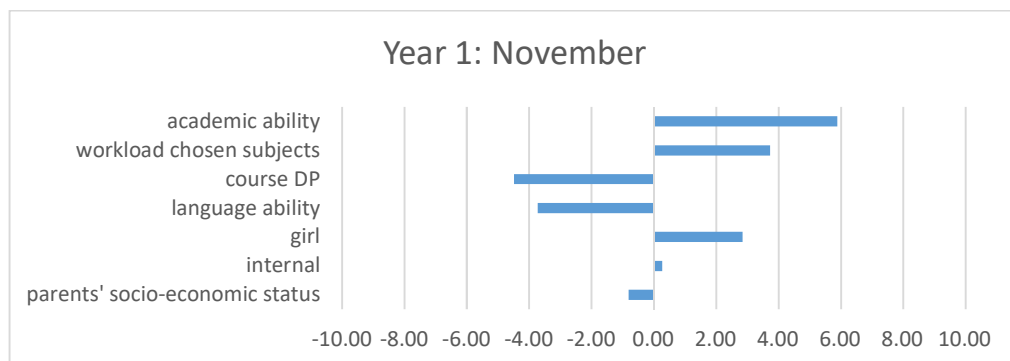
Figure 4.4a Time spent on school work (regular lessons, homework, extra lessons), variations with social and personal background – Year 1: November
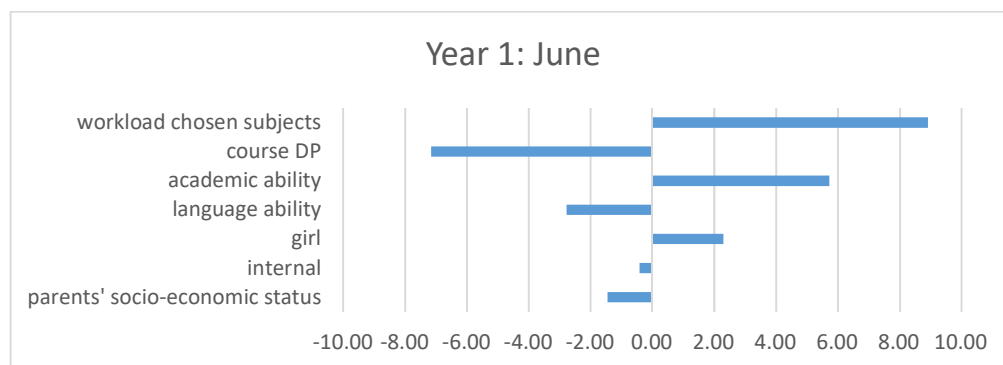
Figure 4.4b Time spent on school work (regular lessons, homework, extra lessons), variations with social and personal background – Year 1: June

- A similar analysis of the time spent on homework and extra lessons pertains to the data gathered in wave 2 at the end of year 1. At this point in time, more so than in the first wave, the time students spend on the DP may be a consequence of the subjective workload they experience. However, as will become clear, variations in the time students spend on the DP according to their personal background are quite similar to those of the previous wave, reflecting a rather stable over time pattern of differences between students. As in the beginning of year 1, students with **higher academic abilities** spend more time on their homework and extra lessons, compared to students with similar personal backgrounds and similar individual curricula. The **workload of the subject selection** also leads to a greater time burden, more strongly so than in the beginning of year one. The differences between students according to their **language abilities** are still related to the time students spend on homework. **Girls** keep spending more time on their school work than boys, still almost three more hours a week. Finally, again unsurprisingly, course students spend less time on DP than full DP students.

4.7 The time DP students spend on school work in relation to other activities

- In this section we examine the relative amount of time students spend on school work in relation to their other weekly or daily activities. **Table 4.3** gives an overview of the time students spend on a broad range of activities, including the time they sleep. The hours spent on several activities are also pictured in **Figure 4.5**. For ease of interpretation, the hours per week (as reported) are transformed into daily averages. We should be aware that the number of hours students spend on several activities should be interpreted with some caution. Some activities could be done jointly, for instance school work can be combined with time spent on social media, and time spent with friends could be during a joint sports activity. In these cases, the hours for separate activities will be overestimated. Still, this is the best approximation of the time students spend on their daily activities .
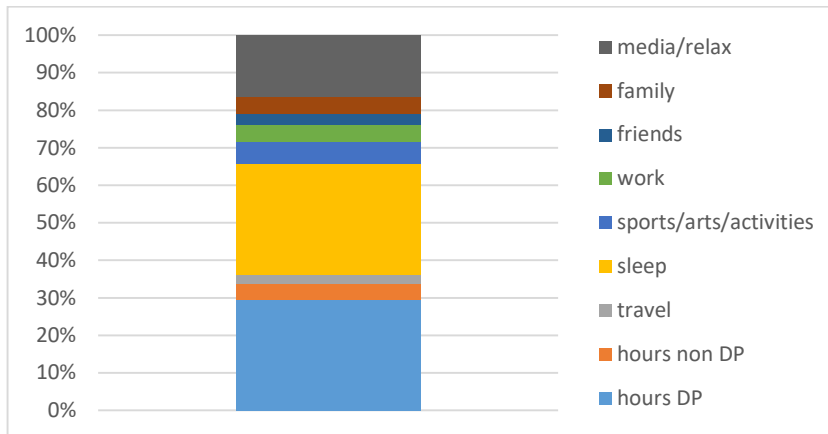
Figure 4.5: Students' time budget: hours spent on school work and other activities. Students wave 2 (N = 1,332)

- As **Table 4.3** shows, students spend on average between seven and eight hours a day on DP and non-DP school work, including both school and weekend days. On average they sleep between six and seven hours a night, also including nights during the school week and the weekend.
Students spend between three and four hours a day relaxing or using internet and (social) media. In addition, they are rather actively involved in sports, arts, and other activities, one and a half hours a day (9.5 hours a week).

- In **Table 4.4** students were split into four quartiles according to the time they spend on DP school work. Quartile 1 represents the 25% share of students who spend the least time on their DP school work, on average between three and four hours a day and includes students who do not follow the full DP programme. The fourth quartile consists of students who spend most time on DP, almost 10 hours per day, including weekend days. How does the time students spend on other activities differ between these four groups?

- As **Table 4.4** and **Figure 4.6** show, particularly the time for sleeping and relaxing is lower among students who spend more time on their DP school work. Ignoring the first quartile, which includes course students, in the second quartile, consisting of full DP students, the time spent on the DP is on average a bit more than six hours a day. Students on average sleep almost **seven hours a night** and spend about **four hours relaxing** and on (social) media. Students of the fourth quartile, sleep half an hour less each night, **about six and a half hours**, and spend three quarters of an hour less relaxing and on (social) media, **about three and a quarter hours** a day.
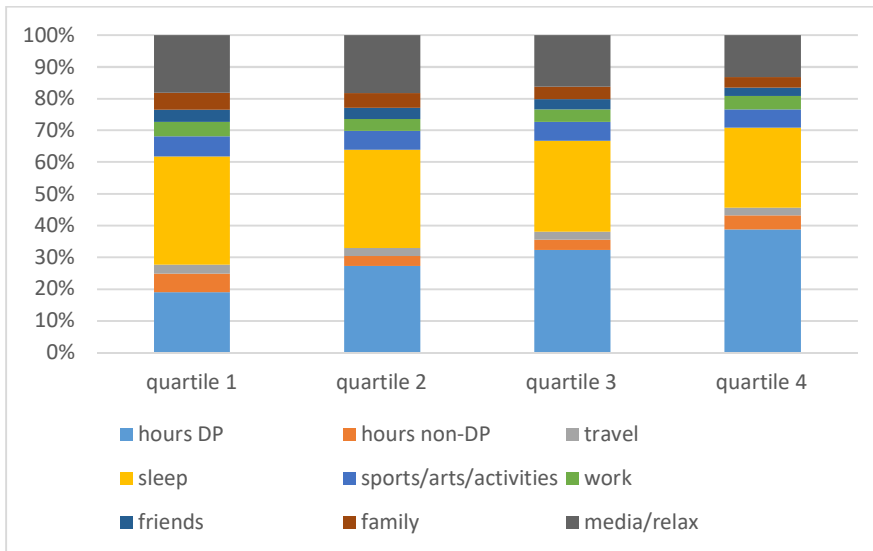
Figure 4.6: Students' time budget: hours spent on school work and other activities. Students wave 2 - quartiles according to hours spent on DP

Construction of time spent on other activities

Students in wave 2 (June) in year 1 of the DP were asked to estimate how many hours they typically spend per week on other activities besides school work, such as sports and cultural activities, (volunteer) work, social activities, internet and social media, as well as on the amount of sleep they get.

Time spent on activities outside school was asked in categories, ranging from no time spent to more than 20 hours a week. To arrive at a measure of hours per week of each activity, generally the class center was taken, for instance 'between 1 and 3 hours a week' was coded as 2, the lowest 'none' category was coded as 0, the upper category was coded as 20 hours a week. The two questions on sleep, on schooldays and on weekend-days, have categories varying between 'fewer than 3 hours per night' to '13 hours or more per night'. Following the same coding procedure to estimate the hours of sleep per night, a measure of weekly hours of sleep was constructed by summing sleep on schooldays five times, and sleep in the weekend two times. For the analyses, only students who answered the two questions on sleep were selected. Students who gave partial information on the time they spend on other activities were assumed to spend no time on that activity (missing values were coded as zero).

To describe the hours students spend on several activities, groups of activities were summarized and very high numbers of hours were recoded to the hours (whole hours) just above three standard deviations above the mean.

Chapter 5: THE WORKLOAD IN THE DIPLOMA PROGRAMME: VIEWS OF DP STUDENTS AND COORDINATORS

Contents

Chapter 5: THE WORKLOAD IN THE DIPLOMA PROGRAMME: VIEWS OF DP STUDENTS AND DP COORDINATORS

This chapter describes how students experience the DP workload. In our overarching model this is referred to as the subjective workload. Subjective workload is in many ways the core concept of the study. It has been measured qualitatively and quantitatively and both DP coordinators and DP students responded to several questions at all three waves of the data collection. The concept of subjective workload also has practical value in being a measurable concept for schools which can be tracked over time and can be used for evaluating which students might need help before they run into problems.

DP students' and coordinators' joint comments provide a varied and nuanced picture of the DP student workload itself as well as the factors that impact its manageability. A majority of both types of respondents evaluates the DP workload as heavy or too heavy, putting high demands on students. It is noteworthy, however, that external factors are mentioned as well that contribute to students' workload, such as university admission procedures or students having to follow courses from other programmes on top of the DP.

By the same token, a picture emerges that students experience the programme differently, with some students struggling while others are coping fine. Manageability is perceived as mainly being contingent on students' predispositions, such as self-management and language skills and to a lesser extent by the school environment.

5.1 The DP student workload according to DP students

- In answering the question '**what are DP students' ideas about the DP workload?**' we first analyse students' responses to open questions. Some of these questions were answered by a majority of students, often with elaborate opinions and suggestions. We have grouped these responses in the following categories (**P-factors**)[1]:
    - **Programme** (elements of the DP that are particularly challenging, eg, subjects or exams)
    - **Policy** (school policy, eg, coordination of assessments)
    - **Personal** (students' personal problems, eg, workload or stress)
    - **Pedagogues** (teacher practice in relation to workload)
    - **Peers** (pressure or competition among peer group)

- In the third wave of the student survey the DP students were asked the following question: '***What did you find the most challenging about the DP?***' This question is suggestive and meant to yield more elaborate information about what kind of challenges DP students experience. This information was analysed to get a better idea of what students experience as challenging and why. It is not used to derive conclusions on what is objectively most challenging for all DP students.

---

[1] Initially, a category 'other' was used in the coding, but this was not needed as virtually all claims fitted in the main categories.

- **1,363 students responded**. Most comments fall into the **Programme (212) and Policy (911)** categories, followed by Personal (about half as many). Few students commented about Pedagogues (teachers) (77) and fewer than ten about Peers[2]. Many students mentioned the final examinations as most challenging, specifically to finish exam papers in time. Furthermore, students single out internal assessments (IAs), the extended essay (EE) and theory of knowledge (TOK) as most challenging. Another frequent comment is that students feel they do not get the recognition they deserve from universities. In the following paragraphs the topics in each category are discussed in more detail.

Policy and Programme-related factors

- As students do not seem to differentiate much between the Diploma Programme itself and school and a lot of comments concern both, we have grouped comments from both categories together.

- Many students comment on the workload or the level of challenge of the programme. They report that the overall workload is the most challenging aspect of the DP. For example:
  - *The workload was immense and although the work itself is not always too difficult to understand, students are not given the time and or support / encouragement to be able to complete and understand the work well enough*.
  - *The workload was what made the IB so much more difficult.*
  - *The large workload was detrimental for my well-being*.

- According to some students, this also led to problems amongst teachers who had trouble covering all the course materials:
  - *In many classes (..) the large volume of material that had to be learned forced students and teacher[s] to skip aspects of materials and not truly understanding the material was disheartening and made doing work I did not understand difficult (..).*
  - *Some of my teachers did not teach us large sections of the section material. I had to teach most of my classes to myself*.

- Many students found the assessments and specific subjects of the Diploma Programme the most challenging.
  - *All of the EE and IAs. I think that it is too much for a student to conduct all those research. Personally [I think], an IA for each subject is a little bit too much…*
  - *Many of our IA and EE final deadlines were in the same week, which put a lot of stress and pressure…*

- Regarding DP subjects, students find it difficult to follow six different subjects at the same time, three of which at higher level (HL). Some say that the choice of subjects is too small and that they ended up following a subject at a higher level which they didn't have any affinity with. According to some students, some higher level courses are way too difficult:
  - *The higher level Sciences were intense*.
- HL Math and Chemistry are often mentioned as the most challenging DP subjects.

---

[2] There were so few comments relating to peers that this category is not further discussed in this chapter.

- Many comments concern deadlines that tend to be scheduled close together and that due to the amount of homework often piles up. One student found the internal assessments (IAs) most challenging:
  - *All the internal assessments were due at the same exact time. It caused high stress and emotional distress.*

- Another comment that was often made is about the content of examinations being unclear. Some students mention that they had to prepare all that was learned in the past two years but that a lot of that ended up not being relevant for the exam. Other students had problems with the marking of exams:
  - *The specificity of the mark schemes and how responses needed to perfectly fit the mark scheme or no points be rewarded.*
  - *The way marks are awarded. There are times where the questions are vague, but they expect a very specific answer.*

- Expectations about the assessments was also a challenging part of the Diploma Programme for many students. They mentioned that guidelines of how much and what to write were unclear and that:
  - *The connections that the student needed to make in order to answer the question was too abstract and not taught in class.*

Personal factors

- At the personal level students often relate issues mentioned in the previous paragraph to their mental and/or physical health. Many students say that the workload causes **pressure, stress and a lack of sleep**:
  - *The most challenging thing about the DP is juggling the demand of each individual subject (..). It eventually takes such a toll on an individual that even 'getting more sleep' can't really help all that much.*
  - *(..) There was a lot of coursework to complete and I would often sacrifice hours of sleep. The lack of sleep led to an increase in my anxiety levels(..) I think that the IB program weakened my self-esteem and caused me unnecessary pain (..).*
  - *It's not the quantity or quality, it's both, and I found it particularly challenging to produce regular high-quality work.*
  - *Students are forced to overwork themselves, spending most of their time and energy on the dp, often going weeks without enough sleep or food.*

- Other students say that they experienced difficulties with the **academic pace** and way of learning:
  - *(..) It is not the difficulty in itself, but it is hard to maintain a high level of academic progress at all times.*

- Also, **time management** was found challenging:

- o *Everything becomes difficult when you don't optimize your time. Whether we procrastinate and use time wastefully or doing things with time makes a difference in outcome.*

- Next to the question 'What did you find the most challenging about the DP?' students were asked how they perceived the level of difficulty of the DP and if they found that the academic level exceeded their capacities.

- Most of the students who did not find the DP difficult make similar comments about the workload as students who do find the DP difficult. They commented about the **workload** being **too high** and found the **amount of study material, the exam schedule, keeping up with deadlines and time management** the most challenging about the Diploma Programme. These students, however, also gave constructive feedback and are more positive about their own abilities in their comments than students who find the DP difficult:
  - o *At the start I had some difficulties managing my time and planning out my work schedule. However, I quickly was able to adjust and feel that I managed it better as time progressed.*
  - o *Getting the hang of thinking outside the box got easier.*
  - o *(…) I often felt as though it was hard for me to become an expert at material before we moved on to a new topic, which made review hard as the exams approached (..). In other ways, however, I did appreciate the challenges and breadth of material exposed to me.*

Pedagogues: the role of teacher practice

- Comments about teacher practice often relate to **inconsistency in teaching**:
  - o *(..) different TOK teachers I consulted had different and often contradictory opinions(..).*

- Another recurrent comment about teachers is the **lack of support** some students feel they get:
  - o *There is too much independence, and we are still in school. The DP makes it so that our teachers are not allowed to guide us enough and we are left on our own.*
  - o *(..) the 'find your answer yourself' attitude was challenging at times.*

5.2 The DP student workload according to DP coordinators and heads of schools

- **Coordinators** were also asked **how they perceived the workload of their DP students**. We have grouped their answers into the following categories[3]:
  - o (too) heavy
  - o somewhat too heavy (also heavy at times, for some students, for some years, for some of the courses belong to this category)

---

[3] When coordinators replied how *students* perceived the workload we did not take these answers into account. These comments form the fourth category, along with comments that weren't specifically about the workload or comments that were not clear on the workload.

- o not heavy (also classified as not heavy; appropriate, manageable, suits the standard of IB and a good preparation for university are)

- **1,318 DP coordinators and heads of school replied** to the question how they perceived their students' workload. **A quarter of the coordinators find the students' workload too heavy**. Almost half of our sample perceives the workload of the students as somewhat heavy, and one in five doesn't perceive the workload as too heavy, but as manageable and not heavy.

- In this paragraph we elaborate on the most common and interesting comments coordinators made about the workload.

Policy and programme-related factors

- Many coordinators comment on the amount of work DP students have to manage.
    - o *I see how stressed out the students are because there are so many requirements (projects, papers, IA's, reading) that are required in all of their classes (..).*

- According to some coordinators, the **workload has increased over the last years**:
    - o *I have been a IB coordinator for 8 years and I am concerned that the increased demands of the program are going to lead to serious mental health issues in our students, an inequitable work/life balance for both teachers and students (..) Students are constantly working; there is very little time to breath and enjoy the process. The internal investigation in the Sciences is far too onerous for [both] students and teachers.*

- Many coordinators say that the problem lies with the **Internal Assessments**:
    - o *There is no need to have an IA for every subject, this creates the greatest workload for both students and teachers.*
    - o *The obsession with recording everything is ruining the learning process for the students and hits especially hard with students who do not have English as their main language. Please reduce the IA demands.*

- Furthermore, coordinators say that the workload expands in the second year of the programme because **students do not have enough knowledge to take the IA's in their first year:**
    - o *It is difficult to start the IA's from September to March (first year) as students have not yet grasped the concepts and skills necessary for completion of IA's (..).*
    - o *Workload is about right; the more difficult part is that everything accumulates at the end of the IB DP. It would be less stressful if some of the assessments were to be examined/moderated earlier (by the end of Year 1 e.g.).*

- Another common answer about the amount of work is that **some subjects have a more extensive curriculum than others**, especially the **Sciences and Mathematic** courses would demand a lot.

- Some coordinators say **teachers create the workload** themselves:

- o *Each wants the best from all students and works beyond the deadlines to get more and more from the students*.

Personal factors

- Many coordinators say that only some students experience difficulties with workload and stress. The programme requires good **time management skills** and some students don't have these and tend to procrastinate. Some schools report to offer time management courses or to have dealt with this problem in another way:
  - o *We have a very innovative timetable and schedule, homework policy, pastoral system that greatly reduces any unnecessary stress and also all their workload is carefully scheduled throughout the whole DP. So that's why even our weakest students can stay on top of what they need to do in a timely manner.*

- Other schools don't seem to have solutions and say time management problems are due to the **students' age**; they are **too young for the high responsibilities**:
  - o *The DP really requires students to be emotionally and self-management wise very mature. In my experience, most 16 through 18 year olds would struggle with the demands of the DP.*
  - o *Though I believe the workload for students is completely manageable, it requires tremendous time management and manipulation of personal scheduling. This has become increasingly difficult for students to balance as they take on greater responsibilities in other areas of their lives (home, work, and the biggest-- extra-curricular involvement).*
  - o *(..) for the typical teenager, it involves a steep learning curve.*

- Other coordinators say the programme tends to attract **students who strive for perfection**, these students experience the highest workload and stress levels, according to them:
  - o *Those who hold themselves to impossible levels of perfection and can't allow for any error often struggle because the expectations of the programme are just too high for them to achieve perfection at all times. Those students who are a little more relaxed are often more successful because they can more readily roll through challenges without allowing it (to) negatively affect their emotional health.*
  - o *We are a school with students in the IB DP who are very, very high achieving students who overall are not going to simply be content with getting a 5 because they have always, always performed very well academically. They will always strive for a 7. And I think that IB attracts these types of students more than it is willing to admit.*

- Finally, according to some coordinators, students whose **first language is not the language of instruction** experience difficulties with the workload.

External factors

- Many coordinators say the high workload and stress are not due to the DP itself, but to other factors the school cannot influence. For instance, **in some countries the DP is not recognized by**

the national authorities, and students need to follow the national programme as well at the same time, which increases students' workload significantly.

- Another factor that influences the workload, especially in the second year of the programme, are the **admission tests for universities** students need to take:
    - *(..) since all my students need to take external standardized language test(s) and university admission tests, and we are only a 3-year program, the amount of work is truly astonishing*.

Expectations and programme reputation

- Some interesting comments were also made about the **high expectations of students, parents and universities**:
    - *The major issue is when parental expectation is not in line with student ability or interest.*
    - *(..) I remember when a diploma was an achievement in itself, at 24 or 25 points, but it seems like the psychological bar from students and universities has raised well above this, which rather crushes the potential pleasure out of the experience of studying subjects at IB diploma level.*

- According to others, the **perception of a high workload already exists before students even start the DP**. It's more the belief in a high workload, than the actual experience of it:
    - *I think IB itself comes with a level of intimidation and stress that is built-in before students ever start*.

- Finally, one coordinator comments that some **false expectations** arise **when students start to compare the DP with other programmes**, such as AP and A-Levels.

- DP coordinators were also asked to rate the workload of the DP and of other programmes, if provided at their school. **Figure 5.1** shows that on average DP coordinators think that DP students experience the workload of their programme as more heavy ($t(703)=32.1$, $p < .05$) and more difficult ($t(709)=29.2$, $p < .05$) than students of other programmes at their school. They also think the DP students are more stressed than students of other programmes ($t(697)=32.5$, $p < .05$) (see also **Appendix 5.1**).
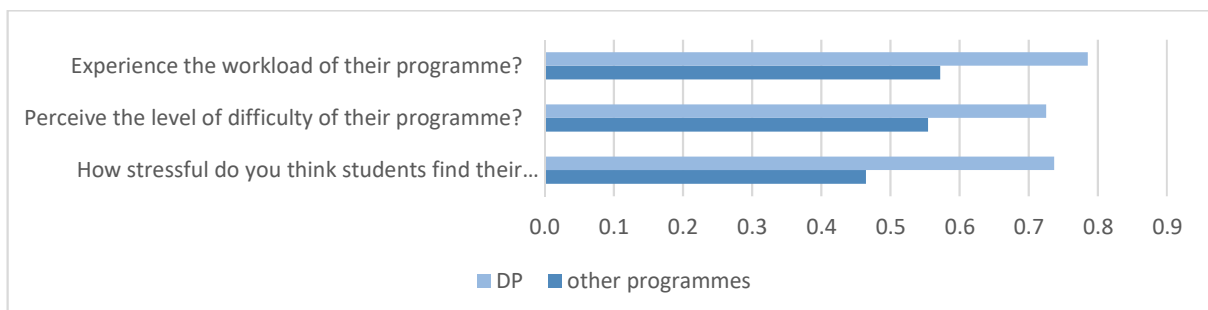
Fig 5.1: Perception of DP coordinators on how students experience the workload, the difficulty and the stress levels of students of DP and other programmes

Chapter 6: SUBJECTIVE WORKLOAD AND STRESS SYMPTOMS: RECIPROCAL CAUSATION

Contents

Chapter 6: SUBJECTIVE WORKLOAD AND STRESS SYMPTOMS: RECIPROCAL CAUSATION

This chapter examines the relationships between students' objective DP workload, how DP students experience their workload ('subjective workload') and the development of stress symptoms. We examine how objective and subjective workload as well as stress symptoms develop during the DP and construct a reciprocal effects model to disentangle the complex causal connections that exist between the three.

Figure 6.1 illustrates which part of the causal model is addressed in the current chapter.

Chapter 6

Objective Workload
Time school work
Selection of subjects

Subjective Workload
Experienced workload,
Manageability &
Level of difficulty
of DP subjects

Stress symptoms
Restlessness
Lack of control
Life dissatisfaction
Physical symptoms

Outcomes
Exam results
Dropout
University admission

P-factors:
Personal student characteristics
Policy of the school
Pedagogues: support or pressure
Parents: support or pressure
Peers: support or pressure

Figure 6.1 Causal model of the student objective and subjective workload and stress

We start by reporting on **the levels of subjective workload and stress symptoms** among DP students. On average, students appear to struggle with their DP workload all across their DP career, with little variation between waves. The stress indicators show rather high mean scores. There is no clear delimitation of when a student is stressed, but if we take the two extreme points of the answering scales as a benchmark, about **16% of DP students could be depicted as "severely stressed" already at the beginning of year 1 (wave 1)**. More worrying still may be that – in contrast to the subjective workload – this percentage clearly rises across waves, in particular from **wave 2 (19%) to wave 3 (35%)**.

There is a rather strong correlation between the subjective workload and stress: **a high workload tends to be reported by the same students who also report feelings of psychological discomfort (and vice-versa)**. Both subjective workload and stress symptoms are impressively stable *within* students over the three waves and thus have a strong personal basis: students who struggle with their workload or suffer from severe stress symptoms have a tendency to be vocal about their problems whatever the situation they are in. Moreover, during the two DP years the relationship between subjective workload and stress symptoms becomes stronger.

There is a somewhat weaker correlation between the objective and subjective workload: **students who work longer hours find their workload harder to manage and vice versa**. Like subjective workload, objective workload too is rather stable within students, meaning that the same students tend to make longer hours all the way through the programme, especially at the first two waves.

But to what extent **is the increase in stress symptoms among DP students caused by the DP workload**? Changes in subjective workload and changes in stress symptoms are found to be related by **reciprocal causation**: causality significantly flows in two directions, both **from subjective workload to stress** and **from stress to subjective workload**. These reciprocal effects are of about equal strength; there is no dominant direction. While complex as a result, the pattern of influences unequivocally warrants the conclusion that the **DP workload causes the development of stress symptoms** in students.

By the same token, changes in the hours students spend on the DP and their subjective workload are found to be related by reciprocal causation as well: causality again significantly flows in two directions, both from **hours spent on the DP (objective workload) to subjective workload** and **from subjective to objective workload**. Here the reciprocal effects, however, are not of equal strength; surprisingly, the dominant direction is from subjective to objective workload. This could mean that the more students evaluate their workload as unmanageable, the more hours they spend doing their homework.

6.1 Subjective workload: measurement and development during the DP

- *Subjective workload* is the pivotal concept in the DP Student Workload study. As implied in the overarching causal model, the objective challenges of the DP workload are conceived to disturb students' well-being, once these students begin to perceive their workload as 'unmanageable'.

> **Measurement of subjective workload**
>
> To measure students' evaluation of the DP workload, we have asked them to react at each of the three waves to **four statements[1]**:
> - "Overall, how do you perceive the level of difficulty of the Diploma Programme?" *(items a23, b12, c06, answering categories: much too easy .. much too difficult)*
> - "The academic level of the DP exceeds my capacity."*(strongly disagree .. strongly agree)*.
> - "How do you perceive the DP workload?"*(much too light .. much too heavy)*.
> - "How manageable do you find the DP Workload?" *(absolutely unmanageable .. absolutely manageable)*
>
> **Table 6.1** shows the descriptive statistics for these four indicators, which are identical across waves[2]. All answers are recoded into 0..1 scales (midpoint .5), and the answers to the fourth item have been reversed, so that the means can be compared between waves. It turns out that **according to all four indicators the subjective workload is, on average, rather stable between waves**.
>
> **Table 6.2** shows how these four items are correlated, within and between the three waves of data-collection. Note that, due to reversed coding of answers to the fourth statement, **all point to the direction, of an unmanageable workload**. There are two important observations on the pattern of the relationships between the four measurements (correlations) of subjective workload:

---

[1] Note that we have used multiple indicators to measure this pivotal concept, with variations of wording of both questions and answers, and different directions ('polarity') of the resulting indicators.

[2] As general in our report, the three waves are marked by a leading a, b and c.

> o The correlations are rather strong, both within and between waves. The strong within-wave correlations (yellow boxes) denote that all four statements measure the same underlying concept, as intended.
> o Correlations between waves are also rather strong, but clearly more pronounced when the correlation refers to the same items (ie, the brown coloured diagonal cells). These indicate the presence of strong so-called 'methods' effects.

- A further analysis of these subjective workload data is represented by the means and correlations in **Table 6.3**, where the four items have been summated by their within-wave average score. The above mid-point means (just under 0.60) indicate that **on average, DP students find their workload hard to manage.** The means of these average scores are very similar which indicates that this was as true in the first wave as it was in the last: **there is no change between waves in how manageable the students find their DP workload**.

- The correlations between the three waves in **Table 6.3** are very strong (> 0.5). This indicates that the students' judgements are above all strongly based in their individual personal situation: students who struggle with their workload in the beginning are to a large extent the same students who say they do one and two years later. The correlations, nevertheless, also indicate some degree of 'true' change, not only because they are far from perfect (if they were, they would be 1), but in particular because the correlation between the non-contiguous waves 1 and 3 (0.53) is clearly lower than that between contiguous waves 1 and 2 (0.68) or wave 2 and 3 (0.59). **Such a correlation pattern can only occur when the scores of subjective workload truly change between waves.** As we can learn from the stable means, these changes are not related to the workload becoming more pressing during the DP years. Rather, some students change in the direction of a lower subjective workload, while other students change in the other direction.

- **Figure 6.2** shows the development over time of the four indicators of subjective workload, including retrospective information when students were asked to look back on year 1 and year 2 (not for all indicators). The four indicators were – for comparability - recoded into a 0-1 range, from the least heavy (0) to most heavy subjective workload (1). Two main conclusions can be drawn from **Figure 6.2**.
  - o The subjective workload remains rather stable throughout the course of the DP programme, from the beginning in year 1 until the end of year 2. This holds for all four indicators equally. An exception is the very beginning of the DP of which students retrospectively report a relatively low subjective workload.
  - o Despite similar developments in subjective workload, students consistently report a higher subjective workload in terms of perceived heaviness and difficulty of the DP than in terms of the manageability and the degree to which the DP exceeds their capacities. At the end of year 1 for instance (see **Appendix 6.1**), 43.4% of the student perceive the DP workload as (much) too heavy, but only 27.6% think it is too difficult. Still, 16.1% of the students say that the DP is (absolutely) unmanageable, and, 14.1% say it exceeds their capacities. To sum up, students perceive the workload of the DP as heavy, but not always to an extent that cannot be managed or exceeds their capacities.
  - o The rather stable development is confirmed by the development in the first year of another indicator of subjective workload, test anxiety (**Figure 6.3**). Around half of the students say to worry about the school tests and grades, both at the beginning of DP and at the end of year 1 (see also **Appendix 6.2**). Students are a little bit more worried in the beginning: there is a slight,

but statistically significant, decrease over the course of year 1 (based on the average of the five indicators, Cronbach's alpha at the two time points respectively is .88 and .86). Still, there is large stability, correlation .73, meaning that students who worry in the beginning also worry at the end of year 1, and vice versa.
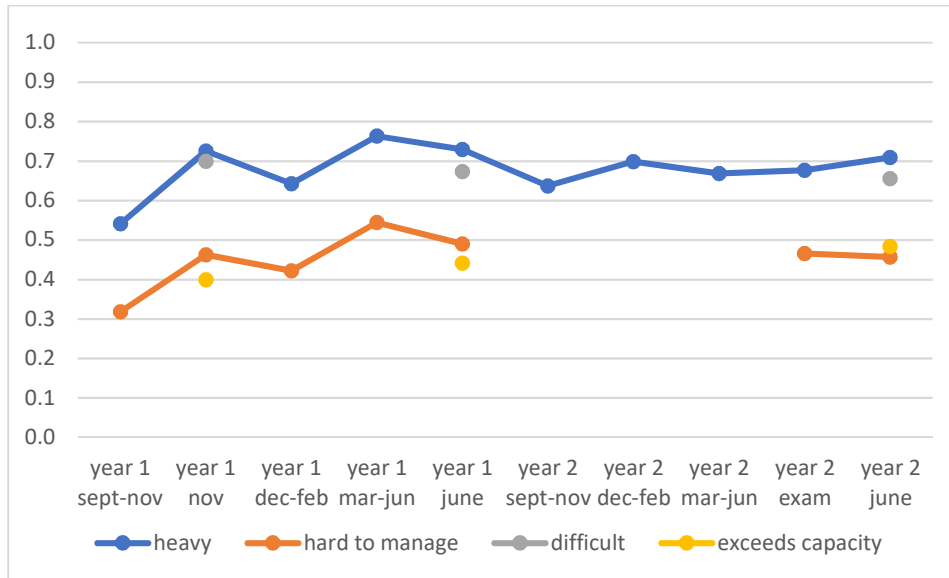
Figure 6.2: Subjective workload over the course of DP
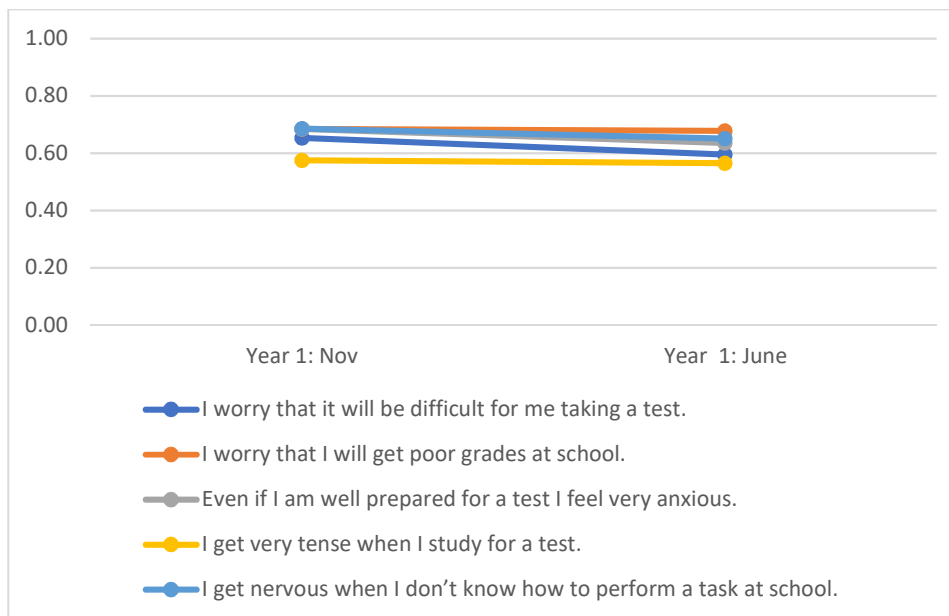
Figure 6.3: Test anxiety: worrying about tests and results

Missing values

Table 6.3 displays another challenge to the analysis of multi-wave panel data: the occurrence of **missing data, due to non-participation** in one or two waves. Altogether, 3,565 students took part in our surveys, but their participation per wave dropped from over 2,458 in wave 1, to just over 2,000 in the third wave. However, only 1,063 students took part in all three waves, whereas the number of students in two waves hovers around 1,000 (see also **Appendix B**). Such patterns of missing values are undesirable, but

common in multi-wave panel research. They are a challenge, because we want to use all available data, which is possible via the use of *Full Information Maximum Likelihood estimation (FIML)*.

One important condition for the valid use of this technique is that the panel attrition is not (too) selective. While it is impossible to determine the selectivity of the non-response (relative to the true population), **Table 6.3** shows that the correlations in the group that participated three times (listwise) are not different from the correlations in all available data (pairwise). The complete cases sample seems in no way different from the available cases sample. This observation makes it plausible that ML estimation is a valid way of handling the missing data and making the most of the data we have.

Simplex or quasi-Markov model

Figure 6.4 explains why the correlation pattern in **Table 6.3** indicates true change. It illustrates the so-called simplex (or: quasi-Markov) model of three-wave panel data, distinguishing between the true (or: 'latent') scores and observed scores of subjective workload. This makes it possible to separate genuine ('true') change in the way students experience the workload (coefficients **aa**) from 'false', measurement-induced effects of the question or response type change (coefficients **bb**). The algebraic expressions that are implied by the simplex model, indicate that the correlation between further removed waves will only[3] increase if **a** < 1.0, ie, no perfect correlation at the true score level.

Our first estimate of stability of the subjective workload are the sub-diagonal correlations in **Table 6.3**: 0.68 and 0.59. These observed correlations must, however, be corrected for measurement error **b**, as portrayed by the model in **Figure 6.4.** If we estimate the model with structural equations (**Table 6.4a**), the estimates of the true (ie, error-corrected) change parameters are higher: 0.88 and 0.80; these effects are weakly (p < .07) different from one another, which may indicate that the changes in the second year of the DP are stronger than those in the first year. **Table 6.4a** shows the pooled estimate (0.84). The measurement reliability b is estimated at 0.88. This last number can be regarded as an alternative estimate of the internal consistency estimate of measurement reliability reported in the third and fourth columns of Table 6.4.a.
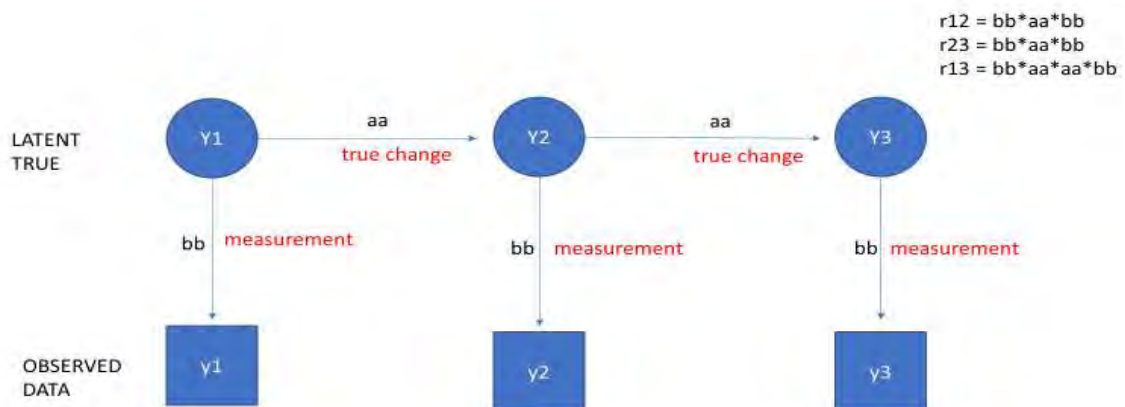
Figure 6.4 Separating change and measurement unreliability in three-wave panel data using a simplex model

---

[3] Alternatively, the correlation pattern could arise when bb (the measurement coefficient) is not stable, but measurement would be worse at wave 1 or wave 3. However, the internal consistency estimates of measurement reliability (Cronbach's alpha) make this scenario highly implausible.

There is still a concern about the method effects that are so clearly visible in **Table 6.2:** part of the stability between waves is due to the fact that the students – even after a year of providing a previous answer –, have a predictable way of responding to one item or the other. This *response tendency* can hardly be attributed to memory effects; it is likely mostly due to a different understanding of either the wording of the items and/or the answers. We can remove this bias by controlling for these method effects in a different type of structural equation model, known as the *Multiple Trait, Multiple Methods model (MTMM)*. This MTMM model (see **Figure 6.5**) leads to estimates of the degree of stability which are slightly lower than the simplex analysis above and still very high (awl -> bwl: aa21 = 0.81 and bwl -> cwl: aa32 = 0.71). Once again, this implies that students' evaluations of the DP workload are very stable across the three waves, but at the same time change as a response to external circumstances. These changes seem more pronounced in year 2 than in year 1 of the DP.

For the sake of clarity only two observed indicators are shown
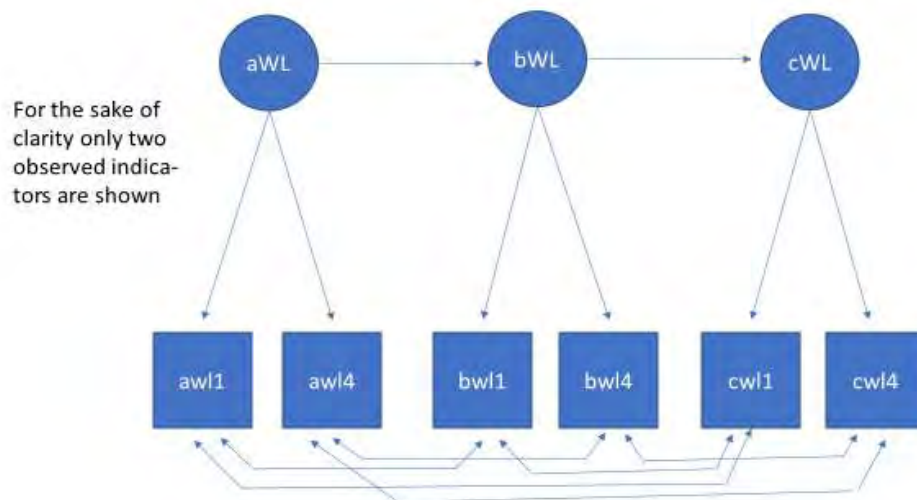
Figure 6.5: An MTMM model to separate true change and measurement unreliability in three-wave panel data.

6.2 Stress symptoms: measurement and development during the DP

- We define stress symptoms as any feeling of psychological discomfort that students report in response to four widely used and validated diagnostic instruments used in this study.

Measurement of stress symptoms

The four instruments we used to measure stress symptoms are:
o *Student Life Satisfaction Scale (SLSS)* (Huebner, 1991)
o *Perceived Stress Scale (PSS)* (Cohen, Kamarck, & Mermelstein, 1983)
o *The World Health Organization Quality of Life Scale* (WHOQOL Group, 1994)
o *The Short-Form-12 Health Survey (SFHS)* (Ware, et al., 1994).

For our own use, we have renamed the four instruments in terms of stress symptoms:
o SLSS   → Stress 1: Restlessness
o PSS   → Stress 2: Lack of Control
o WHOQOL   → Stress 3: Life Dissatisfaction
o SFHS   → Stress 4: Physical Discomfort

An overview of all the items used is given in **Table 6.5a**. All four scales consist of multiple statements, some wording in the direction of discomfort, others rather in the direction of well-being and satisfaction. The answering modes are either five- or six-point scales, but have been recoded to a 0..1 range. We have chosen these four instruments because they are widely used in psychological research, but also because they cover different aspects or dimensions of feelings of stress, which is reflected in our renaming: restlessness, lack of control, life dissatisfaction and physical discomfort.

Notice that none of the statements makes causal attributions, ie, there is no reference to the DP workload as being the cause of stress symptoms development – it may very well be that students report high levels of stress that are induced by other circumstances, such as personality, life events or their family situation. Because of this, it becomes an empirical issue to decide whether and how stress symptoms are related to the DP workload. The lack of causal attribution is the crucial difference with the subjective workload instruments discussed in the previous paragraphs, which do refer directly to the DP workload. This difference between the subjective workload measures and the stress symptoms measures makes it possible to determine whether the workload causes stress, independently of what the students think about this relationship.

Before analysing the relationship between stress symptoms and subjective workload, we examine the quality of our measurements of stress symptoms by addressing the standard methodological questions about their validity (which dimensions of stress do they measure?) and reliability (how stable are the instruments in measuring what they measure)?

The dimensionality of the data has been addressed using factor analysis, for which we chose principal axis factoring in SPSS. Commonly used standards in factor analysis to determine the number of underlying dimensions refer to the size of eigenvalues, which indicate how much additional variance is explained when one assumes that one or multiple dimensions of stress exist. However, in our data the use of factor analysis is complicated by the fact that the measures are repeated at each wave. This makes it necessary to factor-analyse the data separately for each wave, which may and does result in different solutions between the waves. **Appendix 6.3** displays the **pattern of factor loadings** obtained in each wave. These numbers are the correlations between the assumed latent dimensions (fixed at three in each wave) and the observed indicators.

The first and foremost conclusion of the factor-analyses is that the stress symptoms measures strongly manifest **a single underlying dimension**. At all three waves the first eigenvalue covers over 35% of the common variance, and the subsequent dimensions explain only minor parts of the common variance (5% or less). While this result does no exclude that some multi-dimensionality is present in the data, **it certainly justifies that the data can be summarized in a single scale**. A related – but not identical – justification is that when subdimensions are distinguished (see below), scales that represent these dimensions, correlate strongly: between 0.5 and 0.6.

Determining the number and nature of subdimensions in the stress symptom data is much harder than establishing their strong commonality. After a considerable amount of exploration, we decided that **not four, but three factors are the best representation of the sub-dimensionality of the data**. Broadly speaking, our data suggest that the RESTLESSNESS and LACK OF CONTROL scales cannot be distinguished; they appear to measure the same phenomena, despite coming from different literatures.

o   The **clearest subdimension** present in the stress symptom data is the LIFE DISSATISFACTION scale, because it emerges as a separate scale in all three waves. Several items of the LACK

OF CONTROL scale appear to be more associated with the DISSATIFACTION dimensions, but this does not hold in all waves.
o The **second clearest subdimension** in the stress symptoms data is the PHYSICAL STRESS scale, that emerges as a separate dimension in the second and third wave, but much weaker in the first wave, in which these items are not distinguishable from the RESTLESSNESS scale.
o Third and finally, the items initially conceived to measure RESTLESSNESS and LACK OF CONTROL, behave rather irregularly in the dimensional analyses, as some are associated with the PHYSICAL DISCOMFORT dimension in the first wave, but not in the subsequent two waves. Notice, however, that the irregular pattern of factor loadings does not imply that these items do not measure stress: they only fail to measure clearly delineated subdimensions, which may also be taken as evidence that a very strong single dimension underlies the stress symptom data.

Another important question answered by factor analysis is whether all individual items are best treated as being part of the scale according to the conceptualization or are more strongly related to subdimensions that they are not expected to represent. **Appendix 6.3** displays several of such cross-loadings, which in principle indicate that such items should not be included in a scale, when analysing subdimensions.

At this point one clear conclusion is that analysing the data using a single dimension is the best solution – and we suspend judgement of the number and nature of possible subdimensions.
After having examined the validity of the stress measurement, we address the **reliability** of the constructed sub-dimensions and the overall measure. Unreliability of a measure can be defined as the instability of a measure in the condition of no true change. If no true change had occurred, all non-perfect correlation would be due to measurement unreliability. Like with subjective workload, we estimate reliability with two different methods.
o First, we calculate the well-known reliability coefficient (Cronbach's) alpha, using the internal consistency method. This method assumes rather innocently that no true change occurs within waves (ie, within the same survey interview) and that the lack of perfect correlation between parallel indicators is thus entirely caused by random measurement error. However, the tacit assumption of the internal consistency method is uni-dimensionality: all indicators represent the single underlying construct to the same degree. Unfortunately, this assumption does not necessarily hold.
o Our second method exploits the presence of three-wave panel data and estimates simplex reliability, which was already introduced above when addressing the reliability of the measurement of subjective workload.

Table 6.4b shows the results of the reliability analysis for the over-all stress scale. The estimates of reliability for the stress scale are higher than for the subjective workload, which stands to reason, as many more indicators (28) are used to measure stress than to measure subjective workload (4).

• **Table 6.5a** (see **Figure 6.6a**) shows the means of all individual stress indicators as well as the scale means. All responses to the stress items have been recoded to a 0..1-point scale, such that they have values ranging from 0 to 1. Notice that all indicators have been (re)coded in the direction of stress: **higher scores indicate that students experience higher levels of stress.** The standardization to a 0..1 scale makes all answers comparable to one another. While there are no absolute criteria to determine when a student is "stressed" and when not, **Table 6.5b** provides the percentage of

students reporting scores of 0.5 of higher (the midpoint of the answering scales) for each of the statements and at each wave, as an indication.
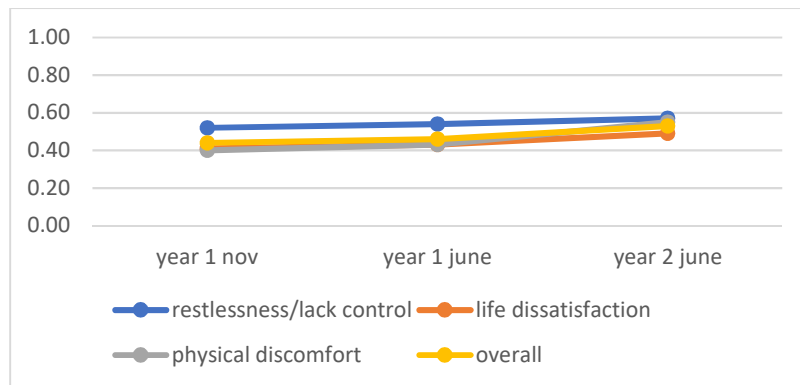
Figure 6.6a Stress levels over the course of the DP

- **Figures 6.6b** shows the development of stress over the course of the second year of DP, in relation to the objective and subjective workload. At the last wave, students were asked to look back on their second year of DP and report their stress levels as they perceive stress. Although only one indicator is used, in contrast to the validated measurement used in the analyses, a similar development can be noticed: an increase over time, also within year 2.
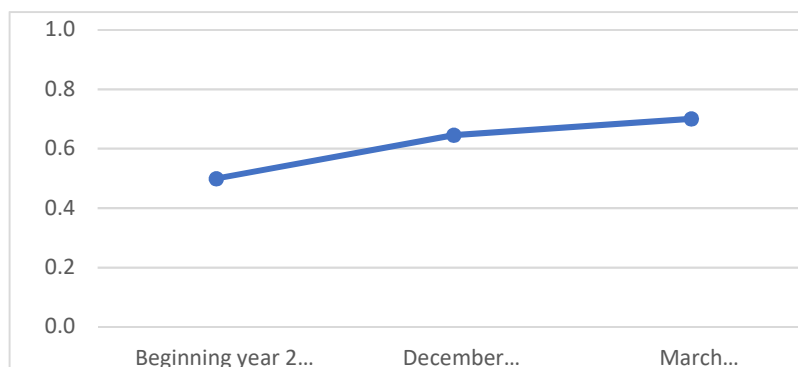
Figure 6.6b: Stress in retrospect: On average, how stressed would you say you were during your second year?

- Conclusions regarding DP students' stress levels:
  - Students report **high levels of stress**, with average scores being just above (restlessness and lack of control) or not far under (life dissatisfaction and physical discomfort) the midpoint of the response scales (0.5).
  - There is a clear **rise of stress levels over the course of the programme**, for all scales and almost all indicators. Notice that the statistical tests that compare students between waves (calculated on the condition that they participated in both compared waves) are all widely significant.
  - The **four scales can be ranked with respect to their mean values** with more students reporting mild stress symptoms, which we labelled 'Restlessness' (increase from 28% to 41%), fewer 'Lack of Control' (increasing from 21% to 32% severely stressed), fewer 'Life Dissatisfaction (increasing from 17% to 29% severely stressed) and still fewer reporting Physical Discomfort (such as lack of energy, sleep problems and physical pain) (rising from 10% to 34% between wave 1 and wave 3). This suggests that the distinction between the scales are foremost along

a **severity of stress** dimension and can be interpreted as a cumulative scale. High scores on the overall stress scale may thus imply not only that those frequently students report stress symptoms, but also that they report more serious levels of stress.

o As **Table 6.5b** shows, overall, taking all dimensions of stress into account, the percentage of DP students who are severely stressed rises from 16% in wave 1 to 35% in wave 3.

6.3 Reciprocal relations between subjective workload and stress

- Our next step is to investigate the relationship between the workload as students perceive it and their reported stress levels. The research question to be answered in this analysis is to what extent the relationship between stress symptoms and subjective workload can be given a causal interpretation, and if so, in what direction the causality flows. Does the workload cause stress or are existing feelings of stress projected in workload reports, or both?

- A three-wave reciprocal causation panel model (**Figure 6.7**) can be used to resolve issues of causality in non-experimental designs in which any association between variables cannot simply be interpreted as cause and effect. If we consider the association between subjective workload and stress symptoms, there are three alternatives to account for this correlation, which is strong and statistically significant in all waves:

  o **Causal**: workload causes stress symptoms. It is plausible that the relationship is positive: a higher (= more unmanageable) workload leads to more stress.

  o **Reversed**: exogenously occurring stress causes students to experience their workload as being unmanageable. In this scenario, having trouble coping with the demands of the DP becomes a by-product of independent stress development, which is projected on to the way students experience the DP workload.

  o **Spurious**: the association between workload and stress symptoms is produced by pre-existing conditions, which could influence both the choice of workload and the level of stress reported. The influence of such 'confounding' may be positive or negative.

Figure 6.7: The reciprocal causation model for three-wave panel data

---

Examples of confounding factors

Examples would be the confounding influence of gender or academic ability: if female students or academically more challenged students are less able to cope with stress and choose lower workloads than male students or academically gifted students, this would produce a spurious, non-causal correlation between workload and stress symptoms. This correlation could be negative or positive, depending upon the direction and strength of the common causality. When one wants to rule out such a common cause interpretation, a major problem is that confounders may be – and likely are – unobserved, and unknown factors affect the relationship in unknown directions.

- Multi-wave panel data such as collected in the DP Student Workload study can resolve the issues of reversed and spurious causation (and hence isolate true causation), due to two related features.
  - First, a **multi-wave panel design makes it possible to identify the direction of causality**, which may still be of a reciprocal nature. Traditionally, this is most often done by a *cross-lagged* design in which characteristics in wave t are caused by characteristics in wave t-1. However, a more direct approach is the *simultaneous* reciprocal causal effects model, in which the causality is estimated within each wave, and the identification of the two arrows is obtained by excluding cross-lagged effects.
  - Second, a **multi-wave panel design allows for the control of previous levels of subjective workload and stress**, and hence of all situations that gave rise to subjective workload and stress in the previous wave – **even the ones that were not measured**. The multi-wave panel model is a 'within-student' design that estimates how changes in subjective workload are related to changes in stress.

Structural equation model

The basic structure of this model was shown in Figure 6.2 (see above)**.** It can be estimated with a structural equation method, for which we used Stata 14. The specification of the model contains the following features:
  - While the multi-wave panel model can be estimated with only two waves, we have access to three-wave data. When estimating the model, we assume that its structure is **stationary** between the two transitions. This makes it possible to constrain similar parameters between waves. Such pooled estimation contributes much to the power of the model.
  - We assume that wave 1 variables can only affect wave 3 variables via wave 2 variables. This **simplex (or Markov) assumption** makes the model also an **instrumental-variables (IV) model**. IV models are widely used in econometrics as an alternative to causal analysis via experimental designs. While the only-indirect-effects assumption is often quite problematic, it is very natural and plausible in multi-wave panel data, in particular when combined with measurement error correction.

- **Table 6.6** shows the correlations between subjective workload and stress symptoms at the three waves (labelled a b c again). The subjective workload measure is the average of the four manageability items of which the means were reported to be high but relatively stable earlier in this chapter. For stress symptoms the table reports correlations for the three subscales found in the dimensional analysis above, but we will concentrate on the combined scores on these scales,

which were shown to be strongly related by a single common factor. Note that the means of subjective workload and stress symptoms play no role here: the correlations indicate how differences between students in subjective workload and stress are associated within and between waves. There are several important observations to be made on these correlations:

o First, **the measures of subjective workload, as well as the measures of stress symptoms, are very strongly related between waves**, with correlations ranging between 0.53 and 0.75. Remember that the three waves were spaced about a year apart, which rules out the possibility that the strong correlations are somehow artificially inflated by memory effects or attempts of the students to present themselves more consistent than they actually are: it is highly unlikely that any student would remember his or her biased answers after a year interval. The strong between waves correlations of the same concepts rather indicate that **much of the between-student variation in subjective workload and stress is a stable individual condition**. As we will see, this stability does not rule out that subjective workload and stress are affecting one another or are being affected by other conditions – but it is necessary and possible to diagnose such variations as small fluctuations in a relatively stable pattern, and thus use the stable pattern as a control. This is what the multi-wave panel model accomplishes.

o A second observation on **Table 6.6** is that the correlations between waves of both subjective workload and stress are consistent with the assumed simplex (or 'hidden Markov') model: the correlations are strongest between contiguous waves and somewhat weaker between wave 1 and wave 3. This is an indication that there is **some degree of within-student change**. If there were no changes within students, the correlations between waves would be equal between all pairs (ab: 1 and 2, bc: 2 and 3, ac: 1 and 3). In a simplex model with perfect measurement the following would hold: r13 = r12*r23, but this is also not perfectly met. This implies our instruments contain some residual measurement error and are not perfectly reliable – this should come as no surprise as the internal consistency estimates reported above were not perfect either.

o A third observation on **Table 6.6** concerns the correlations between the workload and stress measures that are reported in the off-diagonal quadrant. Again, these correlations are strong to moderately strong, although not at the levels of the within-concept correlations (in the yellow quadrant). A striking feature of the correlations is that they are noticeably stronger on the diagonal (marked brown) than off this diagonal. The relationship also appears to increase with wave and is strongest in wave 3 (0.55), as compared to 0.47 in wave 1. This pattern implies that the problems students have with their **workload and stress symptoms become more aligned during the two years of the programme. This may indicate an accumulation of stress over time.**

o A fourth observation on **Table 6.6** is on the off-diagonal correlations between workload and stress between different waves. These correlations are to a large extent symmetric – which they do not need to be.
Notice, for example, that the correlation between stress at wave 1 (aStress) and workload at wave 2 (bWL) is 0.43, while the correlation between stress at wave 2 (bStress) and workload at wave 1 (aWL) is only slightly stronger (0.47). While the pattern is largely symmetric, it is important to note that all the correlations below the brown subdiagonal (in which stress comes after workload) are stronger than the correlations above the brown subdiagonal (in which stress precedes workload). This asymmetry may have consequences for the estimation of reciprocal effects.

o A fifth and final observation on **Table 6.6** refers to the number of cases (N) as reported in the lower part of the table. The N of cases is different for all individual correlations. It is naturally highest for the within-wave correlations and lowest for the wave 1 – wave 3 correlations. The

correlations thus refer to all available data, an approach also known as pairwise deletion of missing values.

- **Table 6.7** shows the estimates of the reciprocal effects of subjective workload on stress and vice versa for the overall stress measure as well as its three subdimensions. While there is some variation in effects between the particular way the stress symptoms are measured, the important finding here is that **causality flows in both ways and is of about equal strength.** Students' subjective workload cause stress symptoms (0.201) and stress symptoms cause increases in students' appraisals of their workload (0.188). Both effects are strongly statistically significant, which means that the model rebuts the interpretation that the strong correlations between subjective workload and stress symptoms are entirely caused by either pre-existing individual differences (ie, spurious causation due to for instance gender) or by reversed causation (stress leading to students struggling more with their workload). In reference to the latter, note, however, that we are not able to determine a dominant direction in the reciprocal causal relationship. The estimated effects are almost equal in size, and it would require much stronger data (either by having a larger sample or by having more than three waves) to firmly establish a conclusion that one direction is stronger than the other.

6.4 Reciprocal relations between subjective and objective workload

- The objective workload caused by the DP can also be measured by the time students invest to absolve the programme's requirement. The particular advantage of this measure is that it allows us to look at developments in objective workload within the DP career and analyse how objective and subjective workload interact. We would expect a relationship between the two not only because the amount of objective workload affects the students' subjective report of it, but also expect reversed causation: when a student feels pressured by the programme, this may lead them to invest more time in their academic career. And again, like with subjective workload and stress, the objective and subjective workload might be related because prior conditions cause both. The reciprocal causation model allows us to separate these possible influences and compare the strength to the reciprocal effects between subjective workload and stress, reported above.

- We have asked the students to report their time expenditure in each of the three waves. For the first two waves the question format was identical and asked for three compartments of the time budget: classroom time, homework time and other tutored activities, which includes the participation in cram schools etc. In the third wave the questions were asked in a different format, as we sought to make an inventory of the time budget in three different episodes in year 2. In wave 3 these questions had a retrospective format, as the students answered them only after the exams – and their concurrent time investments would have been irrelevant.

- **Table 6.8** report the valid N and means of the measures. The means have been obtained by converting the original intervals using mid-category values and refer to average hours per week. Results are very similar for the first two waves. Summed over the three categories, DP students say they spend over 47 hours on their academic work, with some 57% of the time being spent in classes. To this time must be added the time devoted to other programmes, but it should be noted that a considerable number of students are not involved in such activities and chose not to answer these questions on 'other programmes', instead of marking 'none'. If we take into account such

answers as zeroes, the average amount of time spent on other programmes is negligible, although it may matter quite a lot for individual students, who are involved in two programmes.

- The means for the third wave that report on year 2 retrospectively, are not very different from the first two waves, with only a slight decrease of classroom presence in the March-June period and a concomitant rise of the amount of homework in the same period. In order to bring the information to match the measures for the first two periods, we have averaged the time expenditure for year 2 over the three periods.

- **Table 6.9** displays the correlations between the measures of time investment (hours spent) in each of the three waves, together with measures of subjective workload as introduced above. The first finding that strikes, is that the correlation between the time expenditure and subjective workload, is only moderately strong at best: the correlation fluctuates between 0.25 and 0.28, and slightly decreases over the three waves. This suggests that differences in time expenditure cannot be a major reason why some students find the programme unmanageable and others not. A second observation on the correlations in table 6.9 is that the correlations between time expenditure across waves are much weaker than those found for subjective workload (and stress for that matter). Still, the correlations between adjacent waves are strong (0.54 and 0.43), while the correlation between wave 1 and wave 3 is weaker (0.34). As explained above, this information can be used to estimate reliability for the time measurement, using the simplex model. This reliability is still found to be over 0.80, which implies that the weaker correlations between time measures are mostly due to true change, not to measurement error. This at the same time implies that the only moderately strong correlations between subjective workload and time investment are not due to unreliable measurement of the time-investment variable.

- **Table 6.10** shows the results of fitting the reciprocal causation model to this data. These indicate that the reciprocal relationships between time investment and subjective workload are weak, but statistically significant in both directions. However, comparison of model 2 and model 3 suggests that the feedback effect of subjective workload on time spent is stronger than the more obvious effect of Time → Subjective workload. Model 4 shows this, when both effects are estimated simultaneously. In fact, the Time → Subjective workload is only marginally significant. On the other hand, a test that the reciprocal effects are of equal size cannot be rejected (L2 = 70.1, df= 10, which is not significantly different from the L2 for model 4.). In other words, the more students perceive their workload as unmanageable, the more time they end up spending on their homework, rather than the other way around.

References

Antonakis, J., Bendahan, S. Jacquart, P. & Lalive, R. (2010). On making causal claims: A review and recommendations. *Leadership Quarterly* 21 (6): 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010.

Billedo, C.J., Kerkhof, P., Finkenauer, C. & Ganzeboom, H. (2019). Facebook and face-to-face: Examining the short- and long-term reciprocal effects of interactions, perceived social support, and depression among international students. *Journal of Computer-Mediated Communication* online (January): 1–17. https://doi.org/10.1093/jcmc/zmy025.

Finkel, S. E. (1995). *Causal Analysis with Panel Data*. Thousand Oaks CA: SAGE.

https://doi.org/10.4135/9781412983594.

Vaisey, S. & Miles, A. (2014). What you can — and can't — do with three-wave panel data. *Sociological Methods & Research*, 1–24. https://doi.org/10.1177/0049124114547769.

Chapter 7: THE ROLE OF P-FACTORS: PERSONAL BACKGROUNDS, PARENTS, PEERS, PEDAGOGUES AND POLICIES

Contents

Chapter 7: THE ROLE OF P-FACTORS: PERSONAL BACKGROUNDS, PARENTS, PEERS, PEDAGOGUES AND POLICIES

In the previous chapter we have seen that an increase in subjective workload, the sense of having trouble meeting the requirements of the DP, leads to an increase in stress levels. This process need not be the same for all students. In this chapter we examine for which students in particular the feelings of being overwhelmed with their school work turn into stress. We explore how the impact of the DP workload on subjective workload and stress varies with the so-called 'P-factors', as introduced in Chapter 2: Personal background, Parents, Peers, Pedagogues, and the Policies of the school. Each of these factors may reduce or amplify the impact of students' (subjective) programme workload on symptoms of psychological discomfort.

The "P-factors" can be described as follows:

- **Personal background** – in particular students' academic abilities. In Chapter 3 we have established that two dimensions underlie the statements that students have provided about their abilities before entering the DP. We distinguish **general academic ability** and **language ability**. This latter dimension stands out also because for quite a few of the DP students the language of instruction in not their first language.
- **Parental background** – in particular whether the stress development process is aggravated by **lower socio-economic status [SES] and incomplete family structures**. At the subjective side we have included students' reports about emotional and school **support from their parents**, as well as the **academic aspirations** that these parents hold for their children.
- **Pedagogues** – the extent to which stress development depends on **teacher-student relations and teaching quality.**
- **Peers** – whether stress development is dependent upon the nature and functioning of the student body in DP schools. In particular, we will address the rather popular hypothesis that the development of workload-based stress is dependent upon **school pressure, ie, a school climate in which teachers and peers push students to excel**.
- **Policies** – which is our term to summarize situations that vary between DP schools in how the DP is implemented, in particular how the workload (which in principle is not much different between students or between schools) is spread and coordinated over the two DP years.

In our research design we have included multiple measures to tap these different situations. In sections 7.1-7.6 we will first discuss the theoretical rationales behind these measures and their empirical content. In these sections we also examine to what extent the time students spend on DP, their subjective workload, and their stress levels are related to these factors. Finally, in section 7.7, we examine whether and how the process of building up feelings of overload with the DP and the subsequent development of stress symptoms are affected by these P-factors.

Figure 7.1 illustrates which part of the causal model is addressed in the current chapter.
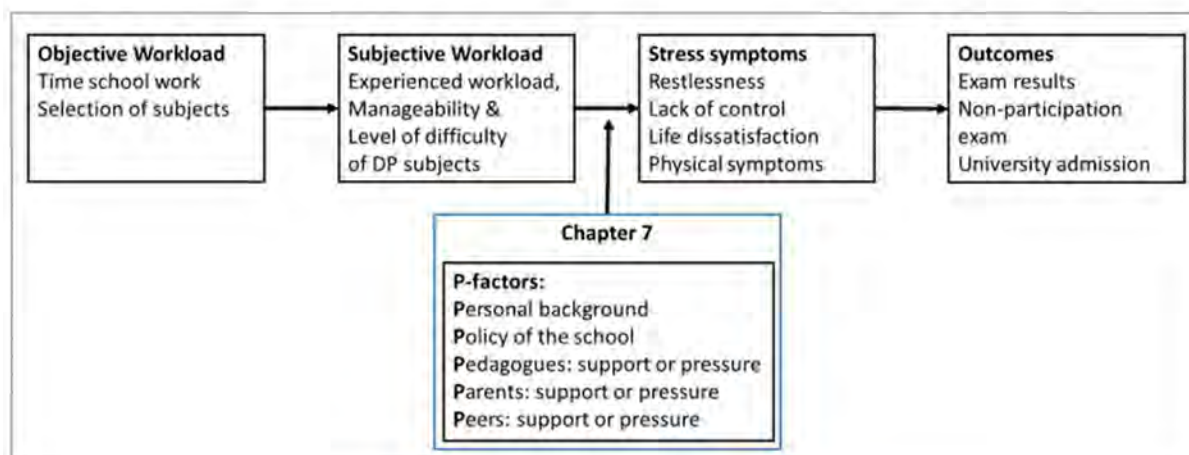
Figure 7.1 Causal model of students' objective and subjective workload and stress

With respect to the relations between the P-factors, the time students spend on the DP, their subjective workload and their stress levels, the findings can be summarized as follows.

- Concerning **Personal background**, we find that stress levels, and to a smaller extent also the subjective workload, are higher among girls than boys and lower among students with higher academic abilities. Students with higher academic abilities also tend to spend more time on their DP school work, as do girls.

- The DP **Programme,** more specifically, **the workload of students' selection of subjects** – defined in terms of the difficulty and time burden of the individual subjects – only affects the amount of time students spend on DP, which increases with the combined workload, without, however, affecting their subjective workload or stress levels.

- Stress levels and, to a smaller degree, levels of subjective workload are lower among students whose **Parents** are involved with their personal lives. The same holds, but less strongly, for the stress levels of students whose parents are concerned with their school work. By contrast, higher level of parental aspirations are related to higher stress levels. No associations were observed between parents' involvement and the time students spend on the DP.

- With respect to teachers – **Pedagogues** – in the DP, we find that the higher the quality of student-teacher relationships the lower the subjective workload and levels of stress. Subjective workload and stress are also lower among students who rate the quality of teaching more positively. More **peer support** is also associated with lower levels of subjective workload and stress. School pressure, by contrast, increases them. We find no variations in students' time investments in the DP according to these factors.

- Surprisingly, we hardly find any relation between school **Policies** and levels of subjective workload and stress. Only schools' entrance requirements and planning are associated with higher time investments in school work.

- We do not, however, find any evidence for the moderating role of these P-factors, meaning that the process of how workload becomes stress does not vary between students of different personal backgrounds, parents, peers, pedagogues and policies. The only exception is gender: the effect of workload on stress is stronger for girls than for boys.

7.1 The relations of P-factors with time students spend on the DP, their subjective workload and stress

- In the following sections, we examine how time investments in the DP, and levels of subjective workload and stress relate to students' personal background, to their relations with parents, teachers, peers at school, and to the school policy. Students and school characteristics are differentiated into P-factors. The first P-factor comprises **personal background**: gender, academic ability, language ability, parents' socio-economic status, and being an internal student or not. As these factors precede enrollment in DP, the effects can be interpreted as *causing* any differences in the time spent on DP, subjective workload and stress. The same holds for the **workload of student's' selection of subjects**, mainly established at the start of the DP in year 1. Variations in our variables of interest between students with a heavier and a lighter selection of subjects can be attributed to their individual DP curriculum.

- Such a causal interpretation does not apply to other student characteristics listed as P-factors, in particular **parents' involvement, teacher characteristics and peer relationships**. Although these may influence the time students spend on the DP, and their levels of subjective workload and stress, the relation here could just as well be reversed. The time students spend on the programme, the subjective workload and stress may affect students' relationship' with parents, teachers and peers as well. However, as these factors are measured in wave 1 and 2, and students' DP time investments, subjective workload and stress levels are also measured in wave 3, the time order allows at least some causal interpretation. Moreover, the causal order regarding the school averages of **teachers' characteristics and peer relationships**, is even less ambiguous. After all, it is unlikely that an individual students' time investment, subjective workload or stress affect the relationship with parents, teachers and peers of all students at the school.

- **School policy** characteristics are largely measured in wave 3, as school averages of individual students' answers or as school reports of DP coordinators. A similar argument applies to interpret a relation between school policy and individual students' time investments in the DP, their subjective workload and stress as causal. It is more convincing that school policy would affect individual students, than that such individual students' characteristics would directly affect school policy.

- **Appendix 7.1** presents the relations between students' individual and school characteristics and their time spent on the DP, and levels of subjective workload and stress, for each of the three waves (beginning of year 1, end of year 1, end of year 2. **Appendix 7.2** presents multilevel standardized regression coefficients of students' individual and school characteristics on their time spent on DP, and levels of subjective workload and stress, but, to make the causal order less ambiguous (see the argument above), only for wave 2 and 3. Moreover, the coefficients are corrected for all personal background variables (and personal background variables are corrected for each other), so that the relations between the P-factors on the one hand and DP time investment, levels of subjective workload and stress on the other, are not due to students' individual backgrounds. We do not take into account that the P-factors may be related as well, as there is no clear causal ordering of the P-factors.

7.2 Personal background

- Students' personal backgrounds were discussed in chapter 3. Here we take the following characteristics into account: gender, academic ability, language ability, parents' socio-economic status, and the students' home situation (see **Appendix 7.1** and **Appendix 7.2**).

- **Figure 7.2** shows the significant effects of personal and social background on time investment, subjective workload and stress, in the order of the effect size on stress. The positive effect of time on stress indicates that stress levels have increased strongly from the end of year 1 to the end of year 2. The effect next in size is that of gender: girls experience higher stress levels than boys. Girls also experience a higher subjective workload, but the effect is smaller than that on stress. Finally, girls spend more time on their DP school work than boys. Language ability negatively influences subjective workload, meaning that the higher the language ability, the lower the subjective workload, or vice versa, the lower students' language ability the higher the subjective workload. Academic ability has the largest negative effects: students who reported higher academic ability prior to the DP have lower levels of stress and subjective workload. The positive effect on time investment shows that they also spend more time on their DP school work.
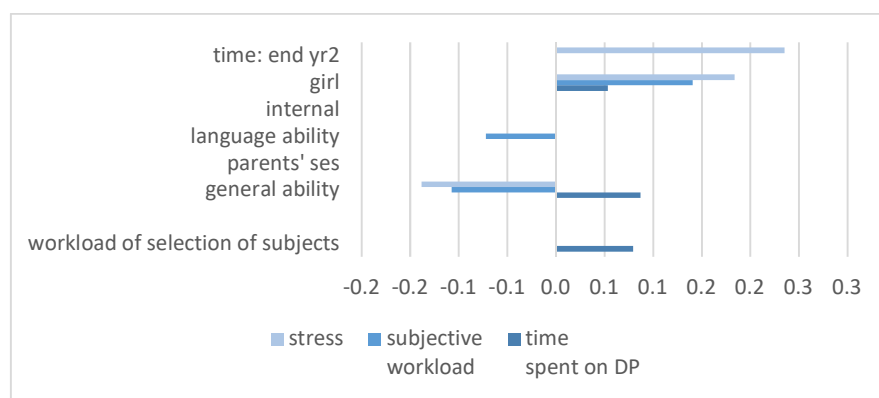
Figure 7.2 Time spent on DP, subjective workload, and stress: effects of personal background and workload of the selection of subjects

7.3 Objective workload: Students' selections of subjects

- **Figure 7.2** also depicts the effects of the workload of the selection of DP subjects, corrected for the effects of personal and social background. There are only variations with respect to the time investment in DP: students with heavier subjects spend more time on the DP. The workload of students' individual subject selection does not, however, impact their levels of subjective workload and stress.

7.4 Parents' involvement

- The home environment may be an important factor in the degree to which students can cope with the demands of the DP. The quality of the relationship with the parents in general may be an important factor in ensuring a safe and warm home environment from which students may be better able to cope with the pressures of their school work. In addition, parents' involvement with school and school work may help students to manage their homework and to prepare for tests and this way prevent them from feeling overloaded. On the other hand, if parents are too closely involved with school, they may also put pressure on students.

- **Figure 7.3** presents several indicators of parents' involvement with the life and school work of their children, the length of the bars reflecting the degree to which students agree with the statements. These statements can be differentiated into three dimensions of parents' involvement (see for more details **Appendix 7.3 and Appendix 7.4**). The indicators of the first dimension refer to parents' involvement in students' life in general, not necessarily their school life. The item students agree with most strongly is: "My parents want me to be happy". Students also strongly agree with the two items reflecting parents' aspirations for their children's school career. The items representing the third dimension are statements students agreed with least. These refer to a more active role of the parents in their child's school work.
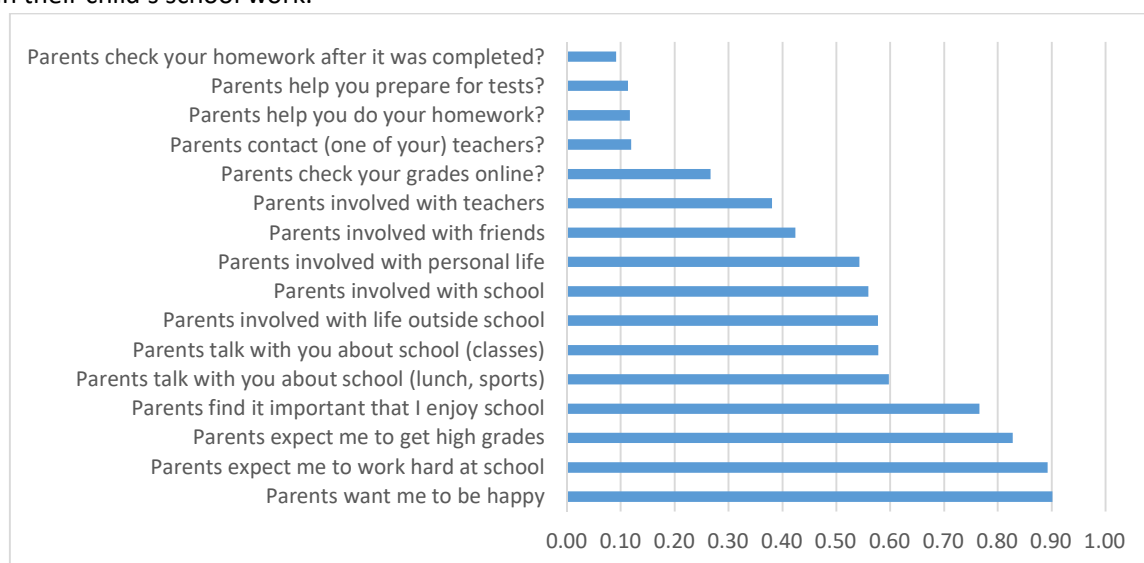
Figure 7.3 Indicators of parents' involvement

- For further analyses we constructed separate dimensions of the three-factor solution, which we name 'parents' involvement with life', 'parents' aspirations', and 'parents' involvement with school':
  - Parents' involvement with student's emotional life – the degree to which parents care about the happiness and social life of their children.
  - Parents' involvement with the student's academic work – the degree to which parents show interest in and support for their children's academic work.
  - Parents' aspirations – the degree to which parents want their children to excel in their school work.

While the concerns of the parents with the student's emotional life and academic work are strongly correlated, there is only a weak correlation with the parents' ambitions.

Figure 7.4 shows the effects of parents' involvement, again in the order of the effect size on stress and corrected for the effects of personal and social background. Higher parents' aspirations increase the stress levels, although the effect is weak. Parents' involvement with their child's school work, by contrast, lowers the stress levels, also weakly. Parents' involvement with the personal life of their child has the strongest effect: more involvement lowers the stress levels quite strongly, and to some extent also the subjective workload. Still, we need to be cautious, as it could also be that students who experience higher stress levels experience that their parents are not involved with their lives.
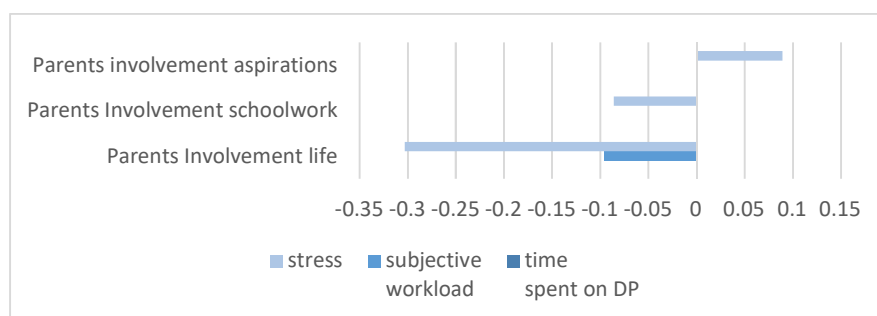
Figure 7.4: Time spent on DP, subjective workload, and stress: effects of parents' involvement

Scale construction

- **Appendix 7.3** summarizes several indicators of the parents' involvement with the life and school work of their children, all recoded into a 0-1 range. Factor analyses, in **Appendix 7.4,** of these indicators show that these indicators could be summarized in one latent factor (explaining 28.5% of the variance), but a three-factor solution also fits the data and is well interpretable (factor 2 and 3 explaining 12.7% and 10.9% of the variance).
- For analyses we constructed three separate dimensions of the three-factor solution that we name 'parents' involvement with life', 'parents' aspirations', and 'parents' involvement with school'. Additional analyses show that most variations (> .90%) are between students within schools. Remarkable is the 10% country variance in parents' aspirations, implying that these vary across countries.

7.5 Pedagogues and peers

- The school environment of the DP as represented by the school staff, teachers and schoolmates may help or hinder students in coping with the pressures of the demands of the programme. We asked students to report on several indicators on teacher conduct, peer relations and the school climate in general. In **Appendix 7.5** these are described in detail. Three dimensions are distinguished: **quality of student-teacher relationships**, **quality of teaching**, and **school pressure** (including also two items that do not refer to teachers directly).

- The items on the quality of student-teacher relationships are depicted in **Figure 7.5**. The bars indicate the degree of satisfaction, agreement or occurrence of the respective statement. As can be noticed, most items have an average score above the scale midpoint of .5, which refers to the positive side of the scale. Indicators of teaching quality are presented in **Figure 7.6**, again showing positive assessments by the students. **Figure 7.7** demonstrates that at the same time relatively high scores are given on items on school pressure (including peer pressure). Apparently, school pressure can go hand in hand with good student-teacher relationships and teaching quality.
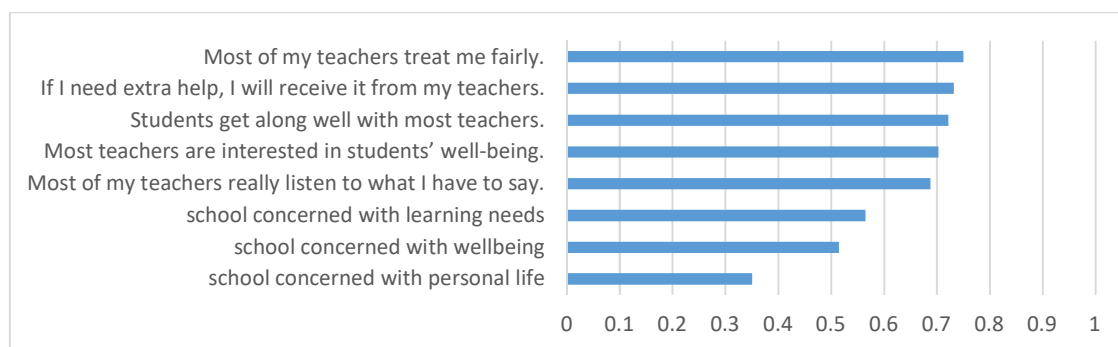
Figure 7.5 Student-teacher relationships
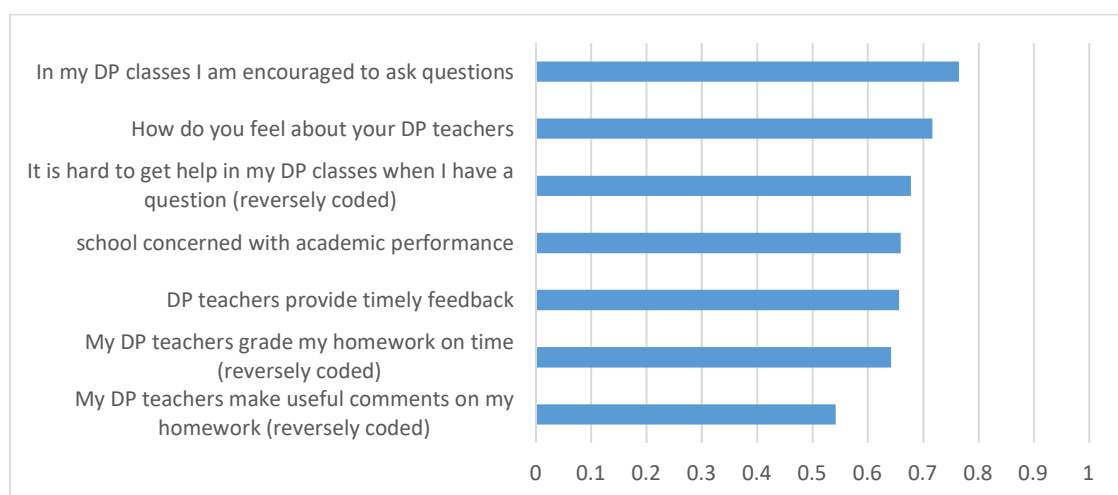
Figure 7.6: Teaching quality
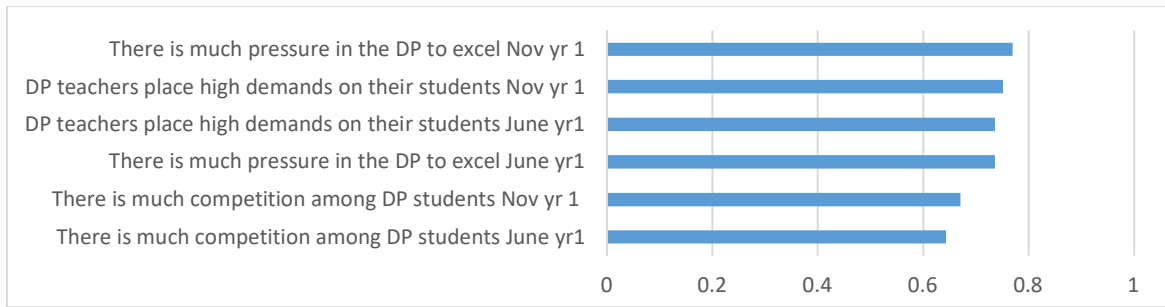
Figure 7.7 School pressure

- **Figure 7.8** presents the effects of teachers involvement and peer support, according to students, again corrected for personal background. The top half of the figure shows the effects of students' individual reports, the bottom half those of the school averages of these measures. Overall, the latter are smaller but also less ambiguous concerning a causal interpretation. School pressure, a mixture of teacher' and peer pressure, is associated with higher levels of stress and subjective workload. At schools where students feel under pressure to excel, stress levels and levels of subjective workload are higher than at schools where students feel more at ease. The remaining aspects of teacher involvement and peer support are related to lower levels of stress and subjective workload. The more students – both individually and on average – think highly of teaching quality, experience positive teacher-student relations, and feel integrated in their school environment, the lower the levels of subjective workload and stress.
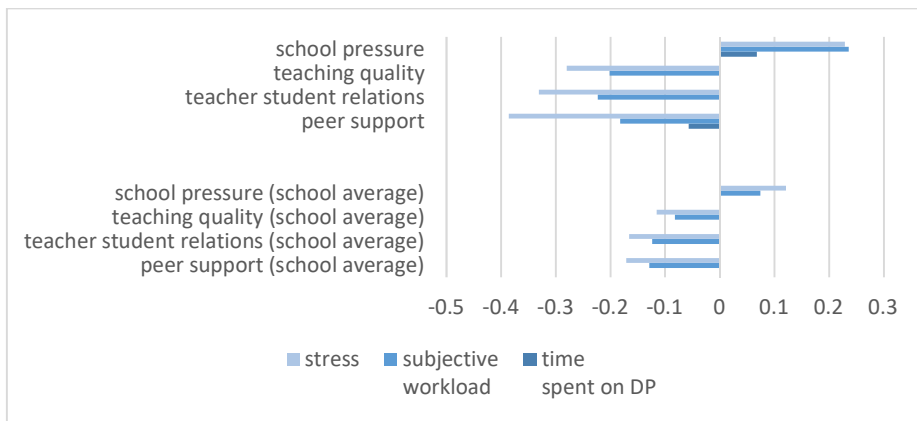
Figure 7.8: Time spent on DP, subjective workload, and stress: effects of teachers' involvement and peer support

- **Figure 7.9** describes the indicators on peer relations on a 0-1 scale. All indicators are above the scale midpoint of .5, so, on average closer to the positive end.
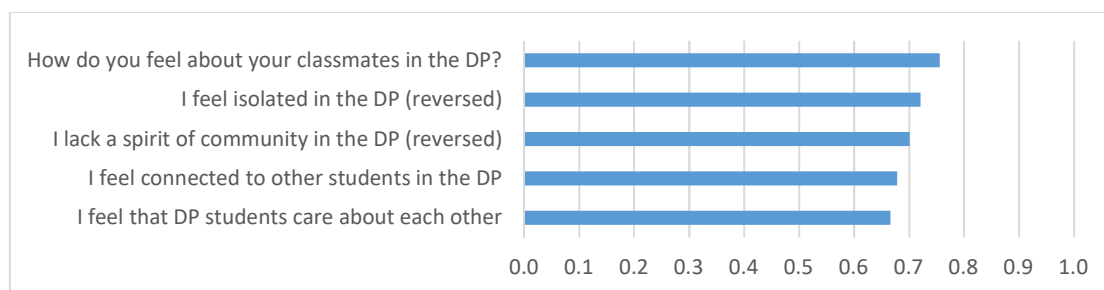
Figure 7.9: Student-peer relationships

Scale construction

Appendix 7.5 presents a description of several indicators on teacher conduct, peer relations and school climate.

- Items that clearly refer to **school climate and teachers** were factor analysed to examine to what extent these could be summarized in one or several dimensions. Factor analysis shows that the answers reflect some (artificial) dimensionality according to the order and position of the items in the questionnaire.
- We differentiated the items into three dimensions: quality of teacher-student relationships, quality of teaching, and school pressure (including also the two items not directly referring to teachers). Factor analyses (with the demand for one single factor per dimension) and reliability analyses of these are presented in **Appendix 7.6**. Three scales were constructed as the average of the separate 0-1 indicators and centred to their means. Quality of teacher- student relationships and quality of teaching correlate positively and rather strongly (.432). The correlations with school pressure are negative, -.025 and -.143 respectively. For further analyses we also included the school averages of the above scales, to represent the school effect. Multilevel analyses show, however, that school variations in the quality of student-teacher relationships is rather low (6.4% school variance, no country variance), that country and school variations in teaching quality are a bit higher, (6.7% country variance and 10.8% school variance), also with respect to school pressure (7.5% country variance, 5.5% school variance).
- Items that clearly refer to **peer relations (Appendix 7.6)** were factor analysed to examine to what extent these could be summarized in one or a few dimensions. Factor analysis shows that the items constitute one single factor, explaining 56.6% of the item variance. All items have factor loadings > .6. Reliability analysis shows that Cronbach's alpha is .80.
- Therefore, one single scale was constructed as the average of the separate 0-1 indicators and centred to the mean. For further analyses we also included the school averages of the above scales, to represent the context effect. Multilevel analyses show that there are some country and school variations in the quality of student-peer relationships (6.3% country variance and 8.1% school variance).

7.6 Policies

- The DP requirements are similar across schools, yet there may be variations in the students' time investments in DP, their subjective workload, and their stress levels, according to schools' practices and policies to balance students' DP workload. We differentiate:
  - o   Information and guidance to help students make appropriate study choices
  - o   Coordination and planning to balance the workload for DP students
  - o   Support services to support students when they are struggling

- As we are interested in the relation between school policies and students' time investments, subjective workload and stress, in this chapter we focus on the practices of the (max 98) schools of the student sample. In chapter 9 we discuss these same measures for the broader sample of schools, including the opinions of both DP coordinators and students.

Information and guidance to help students make appropriate study choices

- Schools consider certain factors before admitting students to the DP, **entrance requirements**. These factors are listed in **Figure 7.10**. A majority of the schools consults students' records of academic performance (66.3% always, 27.5% sometimes) or have an interview with the student (63.8% always, 21.3% sometimes). A language test or recommendations of feeder schools are considered less frequently, merely 34.6% and 29.4% of the schools always takes them into consideration. An ability test is least used, only 15.4% of the schools always uses ability tests, 17.9% sometimes use them. These entrance requirements may lead to variations in student populations that may affect their time investments, the way they manage the DP workload and their stress levels. We therefore created a school measure of the strictness of the entrance requirements, see also **Appendix 7.7**.
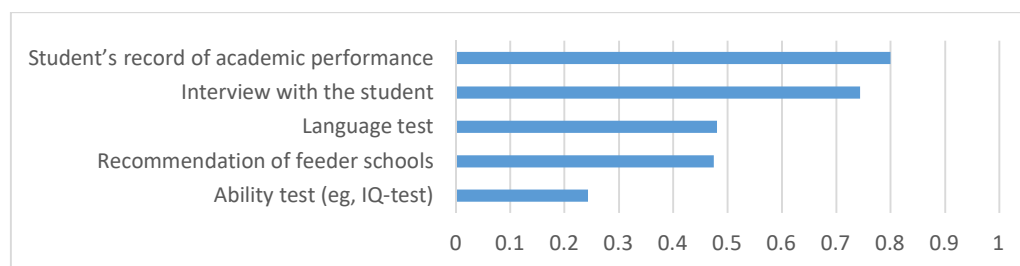
Figure 7.10: Schools' entrance requirements: factors considered when students are admitted to DP

Coordination and planning to balance the workload for DP students

- **Figure 7.11** describes the **number of internal assessments** (IAs) and **core exam components** that students report to have submitted over the course of DP. Schools may have adopted a policy to spread internal assessments and the core components out over the two DP years. A more balanced spread could reduce the impact of the DP workload. Information on the concentration of internal assessments and core component comes from students and DP coordinators. DP coordinators also reported on the

school's policy. The lines in the figures represent a monthly average, adding up to 1.9 IAs in year 1, and to 5.7 IAs in year 2 for the schools of the student sample. The number of IAs in year 2 is around 2.5 to 3 times as high as in year 1, with a peak in the period December - February of year 2. The total number of IAs is on average 7.7, mostly (95% of the schools) varying between 4 and 11 IAs over the course of DP. This is a bit lower than the estimate of 10.6 by the DP coordinators of the sample schools. The monthly average of core components in year 1 is about 1.0, in year 2 3.0, also with a peak in the period December-February in year 2 (see also **Appendix 7.8**).
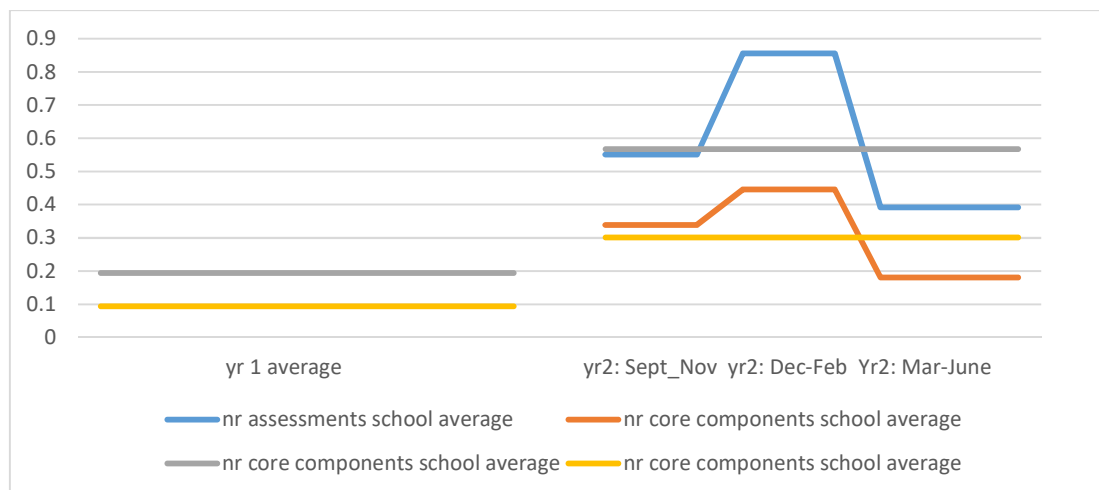
Figure 7.11: Number of assessments and core exam components handed in, estimates by DP students (monthly average)

- DP coordinators were furthermore asked about the ways in which they organize and coordinate assessment dates and homework. **Figure 7.12** (and **Appendix 7.9**) describes different ways of organizing and planning. The coordination of assessments is rather high: over the midpoint of the scale (.5). The majority of the schools always coordinates deadlines for IAs and core components, 64.9% and 71.2% respectively. Schools also organize collaborative meetings on planning issues, on average they are around the midpoint of the scale (.5) meaning that they vary from three or four times a year to once a month. Only a minority of the schools (16.0% to 28.0%) imposes limits to students' homework. For further analyses, we combine these indicators into a measure of the school's policy of organizing and coordinating assessments, aimed at a more equal distribution of workload over time.
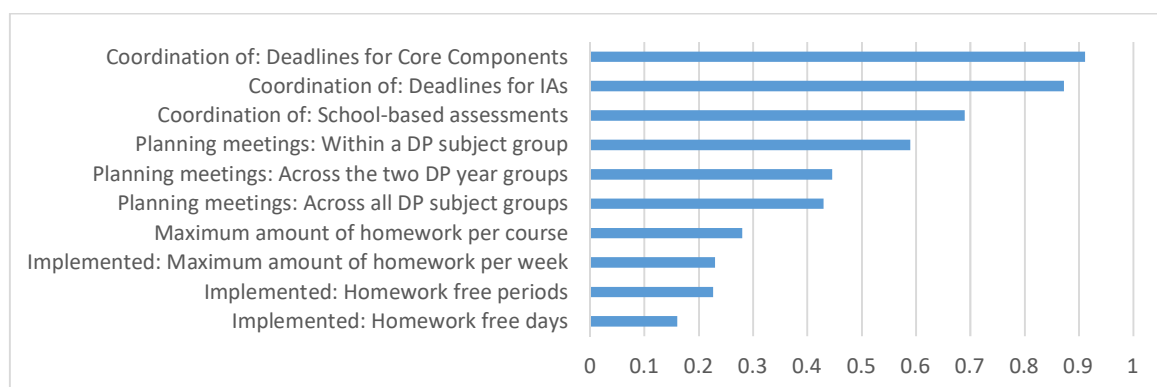
Figure 7.12: School policy on limitations of homework, and spreading of internal assessments and core components, according to DP coordinators (schools student sample)

- **Figures 7.13** and **7.14** (see also **Appendices 7.10** and **7.11**) describe the school's practices to spread students' workload. **Figure 7.13** is about the organization of students' study time in general, according to the students. A 'clear schedule of (non-examination) assessment deadlines' and 'designated study period blocks in the timetable' are most frequently provided, at 43% and 34% of the schools in the sample respectively, still less than half. Measures with regard to the regulation of homework are taken much less often, varying between 6% and 15% of the schools. **Figure 7.14** is about measures with regard to the spreading of IAs and non-examination assessments on which we have reports from students and coordinators. Around 20% of the schools in the sample takes measures with regard to the spreading of IAs and other assessments aimed at finalizing those assessments in year 1 or in November year 2. Finalizing some IAs in year 1 is the most frequent of these measures at the schools in the student sample, practiced by 28.5% of the schools according to students and 39.2% according to the DP coordinators.
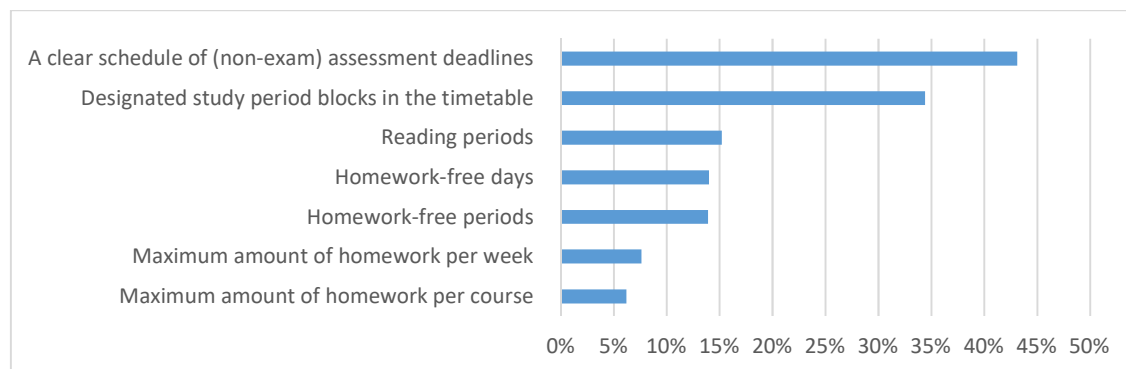
Figure 7.13: Schools' practices to make the workload more manageable
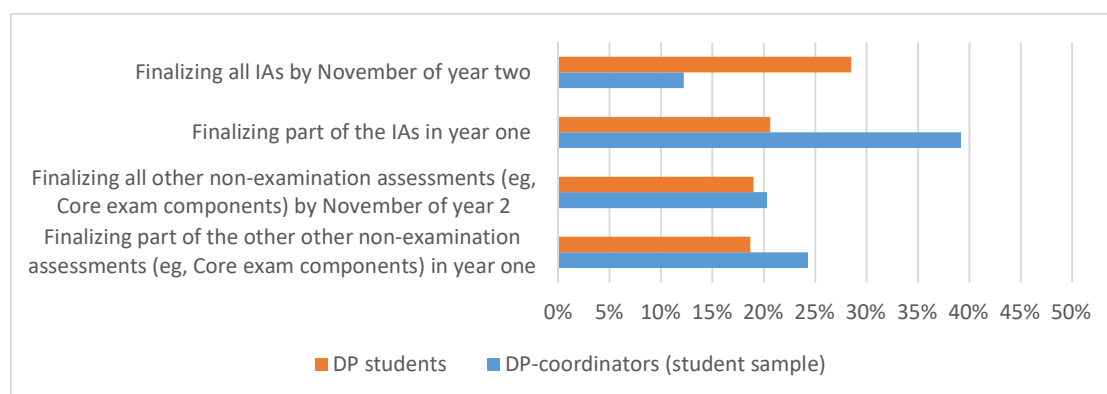
Figure 7.14: Schools' practices to make the workload more manageable

- **Figure 7.15** (and **Appendix 7.12**) presents students' opinions on their school's planning of deadlines of assessments and on homework assignments. The bars represent the degree of 'non-effectiveness' according the students. On average students are in the centre of the scale, indicating a moderate

effectiveness ('sometimes' effective). As assessments of the school policy we would expect them to be inversely related with the schools' policy of organizing and coordinating assessments (school policy aimed at more coordination should receive less complaints of non-effectiveness). This, however, is not the case: both measures correlate positively (.17). Further analyses show that there is large variation between students within schools on the opinion on the effectiveness of the school's planning, which indicates that they reflect students' personal experiences, rather than being an (intersubjective) indicator of school policy.
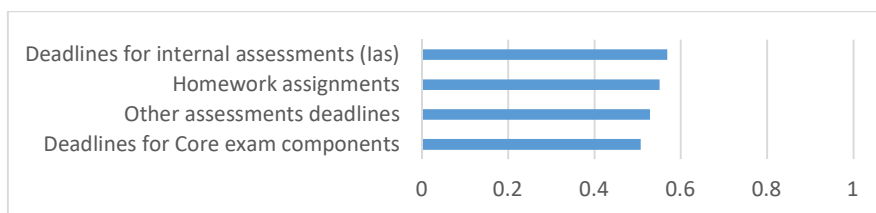
Figure 7.15: Non-effectiveness school planning according to students

Support services to support students when they are struggling

- Students were asked about the presence of **support services** at their school to help them do their school work (ie, coaching, tutoring, additional lessons) and to promote their health and wellbeing (ie, the presence of mental and physical health professionals). **Figure 7.16** shows that a large majority of schools in the sample (77.4%) offer university counselling and this way help to prepare students for their follow-up schooling career. Regarding the DP school work, the majority of the schools offers teacher-led coaching, both in groups (66.1%) and individually (56.6%). 65.5% of the schools offer coaching aimed at special needs (eg, dyslexia, Attention Deficit Hyperactivity Disorder). Student-led coaching, in groups or individually, is also provided, at respectively 42.3% and 38.1% of the schools. Externally-led coaching is much less common, practiced only at 17.6% (in groups) and 20.3% (individually) of the schools.
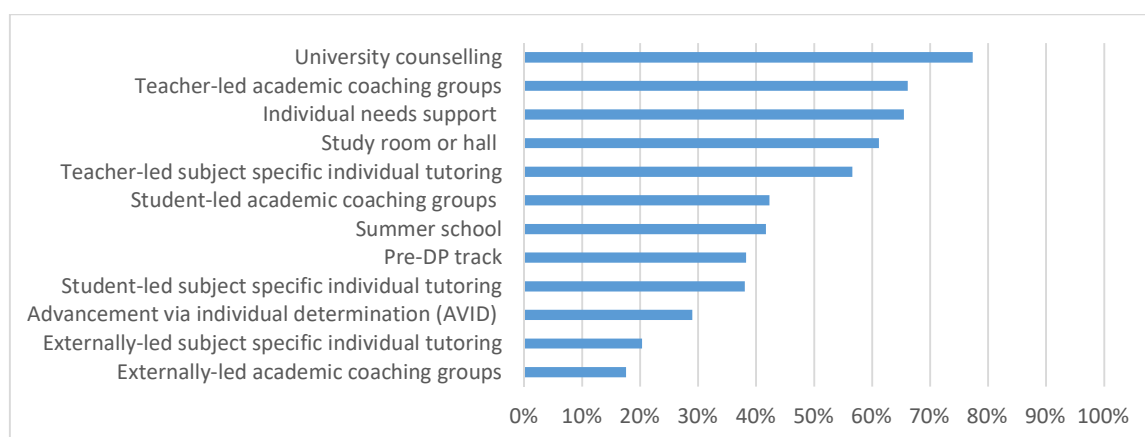
Figure 7.16: Schools' support services with respect to students' needs beyond the classroom according to DP students

- **Figure 7.17** displays the presence of different types of mental and physical **health professionals** at schools. At most schools a school counsellor is available (89.1%). Also, at many schools a nurse (74.0%), a social worker (46.9%) or a psychologist (44.7%) is present. To create a measure of the number of support services, students' answers were averaged per school, and then the overall average of support services, aimed at their school work and at (mental) health was taken. See also **Appendix 7.13**.
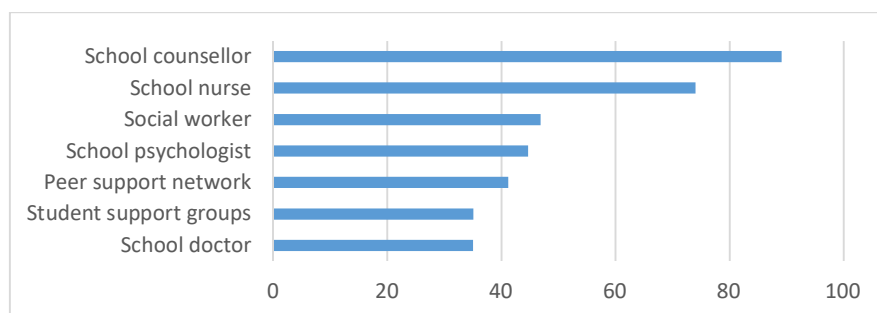
Figure 7.17: Schools' support services with respect to students' health and wellbeing according to DP students

- Finally, **Figure 7.18** presents the reports of DP coordinators on the presence of three programmes schools may offer, aimed at students' learning skills (ie, time management, organization, communication) or their mental health, for instance the programme Social and Emotional Learning (SEL) or relaxation programmes, such as mindfulness or yoga (see also **Appendix 7.14**). Approximately half of the schools provide programmes aimed at students' mental health.
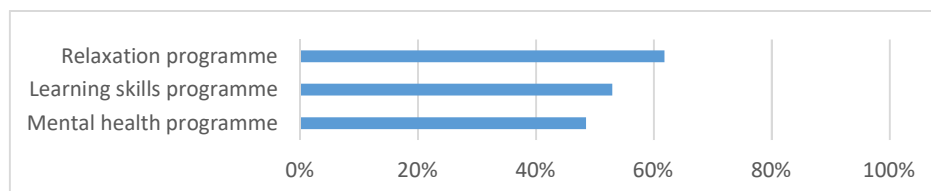
Figure 7.18: Provision of schools' support programmes

Scale construction
- All factors considered as **entrance requirements** correlate positively, between .321 and .521 in the broader sample of DP coordinators, resulting in a Cronbach's alpha of .765. To construct a scale of these items, they were recoded into a 0-1 range, and then the mean was taken as a measure of the severity of entrance requirements (M = .57, sd = .29), see also **Appendix 7.7**.
- **Appendix 7.8** describes the number of **internal assessments (IAs) and core exam components** that students have to submit. Students have estimated the number of assessments as an overall average over year 1 (reported in wave 2), and as averages per three- or four-month periods over year 2 (reported in wave 3).
  - To create a school measure of the concentration of IAs and core components, we took the school averages of **students' estimates** of the **number of IAs[1]** in year 1, in year 2 (Sept-Nov),

---

[1] For core components the sum is not relevant, the maximum is three. We asked for the average number of core components in a three or four month period, the sum refers to the number of months spent on core components.

year 2 (Dec-Feb), and year 2 (March-June). Having four measures of IAs and core components per school, we calculated the **standard deviation as a measure of the school's spreading of IAs and core components across the DP**. A higher within school standard deviation implies that IAs and core components are concentrated in certain periods. The average number of IAs is estimated at 7.7 (std dev 1.8). The within school standard deviation of IAs is on average 1.1 (std dev .5), of core components .6 (std dev .3).

o We followed a similar procedure for the **DP coordinators' estimates** of the number of IAs and core components over de course of DP (DP coordinators of schools in student sample). Here there were six retrospective measures available. The average number of IAs is estimated at 10.6 (std dev 5.0). The within school standard deviation of the IAs is on average 1.6 (std dev .7), of the core components .8 (std dev.4). Students and coordinators do not agree though. On the school level their estimates on the number of IAs correlate .25, whereas the within school standard deviation is for IAs negatively related, -.09, for core components .153. For analyses, we give priority to the students' school average reports, because they are more reliable (more observations per school) and because they are available for all schools (not all DP coordinators of the schools in the sample responded).

o The indicators of different ways of **organizing and planning**, presented in **Appendix 7.12** were coded in a 0-1 range. Factor analysis shows that four factors could be extracted, but we decided to combine them into a single concept. Reliability analysis showed a Cronbach's alpha of .66, with no improvement if one of the items was to be removed. For analyses the concept was constructed as the average of all 0-1 indicators, which again was recoded into a 0-1 range, and centred to its mean.

o The **provisions** presented in **Appendix 7.13** and **7.14** were used as an additional measure of the school's policy to make the workload for students more manageable. For analyses the concept was constructed as the average of all 0-1 indicators, which again was recoded into a 0-1 range, and centred to its mean. Again, for analyses, we give priority to the students' school average reports, because they are more reliable (more observations per school) and because they are available for all schools (not all DP coordinators of the schools in the sample responded).

- The effects of school policy, corrected for students' personal backgrounds, are described in **Figure 7.19**. Only the students' opinions on the effectiveness of the school's planning and organization is related to levels of stress and subjective workload: if students think that the school is not effective in this regard, levels of stress and subjective workload are higher. However, we need to be cautious in interpreting this relation causally. Students report on the (non-)effectiveness at partly the same time (wave 3) as on their subjective workload and stress levels (wave 2 and 3), it could well be that the relationship is reversed, so that higher stress levels could just as well lead to complaints on the schools' organization and planning. No other school policy measures relate significantly to students' levels of subjective workload and stress. The time students spend on the DP is only related to the schools' entrance requirements and planning. At schools with more entrance requirements, students spend more time on their schoolwork than at schools with fewer requirements. And, somewhat surprisingly, the more coordination and collaborative planning meetings, and homework limitations, the more time students spend on DP.
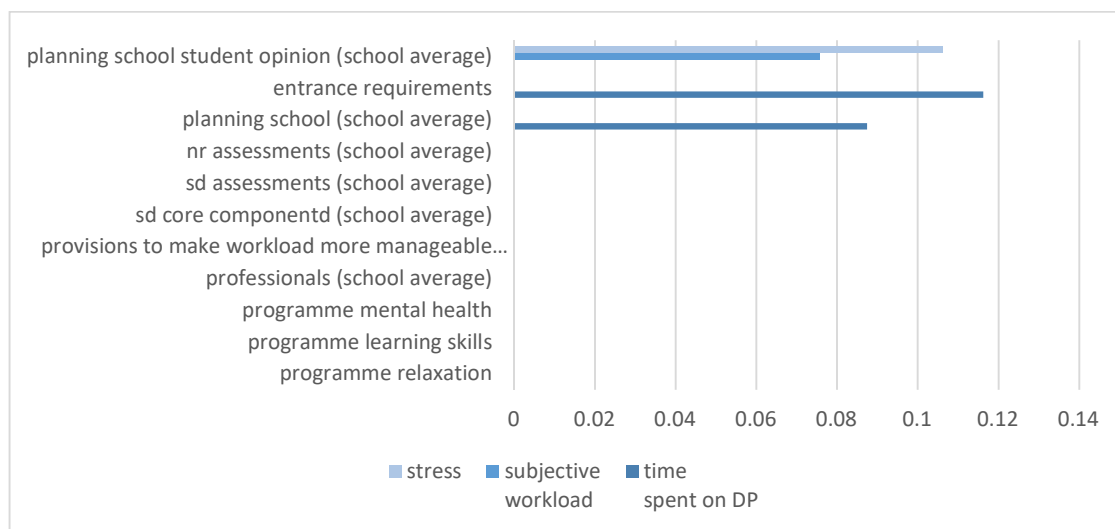
Figure 7.19: Time spent on DP, subjective workload, and stress: effects of school policy

In addition to the information collected on the three waves (beginning year 1, end year 1, end year 2), there is dynamic information available on the **number of assessments in the course of year 2**. The information was collected in wave 3 (June year 2) by asking students to look back on year 2, and to report on their time investment, subjective workload and stress levels[2] in the periods of September-November, December-February, and March-June respectively, as well as the number of internal assessments they had in these periods. Having this information, we are able to examine the impact of the number of assessments (school averages). The results of the multilevel regression analyses, in which coefficients are corrected for time and personal and social background, are displayed in **Figure 7.20**. Time investment, subjective workload and stress levels all increase when the school's number of assessments increases, independent of the time development of year 2, and corrected for personal background, see also **Appendix 7.15**
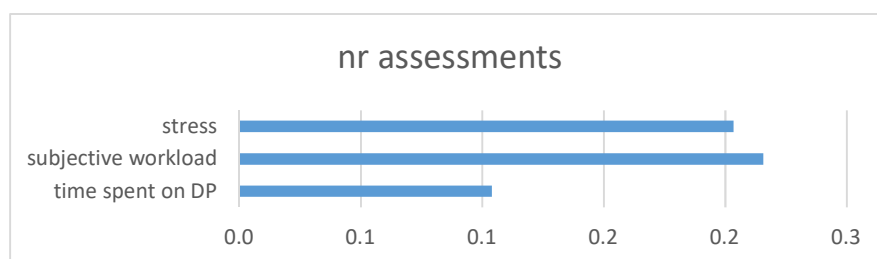
Figure 7.20: Number of assessments over the course of year 2 and the effects on time spent on the DP, subjective workload, and stress.

7.7 When does workload become stress? The role of moderators

- In chapter 6 we have developed the reciprocal causation workload-stress model that – when applied to three-wave panel data –, can decompose the correlations between objective workload (time spent

---

[2] Subjective workload and stress are – contrary to the measurements used elsewhere – measured by a single item.

on DP work), subjective workload (feelings of being unable to manage the programme demands) and stress symptoms (manifestations of psychological discomfort) into a stable part (due to the stable conditions the DP students are in) and a dynamic part, ie, changes that arise during students' DP career. For this dynamic part the model can identify a causal effect of subjective workload on stress, controlled for the reversed (feedback) effect of stress on subjective workload. In Chapter 6 this process was assumed to be the same for all DP students – the model portrays how DP students around the world fare **on average**. Of course, it need not be the case that the process homogeneously applies to all students: it seems likely that some students are less able to cope with the exigencies of the DP than others. Examining attributes and situations of these students may be important clues how the building up of stress during the DP years can be prevented or redressed.

- This type of statistical analysis is technically known as **moderation analysis**. The main question here is not whether and how the students' circumstances affect their outcomes (objective workload, subjective workload, stress symptoms), but rather whether these circumstances affect how these three are related to one another. What makes that some students experience a given time investment as too hard to manage, while others report no problems? **And in which conditions does high subjective workload lead to psychological discomfort and when does it not**?

- In order to do so, we estimate and compare the reciprocal causation model between three (or two) different groups based on the previously introduced P-factors: students who score high on the respective P-factor, students who score low, and students who take a middling position. For P-factors that are by nature dichotomous (such as gender), this reduces to a two-group comparison. Each time the question is whether the workload–stress development works the same for the comparison groups. We will introduce the methodology of comparing the stress development process between groups in a worked example that contrasts two groups: male and female students. Having introduced the methodology and after pointing out the interpretation of the main parameters of the model, the other P-factors will be addressed more briefly.

Gender

- With respect to gender the literature has often found, and our data confirm this, that girls worry quite a bit more about academic pressures than boys. Female DP students systematically express higher levels of subjective workload and also report more severe levels of stress symptoms. As gender is a rather stable student characteristic, it can by itself not possibly explain how workload dynamics lead to changes in stress levels. However, the strength of the effects of workload on stress, can still differ between male and female students. Are girls indeed more sensitive to workload and do they convert a similar level of workload more readily into psychological discomfort than boys?

- While we do not anticipate that gender has indeed a major influence on the workload-stress process, detailing the moderation analysis by gender in an elaborate worked example is useful. Because the moderator is unequivocally divided into two similar-sized groups, the moderation analysis is relatively simple and particularly adequate to serve as an illustration.

- **Table 7.1** shows the basic ingredients of the moderation analysis: the correlations for boys and girls separately, together with their group-specific means and standard deviations.
  First of all, it is important to note that both levels of subjective workload as well as stress, differ between female and male DP students. In fact, the correlations in **Appendix 7.1** indicate that student gender is a stable predictor of both outcomes, with girls consistently reporting higher workload than boys (r = 0.14, 0.13, 0.12, at the three waves) as well as more stress symptoms (r = 0.14, 0.15 and 0.20). While these associations are not strong (but statistically significant), this makes student gender a potential confounder for the workload – stress relationship. However, this relationship itself is much stronger (0.47, 0.48 and 0.53, for boys and girls combined) than could be accounted for by gender alone. It is of interest to see that these correlations are stronger for male than for female students at the first two waves, but not at the third: 0.50 / 0.44, 0.55 / 0.52, 0.49 / 0.57 (the first correlation is for boys, the second for girls).

- But what does this changing pattern of correlations imply for the causal effect of workload on stress? Does it mean that female students are more likely to develop stress symptoms because of the DP-workload? **Table 7.2** displays three versions of the stress development models and how the effect sizes differ between male and female students. We compare a model B in which all four effects are allowed to be different between boys and girls (L2 = 101.8, df=24) to a model A in which all parameters are constrained to be the same (L2 = 123.4, df=20). The difference between the two fit statistics (L2 = 21.6, df=4, p < .001) implies that *some* significant differences exist between male and female students in the workload–stress dynamics. The next model C in **Table 7.2** locates two differences in the stress development process. First, this is in the stability of stress within persons: male students maintain their initial stress levels more strongly than female students. Second, there is a significant difference in how strongly workload leads to stress symptoms: **girls are significantly more vulnerable to workload pressures than boys.**
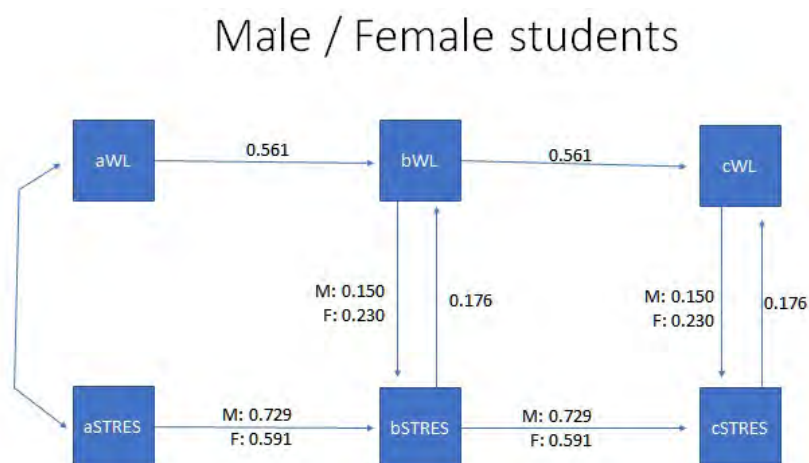
Figure 7.21: Differences in the Workload Stress process between male and female students

Methodological remarks

- The reciprocal causation workload-stress model can only be estimated using structural equation modeling {SEM}. The model does not have a single outcome variable, the point of using it is that it can isolate the causal effect of workload on stress from confounding and feedback mechanisms.

- The SEM methodology can report an overall statistical test on whether a model is significantly different between subgroups of students, but these differences can arise in multiple parameters in the model. As the estimation of the parameters is interdependent, it will not always be clear where a significant difference in the overall test is exactly located.

- Our strategy is intrinsically bivariate with respect to the moderating P-factors: we examine the heterogeneity between the three (or two) levels of the P-factors one-by-one, and hence do not take into account how P-factors may combine to produce heterogeneity. This bivariate strategy can be better justified if the P-factors are more or less uncorrelated and cannot confound each other's moderation effects.

- Our strategy of moderation analysis is discrete: we categorize the moderating variable in two or three groups, between which the process of workload-stress development is compared. **Figure 7.22** redisplays the workload-stress development model as derived in the previous chapter. The effects in this model are labeled **aa**, **bb**, **cc** and **dd** and can be interpreted as follows:
  - Coefficients **aa**, often referred to autoregressive pathways, model the degree of stability of the workload experience between waves.
  - Coefficients **bb**, also autoregressive pathways, model the degree of stability of the stress symptoms between waves.
  - Coefficients **cc** are the causal effect of subjective workload on stress symptoms, our core parameter of interest.
  - Coefficients **dd** are the reversed, feedback effects of stress on subjective workload.
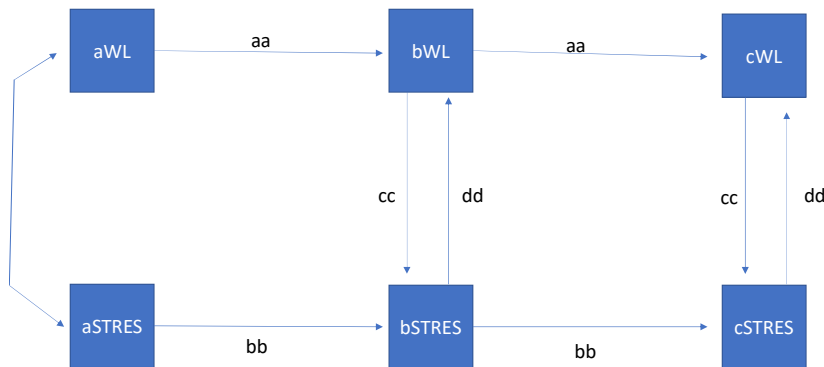
Figure 7.22 The work-stress development model

- Because effects **aa** and **bb** are controlled in the model, any further association between stress and workload can now be interpreted as how *changes* in subjective workload and *changes* in stress are related to one another. This implies that whatever relationship is found between subjective workload and stress cannot be confounded by prior causes, ie, situations that give rise to both subjective workload and stress. These 'backdoor' effects are controlled because both the prior

levels of subjective workload and stress are controlled. Such prior conditions (gender being an important example) will not confound the estimated reciprocal effects between workload and stress. However, it is possible that these reciprocal effects are different between categories such as male and female students.

- Technically, the reciprocal effects model is also an **instrumental variable (IV) model**, a form of non-experimental causality research that is most often used in econometrics[3]. An instrumental variable model is defined by the assumption that a prior condition C can only affect a consequence Y via a mediation variable X, and no other pathways between C and Y exist. The reciprocal effects model makes this assumption twice: prior workload can only affect current stress via current subjective workload; prior stress can only affect the current experience of workload via current stress. This double IV structure makes it possible to identify the two reciprocal effects.

- In addition, the reciprocal effects model in three-wave panel data doubles up the IV assumption, because it is applied both to the wave1 → wave2 transition and the wave2 → wave3 transition. This is not a necessary, but a useful extension, in particular, when all coefficients are constrained to be equal between waves. This is called the **stationarity assumption**. Under stationarity, we assume that we are observing a constant process of individual adaptation, and we take three snapshots of this process. The stationarity assumption is not absolutely necessary to identify the model, but it is helpful, because it leads to much more stable estimates. In fact, the stationarity assumption can be statistically tested, by relaxing the equality constraints between waves. Notice, however, that there are four stationarity assumptions, which makes for 12 possible ways to constrain or relax the model.

Academic ability

- A somewhat more plausible moderator of the workload-stress development process than gender is students' academic ability. As discussed in the Chapter 2, the DP has no formal entry requirement regarding students' academic ability. Nevertheless, we know that a fair amount of selection is at work with respect to academic ability. This exists because the schools that host the DP funnel high-performing students towards this academically challenging track. Similarly, high-performing students and/or their parents may self-select into the DP. Nevertheless, we should not expect all DP students to be equally strong academically, and differences in academic ability are likely to affect the way they experience the programme. This may be all the more be true because the self-selection of students into the DP may also be dependent upon parental resources, which are not necessarily related to their abilities. The most obvious hypothesis is that for academically stronger students there is a weaker relation between workload and stress. They are likely to need less time to complete the work and can use their academic ability to cope with the DP requirements.

- Our earlier analyses, however, have already shown that self-assessed academic ability is not inversely related to time investment in the DP. The relationship is complex: academically stronger students are likely to select a more demanding set of DP subjects, within the restrictions that the DP

---

[3] In econometrics IV models are usually estimated with 2SLS (two-stage least squares) instead of the SEM techniques we use. The SEM technique allows to use equality constraints to make the model estimates more robust. SEM also makes it possible to use incomplete data, as there often arise in panel research.

imposes. They are for instance more likely to choose HL in academically more demanding subject areas, in particular Mathematics and Sciences. As a result, these students spend more rather than less time on their DP work than other students. The question now becomes whether that translates into negative evaluations of the workload, and eventually in stress.

- By contrast, we have also found that students with higher language abilities choose a less demanding set of subjects. The explanation is rather simple: language ability to a large part reflects whether the student was raised in the language of instruction at IB (English, Spanish and French in our sample). Not fully mastering the language of instruction is most common in DP schools in Europe and Asia, but it also arises elsewhere. Students with such language barriers have a preference for Mathematics and Sciences (likely in order to avoid the language barriers) and end up having a more demanding individual curriculum.

- IB does not test students with respect to prior academic ability, and in the likely event that students were tested before entering the DP, no records are available. We therefore asked the students to report subjectively on their prior academic achievement, primarily by comparing themselves to their classmates before entering the DP. We separated the statements into two dimensions of academic ability, one with respect to math and sciences (widely acknowledged as the most difficult and time-consuming subjects) and language abilities which includes the self-report on how well the students are able to communicate in English or the other languages of instruction (Spanish or French).

- For the moderation analysis we separate both ability measures into three groups: low – middle – high. **Table 7.2** (second line) specifies how the core parameters of the workload – stress models are different between these three groups. The models are the same as before (for gender), except that we now have three groups to compare. Model A gives again the pooled estimate of all the parameters – these are more or less the same as the pooled estimate for male and female students. Differentiation by ability group improves the model fit significantly (L2 = 21.0, df=8); so there are significant differences in how the process unfolds. Upon further testing, the **cc** (direct effect) parameter can still be constrained to be equal between the three groups. For the other three parameters, it is the highest ability group (group=3) that stands out. For this group, the feedback effect (**dd**) is much weaker than for the other two groups (although still statistically significant), and they report more stable stress and workload levels (see **Figure 7.23**).

- The take-home conclusion is that academic ability moderates the workload-stress process, but not in the hypothesized direction. The parameter of most interest (Workload → Stress) can be assumed to be equally strong for all three levels of academic ability. What rather seems to be the case is that the academically strongest students have a weaker feedback effect: if something exogenous affects their stress level, they are less likely to translate that in worries about the DP workload.

Three levels of academic ability
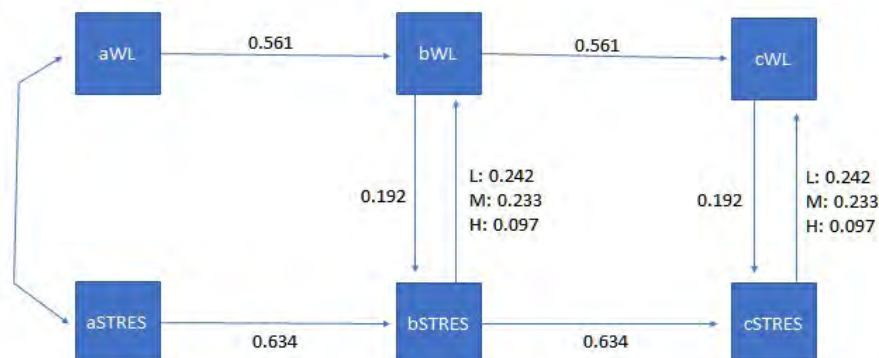
Figure 7.23: Differences in the Workload Stress process between three different levels of academic ability: Low, Middle, High

No other significant moderation

- We have applied this testing strategy on the full batch of personal and school characteristics, including the school variables that measure policies targeted to counter stress development – all of which were found to be non-effective in explaining differences between students. Most of these variables are not dichotomous (like gender), but continuous. In order to use the group comparison strategy, we have organized all continuous variables into three levels of equal size ('tertiles'), which is a more powerful way to detect group differences than resorting to simple dichotomies. Despite this strategy, we have found **no significant moderation** in the workload-stress development model.

- **Table 7.2** lists the relevant L2 tests. In principle we always compare three models: (A) the fully constrained model, with all effects constrained to be equal between groups, (B) the fully heterogenous model in which all effects are allowed to be different between groups, and model (C) in which only selected effects are allowed to be different. However, if the fully constrained model A and heterogenous model are not significantly different, there is no need or possibility to develop a further model C. As it happens, the only moderator for which we have found significant interaction in the effect of workload on stress is indeed student gender, no other.

- At this point our conclusion is that we have not been successful in identifying other factors than gender and prior academic ability that moderate the workload-stress process significantly. This applied not only to different policies that DP schools have in place to diminish programme load or offer support to students who experience psychological discomfort (we have not found that they make any difference to the students' well-being, but also for the school climate differences (related to peers and teachers) that were found to be related to stress levels in the cross-sectional analysis). A positive school climate may produce less stress, but it does not affect some students more than others.

CHAPTER 8: AFTER THE DIPLOMA PROGRAMME

Contents

Chapter 8: AFTER THE DIPLOMA PROGRAMME

This chapter reports on the after-effects of the DP on students. The analysis follows the students in four stages. First comes an analysis of **non-participation in DP exams**: how many of the students from the sampled schools who originally enrolled in the programme in 2016 finalized the DP and took the exams? Second are **exam passing rates**: how many of those students were successful in completing the exams and how many were not? How are exam non-participation and failure related to the workload and stress that students report during their two DP years? Third comes **exam performance**, the grade point average students obtain and how that relates to previous workload and stress levels. Fourth and finally, we look at the continuation of students (prospective) **educational career**: in which universities are DP students continuing their studies? What was their (initial) aspiration level and to what extent were they able to accomplish their aspirations? How are these aspirations related to the workload experienced during the DP? We analyse aspirations and accomplishment in relation to the experienced workload: to what extent do DP students who report excessive workloads, struggle in making their aspirations come true?
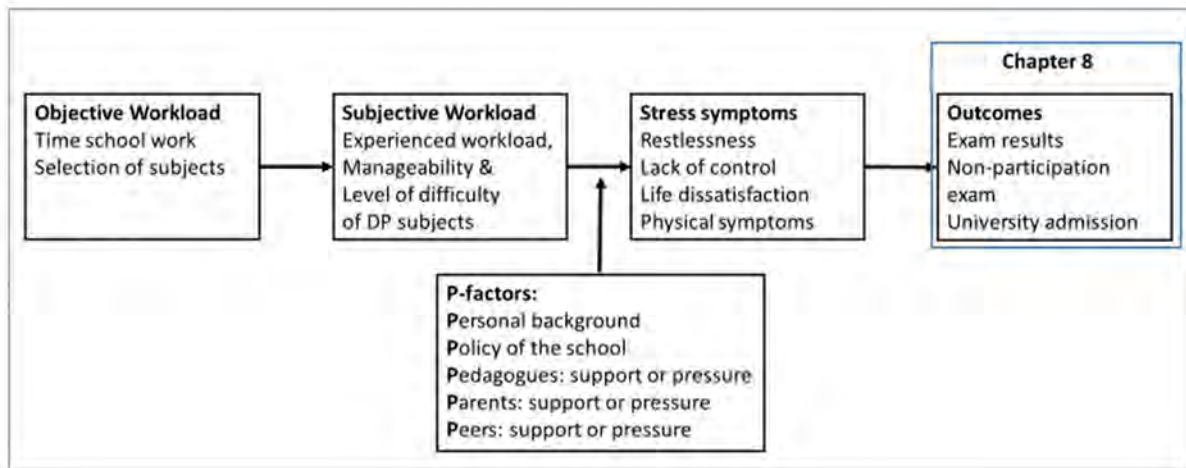
Figure 8.1: Causal model of the student objective and subjective workload and stress

We find that subjective workload and stress impact all four of these student outcome variables. **Students with higher levels of subjective workload and stress at the beginning of the DP are more not to participate in the exam**. **If they do take the exam, they obtain lower results.** In fact, this negative effect is stronger than the positive effects of academic ability or parental socio-economic status. If students struggle more with their workload in the beginning and develop stress symptoms, they are also **less likely to** realize their academic potential and **get accepted at the universities of their choice**. These results not-withstanding, however, **almost 30% of DP students aspire to continue their education at the 100 highest-ranked universities in the world, about 18% even at the top 50.**

We moreover find that DP students start aiming quite high, but during the DP lower their aspirations. These **aspirations appear to be most strongly driven by students' previous academic achievements**, with parental socio-economic status having noticeable indirect effects. **Admission to top-level academic institutions is fully determined by students' DP exam results.**

8.1 Not participating in DP exams

- While students enrol in the DP in the beginning of its first year, their actual participation in the programme remains unknown to the IB until the students enrol for the exams in the middle of year 2. For this reason, it is basically unknown how many students attempt to complete the DP, but at some point, decide to discontinue. The Workload study provides some information on this issue, as we asked DP coordinators in the sampled schools to provide us with a complete list of DP students in the beginning of year 1, November 2016.

- By comparing the complete list of DP enrolments in the sampled schools in year 1 to IB's exam records for these schools, we obtain an estimate of the number of students who – for some reason or other – discontinued the exams. While we were successful in matching the information from the sample to the exam records, this comparison is not without problems, and biases may have arisen. The main problem is that DP coordinators might have provided us with incorrect information and have included students who never actually enrolled in the DP. The reverse problem is also likely to have occurred: there may have been more DP students in a given school than the initial list of sampled students provided by the DP coordinators suggests. In fact, that this has actually happened has become clear after the matching of the sample and the exam records: there is a large number of students who completed the exams in 2018 in the sampled schools but were not listed as DP students in our 2016 sample.

- **Figure 8.2** (and Table 8.1) shows a cross-classification of the exam records and the sampling records. Of all 4,854 sampled students, 3,918 (81%) could be traced in the exam records. This suggests that **some 19%** of the students who were initially listed as DP students by the coordinators, did not take part in the exam. This **19%** may be regarded as an **upper-bound estimate** of students discontinuing the programme. The number may be obscured by several reasons other than dropout why students did not take part in the exam: eg, they may have been ill during the exam session in May 2018 or have otherwise decided not to take part and plan to take the exams later. Some students may have left their DP school during the two-year curriculum and may have continued the programme at another (DP or non-DP) school – possibly in another country, and for that reason not show up in the exam records of the sampled schools.
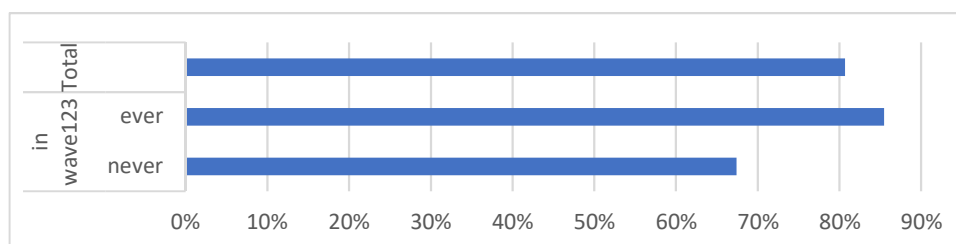
Figure 8.2: Exam participation for sampled students who did and did not respond to the surveys

- **Figure 8.2** also provides an estimate of the size of **non-coverage of DP students by our sample**. Even if this is restricted to full-DP students, the evidence suggests that not all coordinators sent us complete lists of DP students at the onset of the study. This may have happened because DP enrolment was not yet fully transparent at the time, because students transferred from other schools later or because coordinators misread our instructions[1]. As these non-covered students

---

[1] According to the instructions sent to DP coordinators and heads of school, course students were supposed to be included.

were not included in the survey to begin with, we know nothing about them other than the information in the exam record. The available information indicates that most of the non-covered students (over 85%) took less than eight subjects at the exam (79% took only one or two subjects).– This suggests that most of them might have been excluded by the DP coordinators because they were not enrolled in the full DP. The Grade Point Average (see below) of the non-covered students (4.49; 4.60 for those who took eight subjects or more) was significantly lower than for the sampled students (who obtained 4.76), which suggests that the non-coverage may also somehow have been related to the academic potential of these students.

- **Figure 8.2** refers to all students in the sample – however, not all sampled students took part in the survey (see CH3) and it seems likely that those who never responded to the survey are a select subgroup, with a higher likelihood of discontinuing the DP. **Figure 8.2** cross-classifies the exam records with the responding sample, ie, the 3,565 student who participated in the survey in one or more waves. The number of non-participants in the exams in this group provides a **lower-bound estimate of the DP dropout** and amounts to **14%**. This is almost certainly too low an estimate, but it happens to refer to those students for whom we have information about their workload experiences and backgrounds, as they provided this information in either wave 1 or wave 2 or both. They are therefore the only ones for whom we can investigate how the likelihood of discontinuing the programme is related to previous experience of workload and stress.

- **Figure 8.3** (and **Table 8.2**) shows how the likelihood of discontinuing the DP for those students who did take part in the first two waves of the Workload survey is related to workload and stress. There is indeed a statistically significant relationship between high levels of subjective workload and stress and the likelihood of not taking part in the exams. To see this clearly, we have divided the responding students into four equally sized risk groups ['quartiles'], combining information on the subjective workload and stress levels in wave 1 and wave 2. Between the lowest and highest quartiles of risk groups **the likelihood of programme discontinuation doubles from below 9% to over 18%**, but the effect is most pronounced for the highest levels of subjective workload. A linear probability model (not shown) indicates that the relationship is strongly significant (B=0.038, t=4.7, N=1,523). While it must be re-emphasized that our measurement of drop-out is far from perfect, the result appears to be of great relevance. Not only does it show a relationship between workload and DP completion, it also suggests that this relationship can be diagnosed at a relatively early stage of the programme.
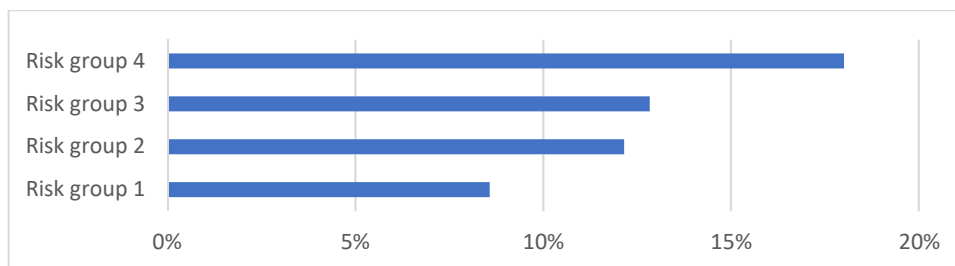
Figure 8.3: Non-participation in exams (not in exam records) by Subjective Workload * Stress in waves 1-2

- We have added control variables to the linear probability model (not shown) to examine whether prior conditions confound or mitigate the relationship between high risk in the early DP career and non-participation in the exam records but have found no differences for prior academic

achievement or parental socio-economic status [SES]: there is **no evidence that academically less able students or students from less favourable socio-economic backgrounds are less likely to take the exams**. The clearest predictor of non-participation we have been able to identify is indeed how demanding students find the DP workload in their first year.

8.2 Taking the exam and making it

- The DP exams are graded on a 1-7 scale for the mandatory subjects from the six DP subject groups; one of the three core subjects, TOK, is also graded, but on a different scale, A-E, which for examination purposes is transformed into a zero/one grade. The other two core DP subjects, EE and CAS are also graded by awarding a pass or fail (zero/one). A DP student can thus obtain a maximum of 45 points, and 24 (53.3%) points is the passing grade. However, the exact calculation of total points earned is rather complicated for the exam records as we received them. Complicating is in particular that not all students complete the full programme in a given year: some may have completed subjects in a previous exam session, others might do in a following session.

- We have therefore simplified the construction of the exam results for our analysis by calculating a **Grade Point Average [GPA] over all graded exam subjects**, which was seven (six regular subjects and TOK) for most students. As averages (unlike sums) are insensitive to the number of subjects taken, calculating a GPA facilitates the valuation of the exam results considerably. When we apply the 53% passing grade to the calculated GPA, we obtain an estimate that about 82% of all students who sat the exams, obtained the GPA passing grade (3.73). While we acknowledge that our GPA criterion is not identical to the actual procedure followed by the IB, we feel that the two procedures are close enough to warrant the choice of GPA as a criterion to judge students' academic achievement.

- **Table 8.3** gives the **mean GPA and the associated estimates of exam passes and failures** respectively for the students from the responding sample. Unlike non-participation in the exam – ie, not sitting the exams – the exam results are clearly related to both their academic achievement of the students before they entered the programme (beta=0.24, t=11.7) and the socio-economic status [SES] of their parents (beta=0.1, t=4.9), with prior academic achievement being the strongest predictor. The question then is how exam results are related to earlier experiences of workload and stress. Here we find a **strong negative effect of the subjective workload** (beta=-0.3, t=5.5) measured, which summarizes again what the students have reported in Waves 1-2 and was coded into quartiles coded between 0 and 1. Students in the highest quartile of subjective workload obtain about a third of a point (of the GPA) lower exam results than the students in the lowest quartile. If expressed in the likelihood of passing the exams, the results are much weaker (which reflects the loss of statistical power when a dependent variable is dichotomized) but still significant. Students who struggle with their workload are also less likely to pass their exams.

8.3 Educational aspirations and actual access to university

- The DP is a worldwide university-preparatory programme and many students opt for it in order to do well in entrance procedures for universities, which may include some of the world's leading academic institutions. In wave 1, over 95% of the responding students indicated that they were

planning to continue to some tertiary education, and **a vast majority (over 80%) said they were aspiring a master's degree or higher** (**Table 8.4**). We asked the sampled DP students in the first and the third wave about their aspirations at the time – which *type* of university they would like to continue to after the DP. **Table 8.5** shows the distributions of the students at the two time points and also the terminology used in these questions. We did not ask the students to identify particular educational institutions in the first wave and it remains up to interpretation what they may have understood to be a **"good", "very good" or "top-level" university**, but at both occasions **over 95% elected one of these three options**. However, there is a clear indication that at wave 3 the aspirations had been adjusted to reality and the percentage of DP students who aspired to go to a "top-level" university had halved from 19% to 9%.

Table 8.6 presents a better perspective on how these aspirations panned out for students by the end of DP year 2. At that point in time we asked the students also to identify the exact institution they were planning to attend after DP, using an open question. We received over 1,800 alternatives to code. They were classified in accordance with the Shanghai Academic Ranking of World Universities (2019)[2]. The Shanghai list identifies the academic qualities of 500 universities, led by Harvard University as #1 to West-Virginia University at #500. It should be noted that individual ranks are only given to the top 50 universities – for the remainder the ranking refers to quantiles (groups of 50 universities that are regarded *ex aequo).* While being produced in China, the Shanghai list is clearly better attuned to universities in North-America than elsewhere, and this may influence our results. North America is, however, also the world in which a large proportion of DP schools are located. At the same time, the Shanghai List's methodology ensures that no major academic institution in the world would be missed: the quality of a university is determined as a weighted average of (Nobel) prizes, highly cited faculty members and publications, and papers published in the leading academic journals Science and Nature.

- **Table 8.6** provides a qualitative impression of the answers given by the DP students who participated in the third wave. The answers are grouped by the level of the universities on the Shanghai List, but we have labelled the categories based on the terminology used in the rest of the questionnaire for aspirations. According to this convention, **almost 30% of the DP students anticipate becoming enrolled in one of the top 100 universities in the world, and about 18% even in the top 50** – which are by all means leading academic institutions. At the other end, over one third of all DP students plan to attend universities that have not made it to the top 400 of the Shanghai listing.

- How are educational aspirations related to pre-DP conditions and how did experience and performance during the programme affect the outcomes? **Table 8.7** unfolds a series of linear models that summarize the statistical relationships and lend credibility to some causal interpretation. The dependent variable here is the broad aspiration, that was asked in the same format in wave 3 and wave 1. Unsurprisingly, in Model A we see that students who regarded themselves academically more able prior to the DP have higher levels of aspiration, also when they reach the end of the programme. There is evidence that these higher aspirations are more prominent among students with better educated and higher status parents. There is no difference in aspiration between boys and girls. Model B shows that these aspirations are also strongly determined by the exam results: students who score higher at the exams have higher ambitions. Model C indicates that the ambitions are not affected by the stress levels that these students

---

[2] http://www.shanghairanking.com/index.html

reported in the previous waves. The table includes all stress measures accumulated over three waves, but essentially the same result is found when assessed by stress levels in the previous wave separately. Model D, finally, brings in another strong predictor of aspirations at wave 3: the aspirations in wave 1. This is no surprise – the point of model D is that it allows us to examine the possible moderation of aspirations by stress: are stressed students more likely to lower their aspiration? The interaction is not statistically significant. **Stress does not affect the development of educational aspirations.**

- **Table 8.8** shows three stepwise models for the choice of university (ranked by the (reversed) Shanghai score), that the DP students say they were ultimately admitted to. Notice that we are not looking at aspirations here as such, but at aspiration come true. After some experimentation, we transformed the Shanghai ranking of these universities by square root transformation (before reversing direction, which makes the analysis particularly sensitive to differences between universities at the top end of the scale. Model A shows that the ranking of the university of admission is not related to self-assigned academic ability but that it is significantly related to the socio-economic status of the parents. Model B shows that the rank of the university of admission is strongly affected by the exam results – and this makes the effect of socio-economic status of the parents insignificant. One way to formulate a conclusion is to say that in the end **only achievement matters for DP students to get access to the world's leading academic institutions**. Finally, model C explores whether the stress experiences in the DP career is related to the choice of university – but this does not seem to be the case. If anything, there a (very weak) positive relationships between stress and choice of high-ranking academic institutions. If so, a possible interpretation is that high ambition makes DP students nervous. This interpretation, however, is somewhat less plausible, as the control for prior aspirations (ie, the variables analysed in **Table 8.7**) makes the positive relationship between stress and the choice of high-ranking academic institutions actually statistically significant (beta=0.086, t=2.1). This indicates that the higher stress of these students is not produced by their higher aspirations, but rather that these aspirations (which themselves are a strong predictor of final choice), suppress the relationship between stress and academic choice. Either higher stress helps these students to realize their ambitions, or the actual admission to a top-level university produced the stress.

Chapter 9: SCHOOLS' PRACTICES AND EVALUATIONS OF PROPOSALS TO ADDRESS WORKLOAD ISSUES

Contents

Chapter 9: SCHOOLS' PRACTICES AND EVALUATIONS OF PROPOSALS TO ADDRESS WORKLOAD ISSUES

In this chapter we report on the **policies and services schools currently have in place to support students in their academic and social-emotional needs and how much services are actually used by students**. Moreover, we display the support of schools and students for various possible ideas to either reduce the workload directly or to make it more manageable for students. As the broader sample of DP coordinators covers a much wider range of schools (N = 1,870) than the student sample (N = 96) and hence provides a more accurate view, we report the opinions of DP coordinators from the broader sample. Where possible, however, we complement this information with students' opinions from the sample schools.

A first important way of addressing workload issues is **prevention**. Roughly two third of DP coordinators believe that **guiding students in their programme and/or subject choice** as well as **training** them **in organizational skills** could be effective ways of averting problems. Another important aspect of prevention is the **planning and coordination of assessments and the restriction of homework**. While schools excel in coordinating deadlines, with roughly three quarters of the schools engaging in at least coordinating the deadlines for IAS and Core components, homework restrictions are much less popular and practiced by less than half the schools. When presented with different options of how to make the workload more manageable, homework-free days or periods do, however, rank high on the students' list right after having a clear schedule of assessment deadlines, which tops it.

A second way of addressing workload problems is the **provision of study and psychological support for students**. Well over 90% of DP schools provide some form of **academic counselling and study facilities**, helping students with their academic work as well as with the choice of an appropriate follow-up education. When available, these provisions tend to be well used by students; in particular, study rooms, university counselling and teacher-led coaching are popular, with about 70% of the students profiting from these services. Almost all schools provide some kind of **professional health or well-being service**, such as a school nurse or counsellor to help students when they run into problems. These latter services are used by about 40% of the students.

A third and last way of addressing students' workload is specific to the **assessment workload**. In a minority of schools **IAs or other non-exam assessments are brought forward** and finalized either in year 1 or by November of year 2 in order to ease the burden during the peak period in year 2. This measure is valued by about 40% of the students.

While no clear preference emerges regarding the best way to regulate exams and other assessments, both coordinators and students believe that having (more) **IB-approved subject or online resources** is the most important way of making the DP workload more manageable. Interestingly, students rank the provision of resources (slightly) even higher than a **reduction of content in DP subjects**, which comes third.

9.1 Information and guidance to help students make appropriate study choices

- It is up to schools whether they consider entrance requirements for the Diploma Programme. DP Coordinators were asked what kind of factors were considered at their school before students are admitted to the DP. **Figure 9.1** lists the requirements that schools use to screen students (see also **Appendix 9.1**).

- The most common **entrance requirement** is the **student's record of academic achievement**. At 69.1% of the schools in our sample this is always considered, at 20.3% sometimes. Another common admission factor is an **interview** with the student. 55.9% of the schools always take an interview, 26.8% sometimes and only 17.3% never. At about a third of the schools a **language test** is always taken (35.4%) while one third of schools never include a language tests in their entrance requirements (36.7%). More than two thirds of the schools ask for **recommendations of feeder schools** (32.8% always and 35.5% sometimes) and an **ability test (eg, IQ-test)** is considered when students are admitted at almost half of the schools in our sample (22.1% always, 20.2% sometimes).

- At 97 (7.1%) of the schools none of these entrance requirements are applied.
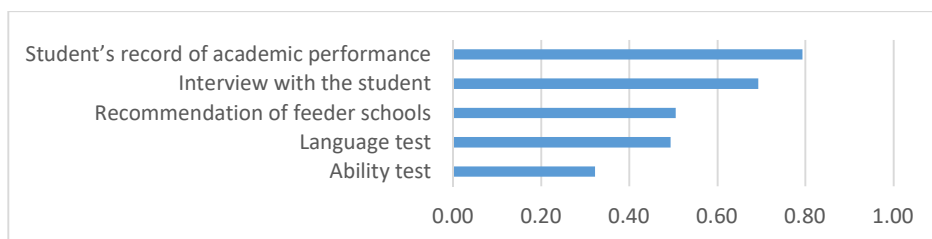
Figure 9.1: Schools' entrance requirements: factors considered when students are admitted to the school

- DP coordinators were asked to give their opinion on the effectiveness of possible ways to guide students in their programme and subject choices as well as prepare them for their study: training students' organizational skills, providing guidance to aspiring DP students about the suitability of the programme for their learning needs, advising students on their subject choice once they start with the programme (eg, which subjects they could best pursue at higher level). The idea behind this kind of guidance and preparation is that students have different needs and abilities and problems can be avoided if they make appropriate choices and know how to organize their studying.
- Figure 9.2 shows that the majority of DP coordinators sees virtue in such practices (the average score is above the scale midpoint). Two thirds of the DP coordinators (68.0%, 65.2%, and 64.9% respectively) indicate that they believe that these measures would be helpful or very helpful in preventing workload problems for students (see also Appendix 9.2).
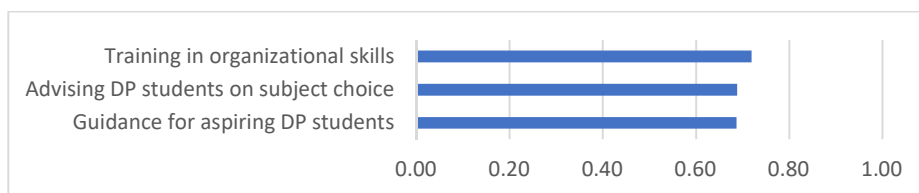
Figure 9.2: DP coordinators' opinions: The effectiveness of guidance and study preparation

9.2 Coordination and planning to balance the workload for DP students

- **Figure 9.3** describes the spreading of internal assessments and core components over the 2 years of DP. The figure shows that the number of assessments (IAs) and core components increases slowly over the course of year 1, starts off in year 2 at a higher rate, and reaches its maximum in the period December-February in year 2 (see also **Appendix 9.3**).
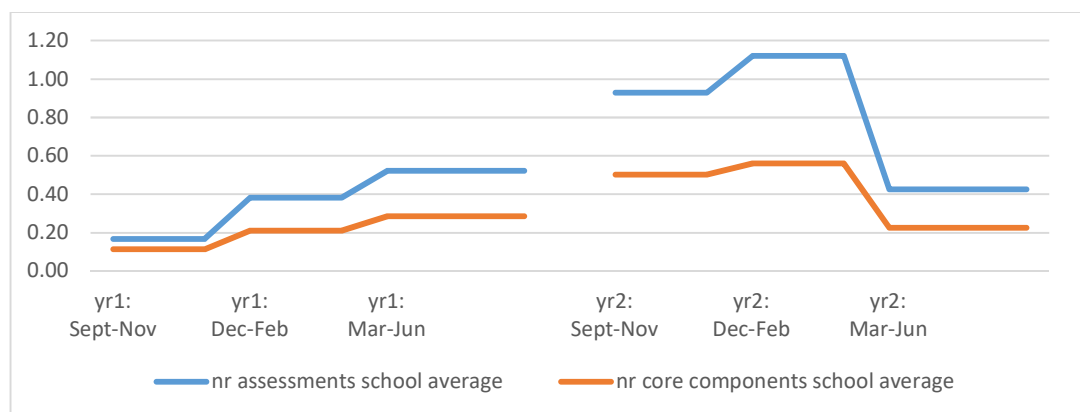
Figure 9.3 Number of internal assessments and core exam components handed in, estimates by DP coordinators (monthly average)

- DP coordinators were asked about how their school **organizes and coordinates assessment dates and homework**. **Figure 9.4** (and **Appendix 9.4**) describes different ways of organizing and planning that are practiced by the schools in the broader sample. The schools' degree of assessment **coordination** is rather high: the average score is above the midpoint of the scale (.5). The majority of the schools always coordinates deadlines for IAs and core components, 72.1% and 76.0% respectively. For school-based assessments this percentage is a lot lower, 38.1%.

- Schools also organize **collaborative meetings on planning issues**, the average score is around the midpoint of the scale (.5) which means that schools regularly hold such meetings, ranging from three or four times a year to once a month. In a minority of the schools, restrictions to students' homework are implemented. 19.7% of DP schools have **homework free days**, 25.1% **homework free periods**.
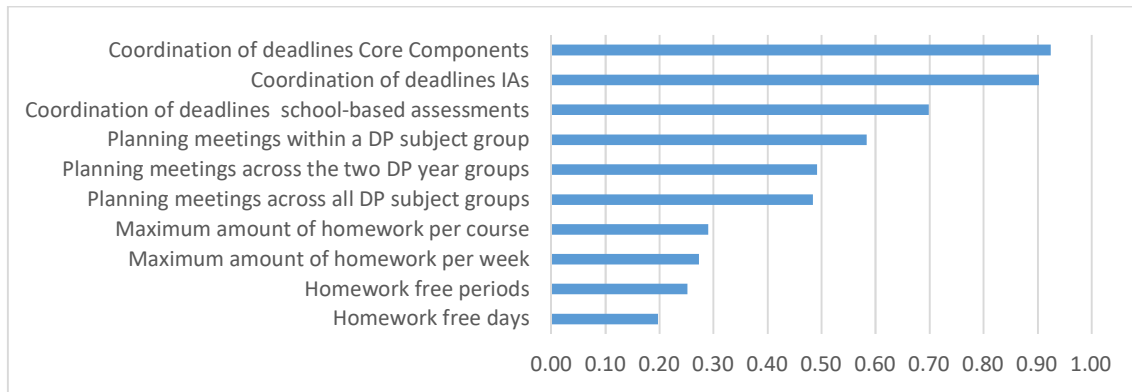
Figure 9.4: DP coordinators' practices: Coordination and planning to support a balanced workload for DP students

- **Figure 9.5** shows how helpful students think the coordination of deadlines and homework restrictions are or would be in making the workload more manageable. There is not much differentiation between the items. The average scores are just above the midpoint of the scale, meaning that students have some confidence that these options would help them manage their workload (see also **Appendix 9.5**).
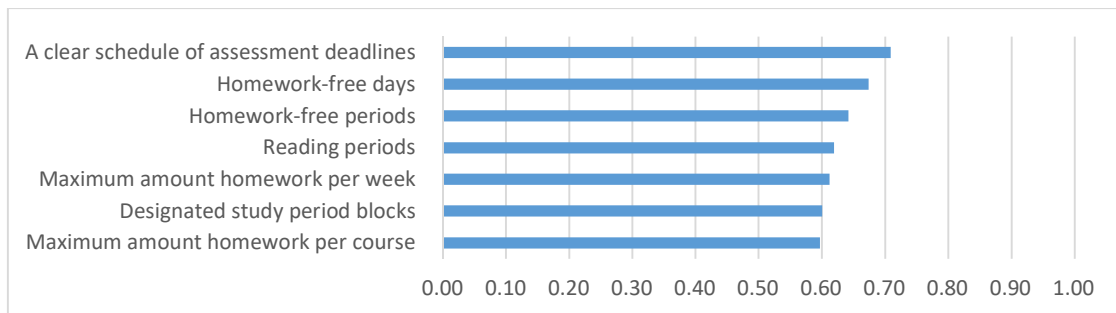
Figure 9.5 Students' opinions: Options to make the workload more manageable for students

- **Figure 9.6a** displays the schools' practices with regard to the **concentration or spreading of IAs and (non-exam) assessments**. Almost half of the schools (46.0%) indicate taking measures to spread IAs and other assessments by finalizing a part of them in year 1 or in November of year 2 respectively. Finalizing IAs in year 1 is the most frequent of these measures; 35.4% of the schools finalize a part of the IAs in year 1 (see also **Appendix 9.6**).
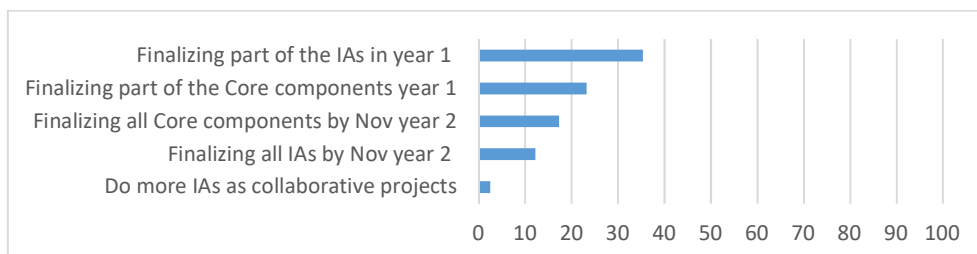
Figure 9.6a: School practices: Spreading students' workload to make the (non-exam) assessment workload more manageable for students

- **Figures 9.6b and 9.6c** present the degree to which DP coordinators and students think these measures are helpful. On average, both DP coordinators and students are in the middle of the scale, **'somewhat helpful'**. Students, however, seem to be somewhat more positive about these measures than DP coordinators.
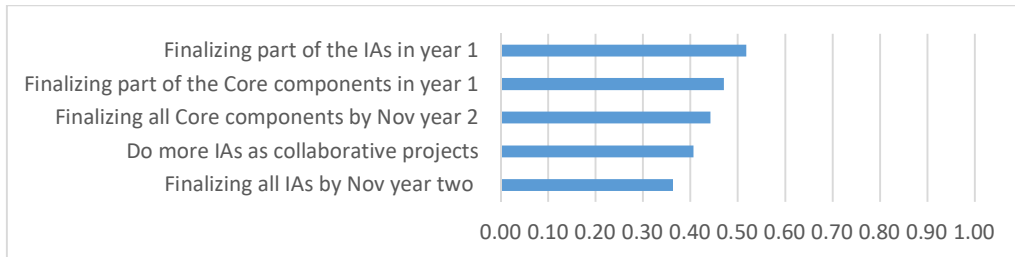
Figure 9.6b: DP coordinators' opinions: Spreading students' workload to make the (non-exam) assessment workload more manageable for students
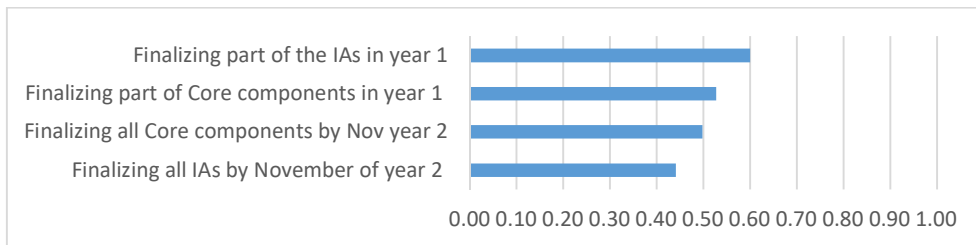
Figure 9.6c: Students' opinions: Spreading students' workload to make the (non-examination) assessment workload more manageable for students

- **Figure 9.7a** displays the opinions of the DP coordinators on various ways of changing the timing and spreading of the SL exam papers, geared at making students' **exam workload** more manageable. Overall the suggestions are only supported by a minority of DP coordinators. For each option less than a quarter of the DP coordinators thinks this would be effective. In **Figure 9.7b** the same suggestions were evaluated by the students, who are generally much more optimistic that these would help them with their exam workload. In fact, the student scores are almost twice as high as those of the coordinators. More than 60% of the students, for example, think that spreading exam papers over six instead of three weeks would make the DP exam workload much more manageable (see also **Appendix 9.7)**.
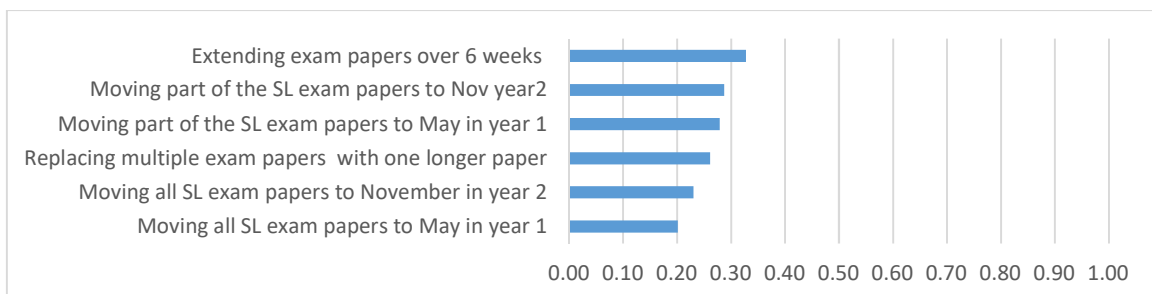
Figure 9.7a: DP coordinators' opinions: options to make DP exam workload more manageable
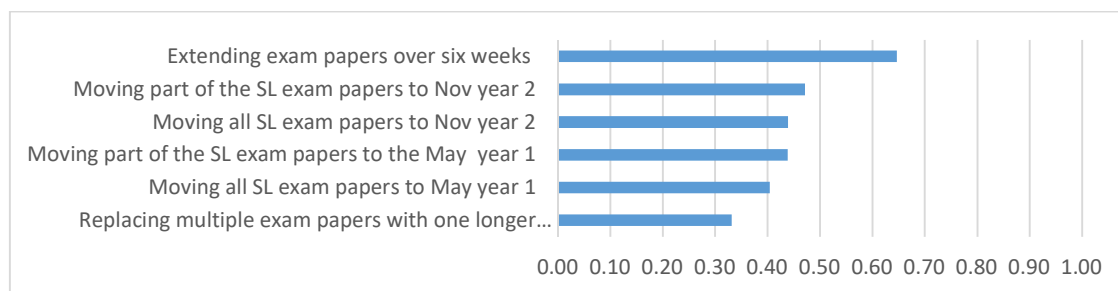
Figure 9.7b: DP students' opinions: options to make DP exam workload more manageable

9.3 Support services for students

- DP coordinators were asked for the presence of support services at their school to help students with their school work (ie, coaching, tutoring, additional lessons) and to promote their health and wellbeing (ie, the presence of mental and physical health professionals).

- **Figure 9.8a** shows that a large majority of schools (87.4%) offer **university counselling** and this way help to prepare students for their follow-up schooling career. Regarding the DP school work, the majority of the schools offer **teacher-led coaching**, both in groups (80.0%) and individually (78.9%). 73.7% of the schools offer coaching aimed at special needs (eg, dyslexia, Attention Deficit Hyperactivity Disorder). See also **Appendix 9.8.**
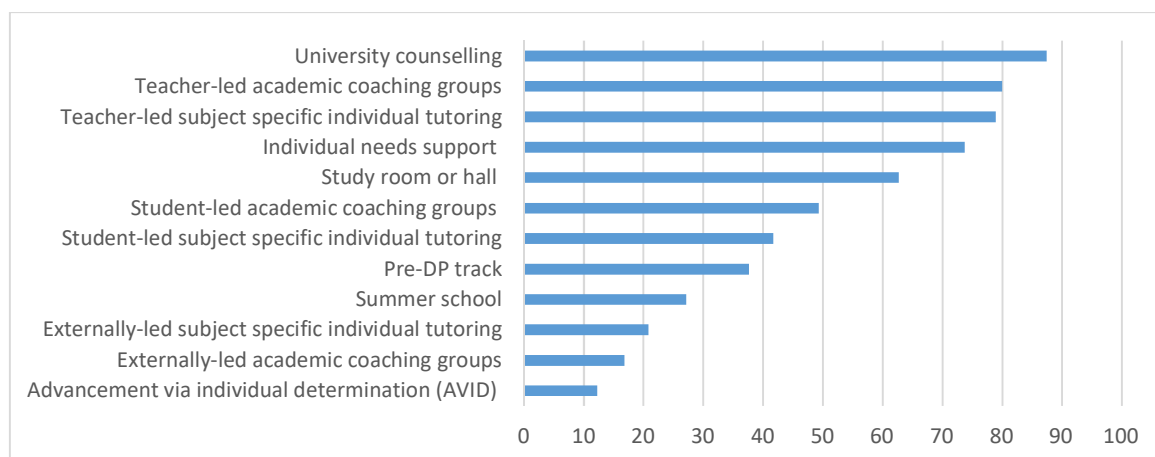
Figure 9.8a: Schools' support services on students' needs beyond the classroom, according to DP coordinators (percentages)

- **Student-led coaching**, in groups or individually, is also provided, at respectively 49.3% and 41.7% of the schools. Externally-led coaching is much less common, offered only at 16.8% (in groups) and 20.8% (individually) of the schools. Overall, **94.9% of the schools offer at least some type of coaching**, be it in groups or individually, and teacher-, student- or externally led (on average 2.9 out of these 6 types).

- **Figure 9.8b** displays the presence of different types of mental and physical **health professionals** at schools. At most schools a school counsellor is available (85.2%). Also, at many schools a nurse (67.2%), a psychologist (59.3%) or an internal student welfare coordinator/officer (59.3) is present. Quite a lot of schools report to have a wellbeing team (excluding students (45.0%) or including students (25.7%)), or an internal welfare coordinator (55.6%). Overall, **69.8% of the schools report to have at least one of these kinds of professional wellbeing support in place[1].** Almost all schools, **98.9%, have some health or wellbeing professional or service** in place (on average 4.5 out of these 10 professionals/services).
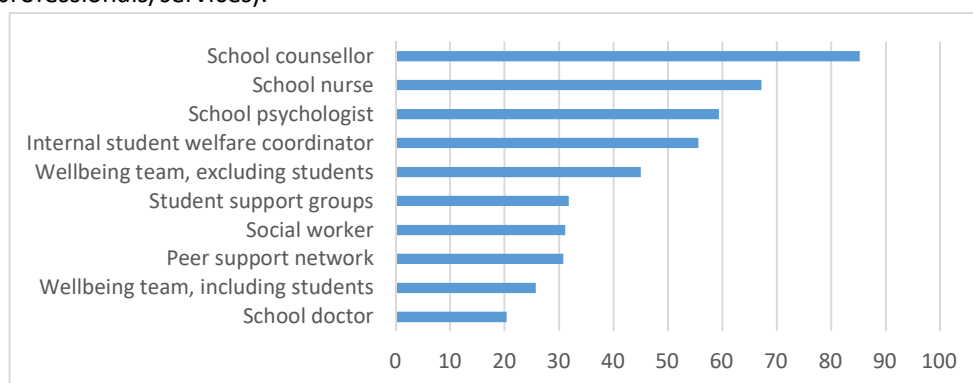
Figure 9.8b: Schools' support services on wellbeing and health, according to DP coordinators (percentages)

- **Figures 9.8c and 9.8d** show the use of these support services by students, if present at their school. Services geared at students' academic needs are more widely used than services geared at students' emotional wellbeing. **Study rooms, university counselling and teacher-led academic coaching groups top the ranks**, with close to 70% of the students using them. School counsellors and nurses are frequented by roughly 40% of the students.
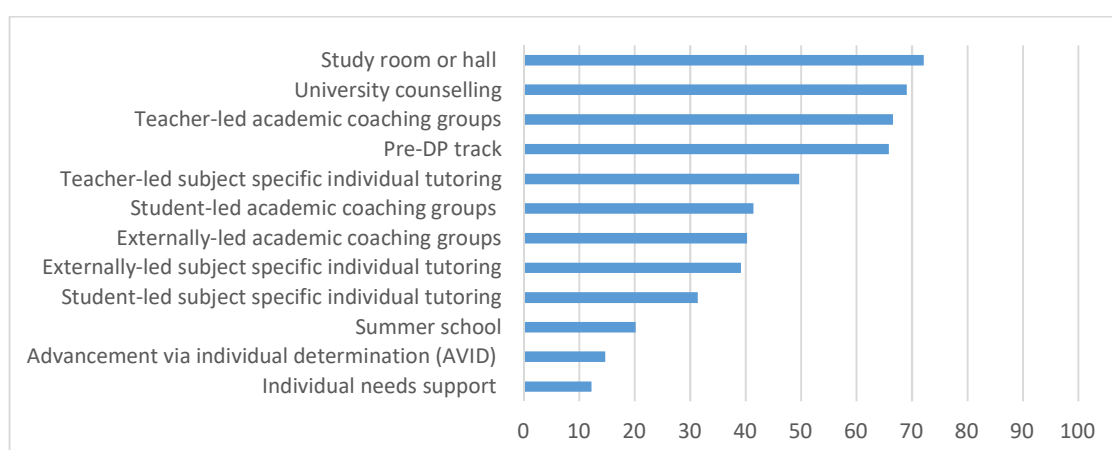
Figure 9.8c: Schools' support services on students' needs beyond the classroom, used by students if present

---

[1] 27.1% of the DP coordinators reports to have one of the three types of wellbeing professionals (a wellbeing team excluding students, a wellbeing team including students, or an internal welfare coordinator); 30.1% reports to have two; 12.7 indicates to have all three.
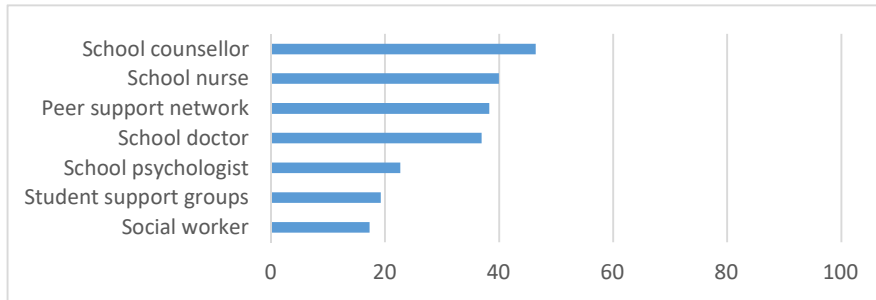
Figure 9.8d: Schools' support services on wellbeing and health, used by students if present

- **Figure 9.9a** describes the provision of school support programmes for students, concerning mental health, learning skills, and relaxation. About half of the schools offer mental health (eg, social and emotional learning [SEL]) (49.7%) or relaxation programmes (eg, mindfulness or yoga) (53.8%); 61.6% of the schools provides a learning skill programme (eg, time management, organization or communication). **Figure 9.9b** shows that if these programmes are available, DP coordinators estimate students' participation in such a programme to be around one third (see also **Appendix 9.9**).
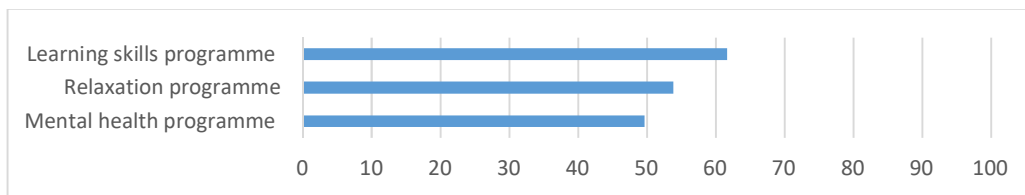
Figure 9.9a: Provision of schools' support programmes
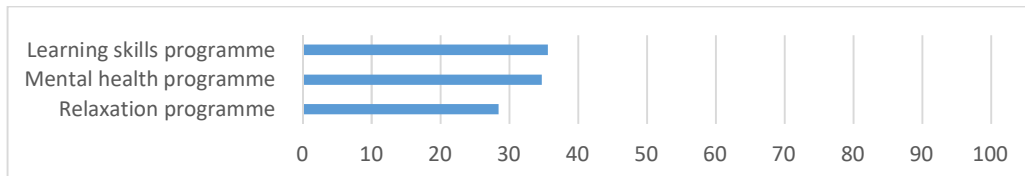
Figure 9.9b: Schools' support programmes: student participation if provided, estimated by DP coordinators

---

The questions on the support services to the DP coordinators were asked in three batteries, the first containing three questions (on the presence of an internal student welfare coordinator/officer, a wellbeing team, including students, and a wellbeing team, excluding students) with response categories yes-no. For the remaining questions DP coordinators were just asked to indicate the presence (or absence) of the service in question.

Missing values were only considered as 'not present' if at least one question of the first three question batteries was answered with 'yes' or 'no' (in order to not mistakenly interpret a partial non-response as a 'no' answer).

9.4 Proposals to adjust the DP workload

- Finally, DP coordinators and students were asked about how they think potential changes or additional assessment options would affect students' workload. **Figure 9.10a** presents the **DP coordinators' views** on several **options on swapping an exam paper for an IA or vice versa** to make the DP workload more manageable for students. On average the responses of the DP coordinators on the four options hover around the midpoint of the scale; about one third thinks the measure would be helpful, one third thinks is would not, and one third is undecided. Still, 'Replacing an exam paper with an additional IA' is the least preferred, with only 20.2% of the coordinators thinking this would help to reduce students' workload (see also **Appendix 9.10**).
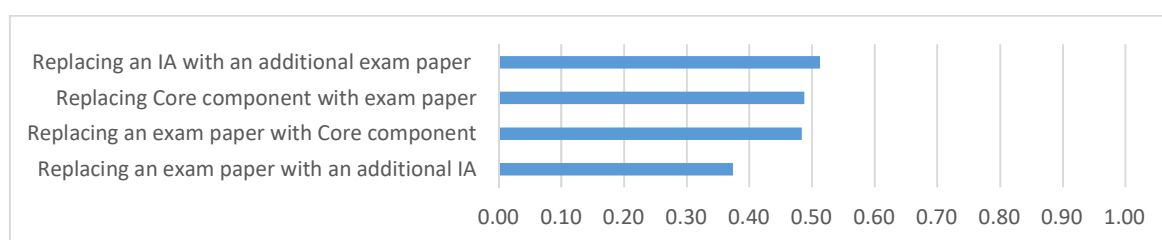
Figure 9.10a: DP coordinators' opinions: Options to make the workload more manageable for students

- **Figure 9.10b** presents **DP students' views** on these options on swapping assessments in order to make the workload more manageable. Less than half the students (34.0% and 38.7% respectively) think that replacing an exam paper with an additional IA or another non-examination assessment would be helpful. Only a small minority of students thinks that, the other way around, an additional exam paper instead of an IA (9.6%) or a core exam component (7.2%) would be helpful (see also **Appendix 9.10**).
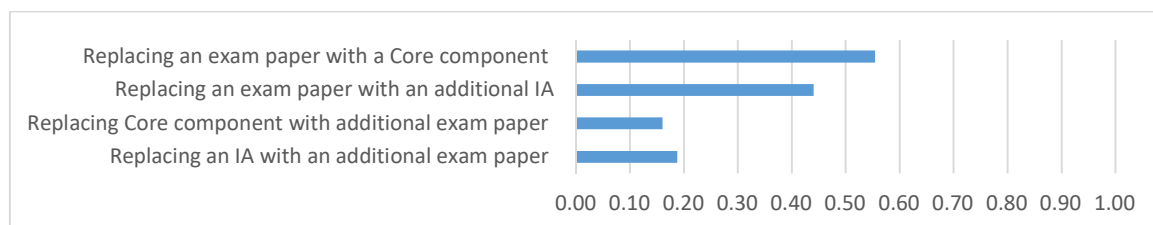
Figure 9.10b: Students' opinions: Options to make the workload more manageable for students

- **Figure 9.11a** displays **DP coordinators' opinions** on another set of possible options to help students with their DP workload (see also **Appendix 9.11**). All measures are on average above the scale midpoint (.5), indicating that on average the DP coordinators are positive on these measures. Most support is for the options of providing **additional IB-approved resources**: two thirds think that these resources (online 65.6% or otherwise 64.2%), would make the workload more manageable. A small majority of 58.6% indicates that reducing the amount of content in DP subjects would help to make the DP workload more manageable, which is relatively low, meaning that in the opinion of DP coordinators DP workload issues are not necessarily caused by subjects being overloaded with content.
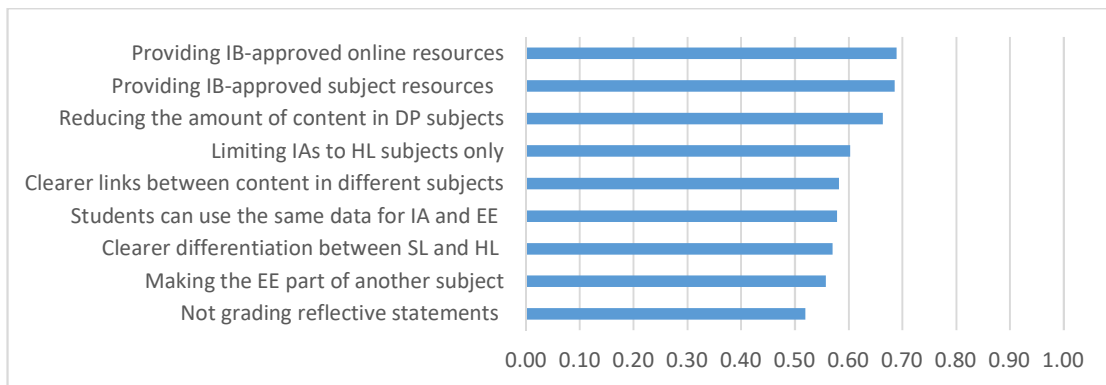
Figure 9.11a: DP coordinators' opinions: options to make DP workload more manageable

- **Figure 9.11b** describes **students' opinions** on some **options to make the DP workload**. Like the DP coordinators, the majority of the students, and even a larger majority (79.5% online; 76.1% other resources) think that providing additional IB-approved resources would help to make the workload more manageable. Interestingly, even for students, 'Having less content to study' is not the most important way to make the DP workload more manageable. Only 57.9% think this would be helpful, still more, however, than any of the remaining seven alternatives. 'Having IAs for HL subjects only', 'Being allowed to use the same data for the Internal Assessment (IA) and the EE for the same DP subject' and 'Having reflective statements not graded' get support from just under half of the students (49.8%, 49.1%, and 46.1% respectively). Other suggestions are perceived to be even less helpful; only about a third of the students think these would help them in managing their workload.
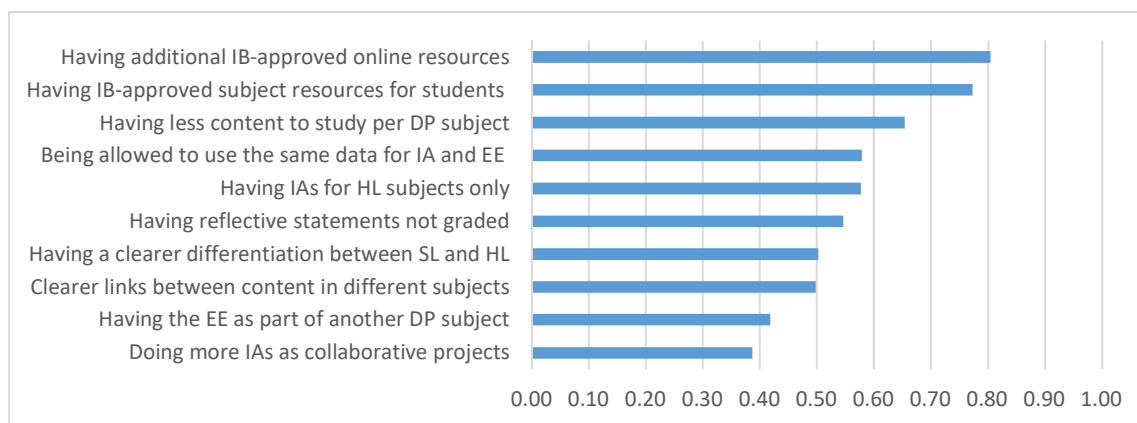
Figure 9.11b: Students' opinions: options to make DP workload more manageable

APPENDIX A: Additional tables

CHAPTER 3

Appendix 3.1: Language choice by region		English	Spanish	French	
	Western Europe	257	0	7	264
		97.3%	0.0%	2.7%	100.0%
	Southern Europe	56	57	0	113
		49.6%	50.4%	0.0%	100.0%
	Eastern Europe	92	0	0	92
		100.0%	0.0%	0.0%	100.0%
	North America	707	0	26	733
		96.5%	0.0%	3.5%	100.0%
	Middle America	22	48	0	70
		31.4%	68.6%	0.0%	100.0%
	South America	31	148	0	179
		17.3%	82.7%	0.0%	100.0%
	Oceania	3	0	0	3
		100.0%	0.0%	0.0%	100.0%
	Asia	263	0	0	263
		100.0%	0.0%	0.0%	100.0%
	Africa	31	0	2	33
		93.9%	0.0%	6.1%	100.0%
	Middle East	119	0	1	120
		99.2%	0.0%	0.8%	100.0%
Total		**1,581**	**253**	**36**	1,870
		84.5%	**13.5%**	**1.9%**	100.0%

Appendix 3.2: School status by region				
			Legal status	
		State	**Private**	**Mixed funding**	
	Western Europe	98	153	13	264
		37.1%	58.0%	4.9%	100.0%
	Southern Europe	22	88	3	113
		19.5%	77.9%	2.7%	100.0%
	Eastern Europe	31	59	2	92
		33.7%	64.1%	2.2%	100.0%
	North America	596	105	32	733
		81.3%	14.3%	4.4%	100.0%
	Middle America	3	66	1	70
		4.3%	94.3%	1.4%	100.0%
	South America	90	84	5	179
		50.3%	46.9%	2.8%	100.0%
	Oceania	0	3	0	3
		0.0%	100.0%	0.0%	100.0%
	Asia	16	242	5	263
		6.1%	92.0%	1.9%	100.0%
	Africa	0	32	1	33
		0.0%	97.0%	3.0%	100.0%
	Middle East	5	114	1	120
		4.2%	95.0%	0.8%	100.0%
Total		**861**	**946**	**63**	**1870**
		46.0%	**50.6%**	**3.4%**	**100.0%**

Appendix 3.3: Size of DP enrollment							
		N DP- coordi-nators	**min**	**max**	**average**	**median**	**Std. Dev.**
q07a	Total DP enrollment year 1 full-time	1,435	0	1,750	45.06	30	*71.51*
q07b	Total DP enrollment year 2 full-time	1,435	0	439	37.15	26	*37.19*
q07c	Total DP enrollment year 1 part-time	1,435	0	700	19.13	0	*53.19*
q07d	Total DP enrollment year 2 part-time	1,435	0	440	17.75	0	*46.47*

Appendix 3.4: Size of DP enrollment of full-time/part-time DP students in year 1 and 2

	DP enrollment			
	year 1 full-time	year 2 full-time	year 1 part-time	year 2 part-time
none	59	48	922	821
up to 50 students	1,045	1,131	462	568
up to 100 students	325	277	81	78
up to 200 students	100	83	58	61
up to 1,750 students	18	8	24	19
none	3.8%	3.1%	59.6%	53.1%
up to 50 students	67.6%	73.1%	29.9%	36.7%
up to 100 students	21.0%	17.9%	5.2%	5.0%
up to 200 students	6.5%	5.4%	3.7%	3.9%
up to 1,750 students	1.2%	0.5%	1.6%	1.2%

Appendix 3.5: Language choice of school and percentage of students whose first language is not language of instruction

	English	Spanish	French	Total
Less than 10%	337	184	9	530
	30.2%	*88.0%*	*31.0%*	*39.1%*
Between 10 and 20%	136	7	0	143
	12.2%	*3.3%*	*0.0%*	*10.6%*
Between 20 and 40%	150	2	5	157
	13.4%	*1.0%*	*17.2%*	*11.6%*
Between 40 and 60%	133	6	8	147
	11.9%	*2.9%*	*27.6%*	*10.8%*
More than 60%	361	10	7	378
	32.3%	*4.8%*	*24.1%*	*27.9%*
Total	1,117	209	29	1,355
	100.0%	*100.0%*	*100.0%*	*100.0%*

Appendix 3.6 Academic abilities before IB

varname	Item	N	mean	stddev	% 2 highest categories
a38/b57	Average grade (mark)	1,923	3.99	0.98	72.4
a39a/b58a	English	1,906	3.45	1.09	48.5
a39b/b58b	Mathematics	1,890	3.42	1.21	50.3
a39c/b58c	Science	1,878	3.45	1.14	51.0
a40/b59	Academic abilities	1,929	3.50	0.99	52.4
a41a/b60a	Language ability: understanding	1,936	4.19	0.89	78.1
a41b/b60b	Language ability: reading	1,934	4.12	0.95	75.2
a41c/b60c	Language ability: speaking	1,933	4.06	1.04	71.9
a41d/b60d	Language ability: writing	1,930	3.93	1.08	65.3

Appendix 3.7 Home situation		N	%	range
		N	%	range
a79a/b62a	lives with mother	1,783	96.7	(0-1)
a79b/b62b	lives with father	1,744	87.5	(0-1)
a79c/b62c	lives with brother(s)	1,569	51.9	(0-1)
a79d/b62d	lives with sister(s)	1,560	51.7	(0-1)
a79e/b62e	lives with grandparent(s)	1,401	9.4	(0-1)
a79f/b62f	lives with other(s) (cousins parer	1,391	9.7	(0-1)
constructed :	constructed :			
withparents	lives without parents	1,787	1.6	(0-1)
withparents	lives with one parent	1,787	14.9	(0-1)
withparents	lives with two parents	1,787	83.5	(0-1)
a50	on a boardingschool	1,622	14.4	(0-1)
a51	internal on a bordingschool	1,613	10.0	(0-1)
aggregate (a50	school is boardingschool	93	14.0	(0-1)

Appendix 3.8 Parents' education and occupation

a84/b66	mothers education	%		a87/b69	fathers education	%	
	Did not complete high school	3.1			Did not complete high school	4.1	
	High school / general educational development (GED)	11.3			High school / general educational development (GED)	9.6	
	Some higher education	16.3			Some higher education	12.3	
	Bachelor's degree	37.2			Bachelor's degree	31.9	
	Master's degree	22.9			Master's degree	26.9	
	Advanced graduate or PhD	9.2			Advanced graduate or PhD	15.2	
	N	1,766	69.3		N	1,733	74

a85/b67	mother's labor market position	%		a88/b70	father's labor market position	%	
	working full time for pay (4 or more working days)	55.8			working full time for pay (4 or more working days)	88.6	
	working part-time for pay (less than 4 working days)	10.6			working part-time for pay (less than 4 working days)	3.5	
	not working, but looking for a job	4.3			not working, but looking for a job	2.4	
	other (home duties, retired)	29.3			other (home duties, retired)	5.5	
	N	1,730			N	1,690	

Appendix 3.9: Student motives for pursuing the DP				N	mean	stddev	% 2 highest categories
A08f	Because we are likely to move to another country			1974	1.70	1.00	21.8%
A08e	Because my parents wanted it			1974	1.95	0.98	27.1%
A08h	The geographical vicinity of the DP school			1969	1.96	0.98	28.3%
A08g	The idea to be with like-minded students			1975	2.65	0.98	57.5%
A08d	The international orientation of the programme			1976	2.91	0.95	68.1%
A08c	The opportunities for personal development			1981	3.11	0.89	77.1%
A08b	The opportunities to eventually get a good job			1976	3.26	0.82	83.3%
A08a	The opportunities to enter good universities			1983	3.53	0.68	92.2%
a	Note: Answering categories 1 = not important, 2 = slightly important, 3 = important, 4 = very important						

Appendix 3.10: Factor analysis students' motives			
	factor loadings		
range	factor 1: personal development	factor 2: career opportunities	factor 3: practical concerns
(1-4)			0.446
(1-4)			0.418
(1-4)			0.456
(1-4)	0.453		
(1-4)	0.503		
(1-4)	0.808		
(1-4)		-0.747	
(1-4)		-0.753	
% variance initially	28.3%	16.5%	15.2%
Cronbach's alpha	0.723	0.611	0.424

Appendix 3.11: Multilevel analyses student motives for pursuing the DP						country variance (%)	school variance (%)	student variance (%)
constructed variable:			range	mean	stddev			
motive_career	motive career opprotunities		(1-4)	3.40	0.67	1.7%	1.9%	96.5%
motive_personal	motiv epersonal development		(1-4)	2.89	0.71	1.8%	4.7%	92.9%
motive_practical	motive practical concerns		(1-4)	1.86	0.67	4.5%	6.1%	90.2%

		motive career opportunities (1-4)				motive personal development (1-4)				motive practical concerns (1-4)			
Appendix 3.12: Student motives for pursuing the DP: individual differences. Multilevel regression.													
constant		3.368	*0.028*	3.128	*0.088*	2.913	*0.033*	2.450	*0.092*	1.926	*0.039*	1.882	*0.093*
girl	(0-1)			0.048	*0.036*			**0.147**	*0.037*			0.019	*0.035*
academic ability	(0-1)			**0.263**	*0.087*			**0.525**	*0.089*			0.047	*0.086*
language ability	(0-1)			0.004	*0.084*			0.024	*0.087*			-0.116	*0.084*
internal	(0-1)			0.039	*0.073*			0.098	*0.078*			**0.267**	*0.080*
father's education	(0-1)			0.093	*0.068*			0.035	*0.070*			0.073	*0.068*
country variance		0.008	*0.006*	0.006	*0.006*	0.009	*0.008*	0.012	*0.009*	0.021	*0.012*	0.018	*0.013*
school variance		0.009	*0.005*	0.013	*0.007*	0.023	*0.009*	0.016	*0.008*	0.027	*0.009*	0.035	*0.012*
student variance		0.430	*0.014*	0.408	*0.016*	0.462	*0.015*	0.430	*0.017*	0.409	*0.013*	0.393	*0.015*
		0.447		0.427		0.494		0.458		0.457		0.446	
R-square				4.5%				7.3%				2.4%	

CHAPTER 4

Appendix 4.1 Number of DP and other subjects		N	average/ %	range	
a07/b05	follow entire DP (report students)	2,547	93.8%	0-1	
exam registration	follow entire DP diploma students)	3,981	82.0%	0-1	
b23b/c42abc2	full DP according to assignments core components wave 3	1,798	83.2%	0-1	
a10bc/b08bc	follow other (non DP) subjects	2,417	30.8%	0-1	
a10abc/b08abc	follow non DP subjects	2,417	30.8%	0-1	
a10bc/b08bc	follow non DP subjects international programs	2,407	14.6%	0-1	
a10bc/b08bc	follow non DP from subjects from national/regional programs	2,415	26.1%	0-1	
a10a/b08a	nr of DP subjects	2,417	6.0	0-14	1.4
a11a-a18c	nr of DP subjects estimated from selection of subjects	1,904	7.8	0-11	1.4
exam registration	nr of DP subjects estimate wave 3 examfile	3,984	7.7	1-12	2.4
a10abc/b08abc	nr of non DP subjects	2,417	1.5	0-22	3.0
a10bc/b08bc	nr of non DP subjects from international programs	2,417	0.6	0-11	1.7
a10bc/b08bc	nr of non DP subjects from national/regional programs	2,417	0.9	0-13	1.9
a10bc/b08bc	nr of non DP subjects from international programs (among students with these subjects)	354	3.3	0-10	2.3
a10bc/b08bc	nr of non DP subjects from national/regional programs (among students with these subjects)	631	2.7	0-12	2.4

Appendix 4.2: Workload of subjects in terms of difficulty and time-burden, according to students and DP coordinators

subject code	workload according to students	workload according to DP coordi-nators	subject					
8030	-1.32	-0.62	DP core subjects Creativity, activity, service (CAS)					
8010	-1.22	-0.57	DP core subjects Theory of knowledge (TOK)					
8020	-1.11	-0.45	DP core subjects Extended essay (EE)					
6041	-1.06	-0.56	DP group 6 (The arts) Theatre, SL					
2032	-0.97	-0.55	DP group 2 (language acquisition) Language B, HL					
6032	-0.96	-0.28	DP group 6 (The arts) Film, HL					
6031	-0.95	-0.51	DP group 6 (The arts) Film, SL					
2031	-0.92	-0.56	DP group 2 (language acquisition) Language B, SL					
2020	-0.92	-0.53	DP group 2 (language acquisition) Language ab initio, SL					
1011	-0.91	-0.61	DP group 1 (language and literature) Language A literature, SL					
1012	-0.85	-0.22	DP group 1 (language and literature) Language A literature, HL					
6012	-0.77	-0.35	DP group 6 (The arts) Dance, HL					
1021	-0.77	-0.59	DP group 1 (language and literature) Language A language and literature, SL					
3091	-0.75	-0.53	DP group 3 (Individuals and societies) Social and cultural anthropology, SL					
4041	-0.71	-0.57	DP group 4 (Sciences) Design technology, SL					
5011	-0.71	-0.38	Mathematical studies, SL					
3031	-0.57	-0.59	DP group 3 (Individuals and societies) Geography, SL					
3101	-0.54	-0.53	World religions, SL					
3092	-0.53	-0.38	DP group 3 (Individuals and societies) Social and cultural anthropology, HL					
3072	-0.53	-0.03	DP group 3 (Individuals and societies) Philosophy, HL					
4042	-0.51	-0.35	DP group 4 (Sciences) Design technology, HL					
3061	-0.51	-0.56	DP group 3 (Individuals and societies) IT in a global society, SL					
7021	-0.51	-0.55	Environmental systems and society, SL					
1022	-0.50	-0.25	DP group 1 (language and literature) Language A language and literature, HL					
3011	-0.47	-0.55	DP group 3 (Individuals and societies) Business and management, SL					
3062	-0.44	-0.14	DP group 3 (Individuals and societies) IT in a global society, HL					
4061	-0.37	-0.65	DP group 4 (Sciences) Sports, exercise and health science; SL					
3021	-0.22	-0.61	DP group 3 (Individuals and societies) Economics, SL					
6011	-0.22	-0.65	DP group 6 (The arts) Dance, SL					

6052	-0.17	0.89	DP group 6 (The arts) Visual arts, HL
3032	-0.16	-0.39	DP group 3 (Individuals and societies) Geography, HL
6051	-0.16	0.04	DP group 6 (The arts) Visual arts, SL
6022	-0.13	0.92	DP group 6 (The arts) Music, HL
6042	-0.10	-0.03	DP group 6 (The arts) Theatre, HL
2011	-0.08	-0.65	DP group 2 (language acquisition) Classical languages, SL
4031	-0.03	-0.48	DP group 4 (Sciences) Computer science, SL
6021	0.16	-0.39	DP group 6 (The arts) Music, SL
4032	0.27	0.04	DP group 4 (Sciences) Computer science, HL
3052	0.32	0.76	DP group 3 (Individuals and societies) History, HL
4062	0.33	-0.24	DP group 4 (Sciences) Sports, exercise and health science; HL
3071	0.35	-0.56	DP group 3 (Individuals and societies) Philosophy, SL
3081	0.35	-0.48	DP group 3 (Individuals and societies) Psychology, SL
3041	0.38	-0.65	DP group 3 (Individuals and societies) Global Politics, SL
3051	0.45	-0.25	DP group 3 (Individuals and societies) History, SL
3012	0.52	-0.19	DP group 3 (Individuals and societies) Business and management, HL
3022	0.55	-0.32	DP group 3 (Individuals and societies) Economics, HL
5021	0.63	-0.03	DP group 5 (Mathematics) Mathematics, SL
3082	0.77	-0.28	DP group 3 (Individuals and societies) Psychology, HL
4011	0.93	0.16	DP group 4 (Sciences) Biology, SL
4051	0.97	0.17	DP group 4 (Sciences) Physics, SL
4021	0.98	0.25	DP group 4 (Sciences) Chemistry, SL
2012	1.70	3.00	DP group 2 (language acquisition) Classical languages, HL
4052	1.80	2.35	DP group 4 (Sciences) Physics, HL
4012	1.80	1.37	DP group 4 (Sciences) Biology, HL
4022	2.25	1.66	DP group 4 (Sciences) Chemistry, HL
5022	2.99	3.84	DP group 5 (Mathematics) Mathematics, HL
5032	3.16	3.22	DP group 5 (Mathematics) Further mathematics, HL

Appendix 4.3: Difficulty and time-burden by subject, correlations

Ratings according to students:

	most difficult	most home-work	most time exam	easiest	least home-work	least time exam
most difficult	1					
most time homework	0.815	1				
most time exam	0.576	0.516	1			
easiest	-0.379	-0.095	-0.247	1		
least time homework	-0.238	-0.226	-0.224	0.751	1	
least time exam	-0.133	-0.038	-0.478	0.349	0.402	1

Ratings according to DP coordinators:

	most difficult	most home-work	most time exam	easiest	least home-work	least time exam
most difficult	1					
most time homework	0.829	1				
most time exam	0.834	0.837	1			
easiest	-0.302	-0.308	-0.388	1		
least time homework	-0.220	-0.254	-0.349	0.936	1	
least time exam	-0.121	-0.192	-0.262	0.857	0.930	1

Appendix 4.4 Hours per week time spent on school work			
	N	**mean**	*stddev*
Year1 (November)			
regular lessons DP	1,813	26.4	*9.7*
homework DP	1,813	18.3	*10.2*
extra lessons DP	1,813	1.8	*2.4*
regular lessons other programmes	1,813	3.7	*7.4*
homework other programmes	1,813	2.9	*6.1*
extra lessons other programmes	1,813	0.4	*1.1*
total hours DP	1,813	**46.4**	*16.7*
total hours non DP	1,813	6.9	*12.5*
total hours school	1,813	53.3	*20.5*
Year1 (June)			
regular lessons DP	1,529	26.7	*10.1*
homework DP	1,529	18.0	*10.1*
extra lessons DP	1,529	2.0	*2.6*
regular lessons other programmes	1,529	3.8	*7.9*
homework other programmes	1,529	2.8	*6.1*
extra lessons other programmes	1,529	0.4	*1.2*
total hours DP	1,529	**46.8**	*17.1*
total hours non DP	1,529	7.0	*13.5*
total hours school	1,529	53.7	*13.5*

Appendix 4.5 Workload due to the number and selection of subjects, variations with social and personal background						
	workload selection of subjects (N = 1,485)					
	(-50 - 50)					
	b	se		b	se	
(Constant)	-0.87	1.04		1.70	1.14	
girl (0-1)				**-3.57**	*0.81*	*
ability_g general ability before IB (0-1, centered)				**24.43**	*2.01*	*
ability_l language ability before IB (0-1, centered)				**-5.57**	*1.96*	*
ses01 parents' socio-economic status (0-1, centered)				**10.09**	*2.58*	*
internal (0-1)				-0.95	*1.90*	
courseDP (0-1)				-1.33	*2.20*	
country variance	0.00			0.00	*0.00*	
school variance	74.15	*14.33*		69.80	*13.22*	
student variance	241.41	*9.13*		211.99	*8.02*	
country variance	0.0%	36		0.0%	36	
school variance	23.5%	97		24.8%	97	
student variance	76.5%	1755		75.2%	1755	

Appendix 4.6 Time spent on school work (homework, extra lessons), variations with social and personal background						
	November year 1 (N = 1,491)					
	(0-50)					
	b	*se*		**b**	*se*	
(Constant)	19.49	*0.60*	*	17.82	*0.72*	*
girl (0-1)				**2.84**	*0.56*	*
ability_g academic ability before IB (0-1, centered)				**5.88**	*1.43*	*
ability_l language ability before IB (0-1, centered)				**-3.72**	*1.33*	*
ses01 parents' socio-economic status (0-1, centered)				-0.82	*1.74*	
internal (0-1)				0.26	*1.23*	
course DP (0-1)				**-4.48**	*1.47*	*
workload of selection of subjects				**3.73**	*1.73*	*
country variance	2.72	*2.65*		3.25	*2.77*	
school variance	9.17	*2.99*		8.45	*2.83*	
student variance	106.01	*3.99*		101.43	*3.82*	
country variance	2.3%			2.9%		
school variance	7.8%			7.5%		
student variance	89.9%			89.7%		

June year 1 (N = 867)					
(0-50)					
b	*se*		**b**	*se*	
20.21	*0.53*	*	18.68	*0.76*	*
			2.29	*0.77*	*
			5.72	*1.93*	*
			-2.78	*1.74*	
			-1.46	*2.21*	
			-0.42	*1.55*	
			-7.16	*2.32*	*
			8.91	*2.27*	*
0.00	*0.00*		0.27	*1.47*	
10.49	*3.49*		8.81	*3.40*	
107.05	*5.38*		101.54	*5.10*	
0.0%			0.2%		
8.9%			8.0%		
91.1%			91.8%		

CHAPTER 5

	Appendix 5.1: Perception of DP coordinators on how students experience the workload, the difficulty and the stress levels of students of DP and other programmes								
		DP				**other programmes** [a]			
	How do you think students at your school …	**N DP coordinators**	**mean**	**stddev**		**N DP coordinators**	**mean**	**stddev**	
q29ab/q32	Experience the workload of their programme?	1091	0.79	0.14		624	0.57	0.17	[a]
q30ab/q33	Perceive the level of difficulty of their programme?	1097	0.73	0.14		628	0.55	0.15	[b]
q31ab/q34	How stressful do you think students find their programme?	1089	0.74	0.17		622	0.47	0.22	[c]

CHAPTER 6

Appendix 6.1: Subjective workload over the course of DP							
Subjective workload: manageability							

Perception		**N**	**Min**	**Max**	**Mean**	**SD**	**% (much) too heavy**
First year of the DP, how did you perceive the DP workload: during the first three months?	year 1: sept-nov	1,449	0	1	0.541	0.200	14.1%
How do you perceive the DP workload?	year 1: nov	1,764	0	1	0.726	0.157	41.5%
First year of the DP, how did you perceive the DP workload: during the second three months?	year 1: dec-feb	1,431	0	1	0.643	0.157	22.3%
First year of the DP, how did you perceive the DP workload: during the last three months?	year 1: mar-jun	1,422	0	1	0.763	0.171	55.9%
How do you perceive the DP workload?	year 1: june	1,462	0	1	0.730	0.149	43.4%
How did you perceive your DP workload in the second year? - Beginning of school year - November	year 2: sept-nov	1,203	0	1	0.638	0.195	27.2%
How did you perceive your DP workload in the second year? - December - February	year 2: dec-feb	1,199	0	1	0.699	0.197	39.8%
How did you perceive your DP workload in the second year? - March - end of school year	year 2: mar-jun	1,192	0	1	0.669	0.202	34.3%
How did you perceive the DP exam workload? (By exam workload, we mean the tot...-	year 2: exam	1,338	0	1	0.677	0.188	
Overall, how did you perceive the DP workload?	year 2: june	1,489	0	1	0.710	0.162	39.2%
Manageability		**N**	**Min**	**Max**	**Mean**	**SD**	**% (absolutely) unmanageable**
First year of the DP, how maneagable was the DP workload: during the first three months?	year 1: sept-nov	1,443	0	1	0.318	0.244	8.2%
How manageable do you find your DP workload?	year 1: nov	1,761	0	1	0.463	0.216	13.2%
First year of the DP, how maneagable was the DP workload: during the second three months?	year 1: dec-feb	1,428	0	1	0.423	0.220	10.9%
First year of the DP, how maneagable was the DP workload: during the last three months?	year 1: mar-jun	1,418	0	1	0.545	0.234	26.9%
How manageable do you find your DP workload?	year 1: june	1,454	0	1	0.490	0.211	16.1%
Overall, how manageable did you find your DP exam workload?-	year 2: exam	1,440	0	1	0.466	0.225	
Overall, how manageable did you find your DP workload?-	year 2: june	1,490	0	1	0.457	0.217	13.2%
Difficulty		**N**	**Min**	**Max**	**Mean**	**SD**	**% (much) too difficult**
Level of difficulty of the DP?	year 1: nov	1,766	0	1	0.667	0.143	23.8%
Level of difficulty of the DP?	year 1: june	1,520	0	1	0.674	0.148	27.6%
Overall, how did you perceive the level of difficulty of the DP?-	year 2: june	1,497	0	1	0.656	0.149	25.3%
Difficulty		**N**	**Min**	**Max**	**Mean**	**SD**	**% (strongly) agree**
Agreement with following statement: The academic level of the DP exceeds my capacity.	year 1: nov	1,768	0	1	0.447	0.244	13.4%
Agreement with following statement: The academic level of the DP exceeds my capacity.	year 1: june	1,518	0	1	0.442	0.247	14.1%
Please indicate your level of agreement with the following statement: The ac...-	year 2: june	1,494	0	1	0.484	0.236	15.9%

Appendix 6.2: School anxiety: worrying about tests and results											
	Wave 1					Wave 2					
	N	Range	Mean	SD	% most or all of the time	N	Mean	SD	% most or all of the time	t	df
I worry that it will be difficult for me taking a test.	1,593	(0-1)	0.654	0.271	54.7%	1,323	0.596	0.281	47.9% a	**6.1**	844
I worry that I will get poor grades at school.	1,590	(0-1)	0.685	0.294	59.9%	1,321	0.678	0.293	61.5%	1.1	838
Even if I am well prepared for a test I feel very anxious.	1,592	(0-1)	0.685	0.303	61.6%	1,321	0.636	0.311	57.1%	**5.6**	841
I get very tense when I study for a test.	1,588	(0-1)	0.576	0.327	44.7%	1,322	0.566	0.312	44.6%	1.7	837
I get nervous when I don't know how to perform a task at school.	1,585	(0-1)	0.686	0.296	62.0%	1,325	0.652	0.298	56.6%	**5.2**	840
	t-test paired										
I worry that it will be difficult for me taking a test.	0.66	0.60									
I worry that I will get poor grades at school.	0.68	0.67									
Even if I am well prepared for a test I feel very anxious.	0.68	0.63									
I get very tense when I study for a test.	0.57	0.55									
I get nervous when I don't know how to perform a task at school.	0.69	0.64									
original scale:											
1 All of the time											
2 Most of the time											
3 Some of the time											
4 A little of the time											
5 None of the time (wave 1: None of the time / not applicable)											

CHAPTER 7

	aHRSx	bHRSx	cHRSx		aWL	bWL	cWL		aStress	bStress	cStress
Appendix 7.1: Correlations of P-factors with Time investment, Workload and Stress indicators at the three waves											
PERSONAL AND SOCIAL BACKGROUND											
Female	**0.111**	**0.089**	**0.068**		**0.139**	**0.129**	**0.122**		**0.136**	**0.145**	**0.203**
Academic ability	**0.051**	0.048	**-0.112**		**-0.108**	**-0.180**	**-0.168**		**-0.096**	**-0.127**	**-0.137**
Language ability	-0.036	-0.023	**-0.123**		**-0.073**	**-0.130**	**-0.149**		**-0.059**	**-0.072**	**-0.091**
Parent SES	-0.012	0.029	-0.029		-0.043	-0.040	-0.020		**-0.088**	-0.051	**-0.076**
One parent					0.007	0.049	0.044		**0.079**	**0.059**	**0.095**
Internal Student	0.012	0.043	**0.154**		**0.079**	**0.131**	**0.084**		0.031	**0.073**	0.022
OBJECTIVE WORKLOAD											
Subjects	**0.106**	**0.108**	0.034		**-0.094**	-0.066	-0.036		-0.054	**-0.083**	-0.029
PARENTS											
Parents Involvement life	0.036	0.034	0.020		**-0.058**	**-0.095**	**-0.129**		**-0.350**	**-0.273**	**-0.269**
Parents involvement aspirations	**0.080**	0.047	0.030		0.057	0.023	0.020		**0.076**	**0.089**	**0.069**
Parents Involvement schoolwork	0.046	0.072	0.067		0.044	0.012	**-0.075**		**-0.109**	**-0.062**	**-0.070**
PEDAGOGUES											
Teachers Involvement (students)											
Teacher student relations	-0.021	-0.044	-0.061		**-0.181**	**-0.265**	**-0.215**		**-0.283**	**-0.344**	**-0.291**
Teaching quality	-0.059	0.010	0.008		**-0.245**	**-0.232**	**-0.200**		**-0.316**	**-0.268**	**-0.278**
School pressure	**0.143**	**0.141**	0.046		**0.319**	**0.253**	**0.202**		**0.272**	**0.210**	**0.227**
Teachers Involvement (school)											
Teacher student relations	0.013	-0.022	-0.048		**-0.108**	**-0.152**	**-0.101**		**-0.148**	**-0.175**	**-0.144**
Teaching quality	**-0.062**	0.018	0.001		**-0.136**	**-0.076**	**-0.063**		**-0.166**	**-0.119**	**-0.119**
School pressure	**0.108**	**0.117**	0.030		**0.123**	**0.090**	0.048		**0.136**	**0.110**	**0.075**

Appendix 7.2: Time spent on DP, subjective workload, and stress: effects of personal background and relations with parents, pedagogues, peers and school policy. Multilevel standardized regression coefficients, corrected for personal and social background

	time spent on DP	subjective workload	stress
TIME			
time: increase end yr2 vs end yr1	-0.014	-0.023	**0.235**
PERSONAL AND SOCIAL BACKGROUND			
Female	**0.053**	**0.141**	**0.184**
Academic ability	**0.087**	**-0.108**	**-0.138**
Language ability	-0.037	**-0.072**	-0.028
Parent SES	-0.003	-0.024	-0.047
Internal student [a]	0.003	0.041	0.030
OBJECTIVE WORKLOAD			
subjects	**0.079**	0.003	0.004
PARENTS			
Parents involvement life	0.008	**-0.096**	**-0.303**
Parents' involvement aspirations	0.043	0.038	**0.089**
Parents' Involvement schoolwork	0.025	-0.014	**-0.086**
PEDAGOGUES			
Teacher involvement (students)			
Teacher-student relations	-0.036	**-0.223**	**-0.331**
Teaching quality	-0.037	**-0.202**	**-0.280**
School pressure	**0.067**	**0.235**	**0.229**
Teacher involvement (school)			
Teacher student relations	-0.006	**-0.124**	**-0.166**
Teaching quality	0.045	**-0.082**	**-0.116**
School pressure	0.023	**0.074**	**0.121**

PEERS			
Peer support (students)	**-0.057**	**-0.182**	**-0.386**
Peer support (school)	-0.050	**-0.129**	**-0.172**
POLICIES			
Entrance requirements	**0.116**	0.043	-0.044
Number of assessments	0.042	0.055	0.048
Spreading of IAs	-0.021	-0.047	-0.034
Spreading of Core components	0.012	0.010	-0.001
Planning (coordinators)	**0.087**	0.037	-0.048
Support measures	0.062	-0.011	-0.032
Planning (students)	0.038	**0.076**	**0.106**
Support professionals	0.007	0.033	-0.053
Programme mental health	0.040	-0.021	-0.094
Programme learning skills	0.016	-0.016	-0.023
Programme relaxation	-0.009	0.009	-0.002
[a] controlled for parameter for missing values			

Appendix 7.3: Parents' involvement - descriptives

			N	range	mean	stddev
a80a	pschool01	Parents involved with school	1,526	0-1	0.56	0.29
a80b	pteachers	Parents involved with teachers	1,522	0-1	0.38	0.29
a80c	ppersonal	Parents involved with personal life	1,525	0-1	0.54	0.28
a80d	plife01	Parents involved with life outside school	1,524	0-1	0.58	0.28
a80e	pfriends0	Parents involved with friends	1,521	0-1	0.42	0.28
a81a/b63a	phappy01	Parents want me to be happy	1,793	0-1	0.90	0.17
a81b/b63b	pwork01	Parents expect me to work hard at school	1,787	0-1	0.89	0.16
a81c/b63c	penjoy01	Parents find it important that I enjoy school	1,789	0-1	0.77	0.24
a81d/b63c	pgrades01	Parents expect me to get high grades	1,521	0-1	0.83	0.20
a82a/b64a	pcheck01	Parents check your homework after it was completed?	1,795	0-1	0.09	0.21
a82b/b64b	phw01	Parents help you do your homework?	1,793	0-1	0.12	0.22
a82c/b64c	ptests01	Parents help you prepare for tests?	1,778	0-1	0.11	0.23
a83d/b64c	ptalk01	Parents talk with you about your experience at school (classes)	1,784	0-1	0.58	0.36
b64e	ptalk201	Parents talk with you about your experience at school (lunch, sports	268	0-1	0.60	0.36
a83a/b65a	pcheck201	Parents check your grades online?	1,788	0-1	0.27	0.33 [a]
a83b/b65b	pcontact0	Parents contact (one of your) teachers?	1,782	0-1	0.12	0.21 [a]

[a] answering categories different between waves 1 and 2

Appendix 7.4: Parents' involvement - factor analysis - pricipal axing factoring - oblimin rotation

	parents involved with ...			
	life	aspirations	school	
Parents involved with personal life	0.718			
Parents involved with life outside school	0.712			
Parents want me to be happy	0.605			
Parents find it important that I enjoy school	0.592			
Parents involved with friends	0.577			
Parents talk with you about your experience at school (classe	0.391			
Parents expect me to work hard at school		0.893		
Parents expect me to get high grades		0.704		
Parents check your homework after it was completed?			0.594	
Parents help you do your homework?			0.582	
Parents help you prepare for tests?			0.573	
Parents involved with teachers			0.529	
Parents involved with school	0.352		0.483	
Parents contact (one of your) teachers?			0.401	
Parents check your grades online?			0.383	
Initial % variance	28.6%	12.7%	10.9%	
Cronbach's alpha	0.77	0.77	0.73	
Cronbach's alpha overall	80.0%			
Factor Correlation Matrix				
Factor	1	2	3	
1	1			
2	0.131	1		
3	0.303	0.017	1	
a for scale construction considered as indicator of factor 3				

Appendix 7.5: Pedagogues and peers: Teachers' involvement and peer relations - descriptives

				N	Range	Mean	
sf_subjects01	a53a	How do you feel about your DP subjects	a	1,638	0-1	0.72	0.19
sf_teachers01	a53b	How do you feel about your DP teachers	a	1,636	0-1	0.72	0.20
sf_peers01	a53c	How do you feel about your classmates in the DP	a	1,634	0-1	0.76	0.21
teacher_questions01	a54a	In my DP classes I am encouraged to ask questions	b	1,637	0-1	0.76	0.19
teacher_help01	a54b	It is hard to get help in my DP classes when I have a question (reversely code	b	1,634	0-1	0.68	0.24
teacherfeedback01	a54c	DP teachers provide timely feedback	b	1,621	0-1	0.66	0.21
DP-learning01	a54d	The DP results in only modest learning (reversely coded)	b	1,603	0-1	0.57	0.25
peers_care01	a55a	I feel that DP students care about each other (reversely coded)	c	1,637	0-1	0.67	0.23
peers_isol01	a55b	I feel isolated in the DP	c	1,634	0-1	0.72	0.26
peers_connect01	a55c	I feel connected to other students in the DP (reversely coded)	c	1,630	0-1	0.68	0.25
peers_lackcomm01	a55d	I lack a spirit of community in the DP	c	1,634	0-1	0.70	0.28
pressure_excel_a01	a56a	There is much pressure in the DP to excel.	b	1,593	0-1	0.77	0.23
pressure_demands_a01	a56b	DP teachers place high demands on their students.	b	1,590	0-1	0.75	0.21
pressure_competition_a01	a56c	There is much competition among DP students.	b	1,590	0-1	0.67	0.27
pressure_excel_b01	b47a	There is much pressure in the DP to excel.		1,304	0-1	0.74	0.23
pressure_demands_b01	b47b	DP teachers place high demands on their students.	b	1,300	0-1	0.74	0.22
pressure_competition_b01	b47c	There is much competition among DP students.	b	1,294	0-1	0.64	0.28

sc_hwcomplete01	a57a	I complete my homework on time (reversely coded)	c	1,601	0-1	0.79	0.20
sc_hwquiet01	a57b	I do my homework in a quiet space (reversely coded)	c	1,602	0-1	0.72	0.23
teacher_grade01	a57c	My DP teachers grade my homework on time (reversely coded)	c	1,598	0-1	0.64	0.28
teacher_comments01	a57d	My DP teachers make useful comments on my homework (reversely coded)	c	1,596	0-1	0.54	0.28
teacher_along01	b46a	Students get along well with most teachers.	b	1,299	0-1	0.72	0.21
teacher_interest01	b46b	Most teachers are interested in students' well-being.	b	1,303	0-1	0.70	0.23
teacher_listen01	b46c	Most of my teachers really listen to what I have to say.	b	1,299	0-1	0.69	0.23
teacher_extra01	b46d	If I need extra help, I will receive it from my teachers.	b	1,302	0-1	0.73	0.21
teacher_fairly01	b46e	Most of my teachers treat me fairly.	b	1,301	0-1	0.75	0.21
		How concerned is your school with:					
schoolconcerned_ac01	b48a	academic performance	d	1303	0-1	0.66	0.27
schoolconcerned_learning01	b48b	learning needs	d	1301	0-1	0.56	0.28
schoolconcerned_wellbeing01	b48c	wellbeing	d	1301	0-1	0.51	0.30
schoolconcerned_personal01	b48d	personal life	d	1305	0-1	0.35	0.28
constructed scales							
t_relations		quality of student-teacher relations		1305	-.63 - .	0.00	0.18
t_teaching		teaching quality		1642	-.59 - .	0.00	0.15
t_pressure		school pressure		2064	-.72 - .	0.00	0.19

a original scale 1-6: completely dissatisfied - completely satisfied
b original scale 1-6: strongly disagree - strongly agree
c original scale 1-5: all of the time - none of the time
d original scale 1-5: not at all concerned - extremely concerned

Appendix 7.6: Factor analysis questions on teachers and school climate in three dimensions

Student-teacher relationships			
teacher_interest01	Most teachers are interested in students' w	0.842	
teacher_listen01	Most of my teachers really listen to what I h	0.827	
teacher_extra01	If I need extra help, I will receive it from my	0.741	
teacher_fairly01	Most of my teachers treat me fairly.	0.739	
teacher_along01	Students get along well with most teachers.	0.674	
schoolconcerned_wellbeing01	school concerned with wellbeing	0.610	
schoolconcerned_personal01	school concerned with personal life	0.508	
schoolconcerned_learning01	school concerned with learning needs	0.470	
Cronbach's alpha		*0.86*	
Average, stddev		0.63	*0.18*
School average, stddev		0.62	*0.07*
Teaching quality			
teachers_feedback01	DP teachers provide timely feedback	0.654	
sf_teachers01	How do you feel about your DP teachers	0.601	
teacher_comments01	My DP teachers make useful comments on r	0.599	
teacher_questions01	In my DP classes I am encouraged to ask que	0.513	
teachers_help01	It is hard to get help in my DP classes when I	0.451	
teacher_grade01	My DP teachers grade my homework on time	0.401	
schoolconcerned_ac01	school concerned with academic performan	0.188	
Cronbach's alpha		*0.65*	
Cronbach's alpha after removing school concerned with academic perfoman		*0.70*	
Average, stddev		0.67	*0.15*
School average, stddev		0.67	*0.07*
School pressure			
pressure_excel_b01	There is much pressure in the DP to excel.	0.745	
pressure_excel_a01	There is much pressure in the DP to excel.	0.655	
pressure_competition_a01	There is much competition among DP stude	0.614	
pressure_demands_b01	DP teachers place high demands on their stu	0.613	
pressure_competition_b01	There is much competition among DP stude	0.605	
pressure_demands_a01	DP teachers place high demands on their stu	0.562	
Cronbach's alpha		*0.80*	
Average, stddev		0.72	*0.18*
School average, stddev		0.72	*0.09*

Appendix 7.7: Schools' entrance requirements: factors considered when students are admitted to the DI

			N DP coordinators	average - never - sometimes - always (0-1)	Std. Dev.	% always		
	q16a	Ability tes	78	0.24	*0.38*	15.4%		
	q16b	Interview	80	0.74	*0.37*	63.8%		
	q16c	Language	78	0.48	*0.43*	34.6%		
	q16d	Recomme	78	0.47	*0.40*	29.4%		
	q16e	Student's	80	0.80	*0.30*	66.3%		

Appendix 7.8: Number of assessments and core exam components handed in, estimates by DP students

		yr 1	yr2: Sept-Nov	yr2: Dec-Feb	yr2: Mar-June	yr 2	
nr assessments	mean	**1.94**	1.65	2.57	1.57	**5.67**	
	range	0-4.3	0-6	0-6	0-6	0-6	
	stddev	*0.83*	*1.11*	*0.98*	*1.15*	*1.64*	
	low school	0.3	-0.6	0.6	-0.7	2.4	
	high school	3.6	3.9	4.5	3.9	9.0	
	N (schools)	98	96	96	96	96	
nr core components	mean	**0.94**	1.02	1.34	0.72	**3.01**	
	range	0-2.1	0-3	0-3	0-3	0-3	
	stddev	*0.54*	*0.58*	*0.55*	*0.60*	*0.92*	
	low school	-0.1	-0.1	0.2	-0.5	1.2	
	high school	2.0	2.2	2.4	1.9	4.8	
	N (schools)	96	95	96	94	93	

		N	mean (0-1) recoded items	stddev (0-1) recoded items		original categories	
Appendix 7.9: School policy on limitations of homework, and spreading of internal assessments and core components, according to DP coordinators - school sample							
	To what extent are the following dates coordinated at your school:				**% always**		
Q47a_imp	Deadlines for IAs	74	0.87	0.20	64.9%	1-5	a
Q47b_imp	Deadlines for Core Components	73	0.91	0.15	71.2%	1-5	a
Q47c_imp	School-based assessments	74	0.69	0.28	32.4%	1-5	a
	How often do DP teachers hold collaborative planning meetings:				**% once a week or more often**		
Q48a_imp	Across all DP subject groups	74	0.43	0.19	5.4%	1-6	b
Q48b_imp	Within a DP subject group	74	0.59	0.21	35.1%	1-6	b
Q48c_imp	Across the two DP year groups	74	0.45	0.24	14.9%	1-6	b
	Implemented at your school:				**% implemented**		
Q58a_imp	Homework free days	75	0.16	0.37	16.0%	1-2	c
Q58b_imp	Homework free periods	75	0.23	0.42	22.7%	1-2	c
Q58c_imp	Maximum amount of homework per course	75	0.28	0.45	28.0%	1-2	c
Q58d_imp	Maximum amount of homework per week	74	0.23	0.42	23.0%	1-2	c
	Cronbach's alpha		*0.65*				
[a] never - always							
[b] never - more than once a week							
[c] yes-no							

Appendix 7.10: Schools' practices to practices to spread students' workload				N schools	% provided at my school - student school average
Designated study period blocks in the timetable				94	34.4%
A clear schedule of (non-examination) assessment deadlines (ie, IAs and (				94	43.1%
Homework-free days				94	14.0%
Homework-free periods				94	13.9%
Maximum amount of homework per course				94	6.2%
Maximum amount of homework per week				94	7.6%
Reading periods (ie, scheduled blocks without instructions during which s				94	15.2%
c48a1-c48g6					

Appendix 7.11: Schools' practices to spread students' workload		N schools	% provided at my school - student school average	N DP-coordinators student sample	% provided - DP coordinators
c44a/q42a	Finalizing all IAs by November of year two	96	19.0%	74	14.9%
c44b/q42b	Finalizing part of the IAs in year one	96	28.5%	74	39.2%
c44c/q42c	Finalizing all other non-examination assessments (	96	18.7%	74	20.3%
c44d/q42d	Finalizing part of the other other non-examination	96	20.6%	74	24.3%

Appendix 7.12: Effectiveness schoolplanning according to students							
				N	**range**	**mean**	**stddev**	**% two lowest categories (not effective)**
How effective was your school's planning … [a]								
	Deadlines for internal assessments (Ias)			1,150	0-1	0.57	0.29	41.7%
	Deadlines for Core exam components			1,149	0-1	0.51	0.29	33.9%
	Other assessments deadlines			1,149	0-1	0.53	0.28	33.9%
	Homework assignments			1,148	0-1	0.55	0.28	37.2%
	Cronbach's alpha .848							
					country variance	**school variance**	**individual student variance**	
	Average non-effectiveness				3.3%	11.9%	85.1%	

[a] reversely coded: 1-5: extremely effective - not at all effective

variables: C47a-c47d

Appendix 7.13: Schools' support services according to DP students		N schools	% provided at my school - student school average
Support services students' needs beyond the classroom			
c49a	Teacher-led academic coaching groups	94	66.1%
c49b	Student-led academic coaching groups (eg, peer tutor groups)	94	42.3%
c49c	Externally-led academic coaching groups	94	17.6%
c49d	Teacher-led subject specific individual tutoring	94	56.6%
c49e	Student-led subject specific individual tutoring	94	38.1%
c49f	Externally-led subject specific individual tutoring	94	20.3%
c49g	Study room or hall (for students to do their homework under supervision)	94	61.2%
c49h	Individual needs support (eg, dyslexia, Attention Deficit Hyperactivity Disorder)	94	65.5%
c49i	Summer school	94	41.7%
c49j	Advancement via individual determination (AVID) for low-income students	94	29.0%
c49k	Pre-DP track	94	38.3%
c49l	University counselling	94	77.4%
Support services health and wellbeing			
c50a	School counsellor	94	89.1%
c50b	Social worker	94	46.9%
c50c	School doctor	94	35.0%
c50d	School nurse	94	74.0%
c50e	School psychologist	94	44.7%
c50f	Student support groups	94	35.1%
c50g	Peer support network	94	41.2%

		N DP coordinators	% provided
Appendix 7.14: Schools' support programmes: provision and student participation, estimated by DP coordina			
q61a	Mental health programme (eg, Social and Emotional Learning [SEL])	68	48.5%
q61b	Learning skills programme (eg, time management, organization, communicat	68	61.8%
q61c	Relaxation programme (eg, Mindfulness, Yoga)	68	52.9%

Appendix 7.15: Number of assessments over the course of year 2 and the effects on time spent on DP, subjective workload, and stress (retrospective information). Multilevel standardized regression coefficients, corrected for personal and social background

	time spent on DP	subjective workload	stress
TIME			
yr 2: Sept-Nov (constant)			
yr 2: Dec-Feb (deviation from Sept-No	**0.054**	**0.174**	**0.421**
yr 2: Mrch-June (deviation from Sept-N	0.028	**0.124**	**0.599**
PERSONAL AND SOCIAL BACKGROUND			
Female	**0.083**	**0.121**	**0.233**
Academic ability	0.031	0.034	0.030
Language ability	-0.041	**-0.073**	-0.027
Parent SES	-0.022	-0.016	-0.019
Internal student [a]	0.045	-0.001	-0.014
internal missing	-0.034	0.059	-0.056
nr assessments	**0.104**	**0.216**	**0.203**
[a] controlled for parameter for missing values			

Appendix 7.16: Test of Moderators by SEM group comparisons

	Groups	Model A		Model B		Model C	
		All effects the same		All effects different		Selected effects different	
	Groups	L2	df	L2	df	L2	df
Gender	2	123.4	24	101.8	20	102.8	22
Academic Ability	3	123.9	38	102.9	30		
Parents SES	3	130.2	38	116.5	30		
Family Situation	3	122.4	38	118.9	30		
Parental Involvement Life	3	118.2	38	108.5	30		
Parental Involvement Aspirations	3	130.3	38	119.1	30		
Parental Involvement Schoolwork	3	126.2	38	122.3	30		
Number of Assessments	3	127.6	38	120.4	30		
Spread of Assessments	3	135.0	38	129.0	30		
Spread of Core	3	125.0	38	119.2	30		
Planning	3	102.0	38	94.9	30		
Ineffective Planning	3	116.5	38	109.6	30		
Provisions	3	125.2	38	120.2	30		
Note 1: Results of model A differ between moderators because of differen N							
Note 2: Estimation of Model C is unnecessary when difference between A and B is not significant							
Note 3: Critical values (p<.05, two taled): L2 = 9.5 / df=4 and L2 = 15.5 / df=8							

CHAPTER 9

Appendix 9.1: Schools' entrance requirements: factors considered when students are admitted to the DP					
	N DP-coor-dinators	**never - sometimes - always (0-1)**	**Std. Dev.**	**% always**	
Ability test (eg, IQ-test)	1,309	0.32	*0.41*	22.1%	
Interview with the student	1,347	0.69	*0.38*	55.9%	
Language test	1,323	0.49	*0.42*	35.4%	
Recommendation of feeder schools	1,312	0.51	*0.40*	32.8%	
Student's record of academic performance	1,352	0.79	*0.34*	69.1%	

Appendix 9.2: DP coordinators' opinions: options to prevent students from running into problems with the DP workload				
	N DP- coor-dinators	**average not at all extremely helpful (0-1)**	**stddev**	**% would be (very) helpful**
Training in organizational skills	1,216	0.72	*0.24*	68.0%
Guidance for aspiring DP students	1,217	0.69	*0.24*	65.2%
Advising DP students on their subject choice	1,214	0.69	*0.25*	64.9%

Appendix 9.3 Number of assessments and core exam components handed in, estimates by DP coordinators						
DP Coordinators broader sample						
	yr1: Sept-Nov	yr1: Dec-Feb	yr1:Mar-Jun	yr2: Sept-Nov	yr2: Dec-Feb	Yr2: Mar-June
nr assessments in this period:	**0.50**	**1.15**	**2.09**	**2.79**	**3.36**	**1.70**
monthly average:	0.17	0.38	0.52	0.93	1.12	0.43
	0-6	0-6	0-6	0-6	0-6	0-6
	1.00	*1.24*	*1.14*	*1.51*	*1.68*	*1.83*
nr core components in this period:	**0.34**	**0.63**	**1.14**	**1.51**	**1.68**	**0.90**
monthly average:	0.11	0.21	0.29	0.50	0.56	0.23
	0-3	0-3	0-3	0-3	0-3	0-3
	0.69	*0.85*	*0.91*	*0.91*	*0.88*	*1.07*

Appendix 9.4: DP coordinators' practices: Coordination and planning to support a balanced workload for DP students				
Coordination of deadlines ….	**N DP coordinators**	**average never (0) - always (1)**	**stddev**	**% always**
… for IAs	1,193	0.90	*0.18*	72.1%
… of deadlines for Core Components	1,184	0.92	*0.15*	76.0%
… of school-based assessments	1,185	0.70	*0.29*	38.1%
Collaborative planning meetings …	**N DP coordinators**	**average never (0) - more than once a week (1)**	**stddev**	**% once a week /more often**
… across all DP subject groups	1,192	0.48	*0.22*	15.2%
… within a DP subject group	1,177	0.58	*0.23*	34.6%
… across the two DP year groups	1,174	0.49	*0.23*	18.8%
Homework limitations	**N DP coordinators**			**% implemented**
Homework free days	1,161			19.7%
Homework free periods	1,158			25.1%
Maximum amount of homework per course	1,162			29.0%
Maximum amount of homework per week	1,149			27.2%

Appendix 9.5: Students' opinions: Options to make the workload more manageable for students				
	N students	**average not at all(0)- very(1) helpful**	**SD**	**is/would be very helpful**
Designated study period blocks in the timetable	957	0.60	*0.31*	50.8%
A clear schedule of (non-examination) assessment deadlines (ie, IAs and Core exam components)	920	0.71	*0.27*	67.5%
Homework-free days	1,076	0.67	*0.32*	62.7%
Homework-free periods	1,075	0.64	*0.34*	57.9%
Maximum amount of homework per course	1,092	0.60	*0.35*	51.7%
Maximum amount of homework per week	1,084	0.61	*0.35*	54.2%
Reading periods (ie, scheduled blocks without instructions during which students prepare for assessments)	1,059	0.62	*0.34*	53.5%

Appendix 9.6: DP coordinators' practices and opinions: Spreading students' workload to make the non-examination assessment workload more manageable for students						
DP coordinators - broader sample (N schools = 1182)						
	N DP- coor- dinators	**% practiced**	**N DP- coor- dinators**	**average not at all(0)- extremel y(1) helpful**	**SD**	**% would be very helpful**
Finalizing all IAs by November of year two	1,182	12.2%	1051	0.36	*0.31*	22.3%
Finalizing part of the IAs in year one	1,182	35.4%	950	0.52	*0.30*	40.0%
Finalizing all other non-examination components (eg, written assignment) by November of year 2	1,182	17.3%	1010	0.44	*0.31*	31.3%
Finalizing part of the other other non-examination components (eg, written assignment) in year one	1,182	23.2%	984	0.47	*0.30*	33.4%
Do more IAs as collaborative projects.	1,182	2.5%	1116	0.41	*0.31*	26.9%
Students of sampled schools (N schools = 96)						
	N students	**% practiced**	**N students**	**average not at all(0)- very(1) helpful**	**SD**	**% would be very helpful**
Finalizing all IAs by November of year two	1,158	18.2%	1032	0.44	*0.33*	31.8%
Finalizing part of the IAs in year one	1,158	29.3%	1011	0.60	*0.32*	53.2%
Finalizing all other non-examination assessments (eg, Core exam components) by November of year two	1,158	18.4%	1052	0.50	*0.33*	39.6%
Finalizing part of the other other non-examination assessments (eg, Core exam components) in year one	1,158	20.4%	1053	0.53	*0.34*	41.9%

Appendix 9.7: DP coordinators' and students' opinions: Options to make DP exam workload more manageable

DP coordinators - broader sample (N schools = 1,870)

	N DP- coor-dinators	average not at all(0)-extremely(1) helpful	SD	% would be very helpful
Coordinators				
Extending exam papers over 6 weeks instead of 3 weeks	1,188	0.33	*0.34*	23.7%
Replacing multiple exam papers per subject with one longer	1,181	0.26	*0.30*	14.2%
Moving all SL exam papers to November in year two	1,174	0.23	*0.30*	14.2%
Moving part of the SL exam papers to November in year two	1,168	0.29	*0.32*	19.6%
Moving all SL exam papers to May in year one	1,161	0.20	*0.30*	12.4%
Moving part of the SL exam papers to May in year one	1,171	0.28	*0.32*	18.1%

Students of sampled schools (N schools = 96)

	N DP- coor-dinators	average not at all(0)-extremely(1) helpful	SD	% would be very helpful
Students				
Extending exam papers over six weeks (instead of the currer	1,342	0.65	*0.31*	61.1%
Replacing multiple exam papers per subject with one longer	1,343	0.33	*0.28*	17.0%
Moving all SL exam papers to the November session in year t	1,335	0.44	*0.33*	32.2%
Moving part of the SL exam papers to the November session	1,332	0.47	*0.33*	37.0%
Moving all SL exam papers to the May session in year one	1,333	0.40	*0.32*	28.3%
Moving part of the SL exam papers to the May session in yea	1,320	0.44	*0.32*	31.5%

Appendix 9.8: Schools' support services according to DP coordinators, and, if present, used by students			
Support services students' needs beyond the classroom	N DP coordinators	% present	N students at schools with support service present
Teacher-led academic coaching groups	1,168	80.0%	739
Student-led academic coaching groups (eg, peer tutor groups)	1,168	49.3%	560
Externally-led academic coaching groups	1,168	16.8%	236
Teacher-led subject specific individual tutoring	1,168	78.9%	632
Student-led subject specific individual tutoring	1,168	41.7%	469
Externally-led subject specific individual tutoring	1,168	20.8%	268
Study room or hall (for students to do their homework under supervision)	1,168	62.7%	656
Individual needs support (eg, dyslexia, Attention Deficit Hyperactivity Disorder)	1,168	73.7%	760
Summer school	1,168	27.1%	411
Advancement via individual determination (AVID) for low-income students	1,168	12.2%	286
Pre-DP track	1,168	37.7%	498
University counselling	1,168	87.4%	892
Support services wellbeing and health			
Internal student welfare coordinator/officer	1,162	55.6%	
Wellbeing team, including students	1,158	25.7%	
Wellbeing team, excluding students	1,152	45.0%	
School counsellor	1,168	85.2%	1042
Social worker	1,168	31.2%	579
School doctor	1,168	20.4%	406
School nurse	1,168	67.2%	866
School psychologist	1,168	59.3%	512
Student support groups	1,168	31.8%	467
Peer support network	1,168	30.8%	517

Appendix 9.9: Schools' support programmes: provision and student participation, estimated by DP coordinators			
	N DP coordinators	% provided	% student participation, if provided
Mental health programme (eg, Social and Emotional Learning [SEL])	1,043	49.7%	34.7%
Learning skills programme (eg, time management, organization, communication)	1,040	61.6%	35.6%
Relaxation programme (eg, Mindfulness, Yoga)	1,041	53.8%	28.4%

Appendix 9.10: DP coordinators' and studens' opinions: Options to make the workload more manageable for students

DP coordinators - broader sample (N schools = 1,870)

	N DP coordinators	average much heavier (0)- much lighter (1)	SD	% would make the workload (much) lighter
Replacing an exam paper with an additional IA	1,209	0.37	0.28	20.2%
Replacing an exam paper with non-examination assessment component	1,200	0.48	0.27	32.7%
Replacing an IA with an additional exam paper to take in the exam period	1,205	0.51	0.26	38.6%
Replacing non-examination assessment component with exam paper	1,198	0.49	0.25	33.2%

Students of sampled schools (N schools = 96)

	N students	average not at all(0)- very(1) helpful	SD	% would be very helpful
Replacing an IA with an additional exam paper (to take in the exam period)	1,168	0.19	0.28	9.6%
Replacing a Core exam component with an additional exam paper (to take in the exam period)	1,161	0.16	0.26	7.2%
Replacing an exam paper with an additional IA	1,335	0.44	0.33	34.0%
Replacing an exam paper with another non-examination assessment component (eg, written assignment)	1,335	0.55	0.31	48.7%

Appendix 9.11: DP coordinators' and students' opinions: options to make DP workload more manageable

DP coordinators - broader sample (N schools = 1,870)

	N DP coordinators	average not at all(0)-extremely(1) helpful	SD	% would be very helpful
Clearer differentiation between SL and HL requirements.	1,216	0.57	*0.28*	46.1%
Clearer links between topics studied in different subjects.	1,212	0.58	*0.27*	49.3%
Providing IB-approved subject resources for students (eg books).	1,217	0.69	*0.27*	64.2%
Providing additional IB-approved online resources.	1,214	0.69	*0.27*	65.6%
Reducing the amount of content in DP subjects.	1,211	0.66	*0.30*	58.6%
Making the EE part of another subject.	1,213	0.56	*0.37*	50.1%
Students can use the same data for the IA and the EE for the same subject.	1,210	0.58	*0.34*	50.0%
Limiting IAs to HL subjects only.	1,211	0.60	*0.35*	55.4%
Not grading reflective statements across the programme (eg for CAS).	1,205	0.52	*0.34*	40.8%

Students of sampled schools (N schools = 96)

	N students	average not at all(0)-extremely(1) helpful	SD	% would be very helpful
Having IB-approved subject resources for students (eg, textbooks, study guides)	1,169	0.77	*0.27*	76.1%
Having additional IB-approved online resources	1,168	0.80	*0.26*	79.5%
Having IAs for HL subjects only	1,167	0.58	*0.34*	49.8%
Doing more IAs as collaborative projects	1,164	0.39	*0.35*	29.4%
Having a clearer differentiation between Standard Level (SL) and Higher Level (HL) requirements	1,487	0.50	0.30	36.3%
Having clearer links between content studied in different subjects	1,477	0.50	0.30	36.5%
Having the Extended Essay (EE) as part of one of the other DP subjects	1,476	0.42	0.34	29.7%
Being allowed to use the same data for the Internal Assessment (IA) and the EE for the same DP subject	1,483	0.58	0.33	49.1%
Having reflective statements not graded	1,477	0.55	0.35	46.1%
Having less content to study per DP subject	1,483	0.65	0.32	57.9%

Appendix B: DATA DOCUMENTATION

Contents

Appendix B: DATA DOCUMENTATION

1. Study description

Study title:	DIPLOMA PROGRAMME STUDENT WORKLOAD STUDY
Fieldwork dates:	December 2016 - August 2018
PI:	Heike Schröder (Principal Investigator / IB research manager)
	Harry BG Ganzeboom (Co-PI and data administrator)
	Ineke Nagel (Co-PI)
Funding:	IB-International, Geneva
Sampling frame:	Selection of 10% of IB-DP schools from 50 countries:

* With May exam schedules
* From English, Spanish or French language DP schools
* Stratified random sample of 258 DP schools in 36 countries
* N=4855 DP Year 1 students identified by DP school coordinators

Sample type:	Three waves of data collection among the DP students:

* Wave 1: after the start of the Diploma Programme, December 2016
* Wave 2: after the first year of IB, June 2017
* Wave 3: after the end of the second year of, May 2018 and Exam data

One wave of data collection among DP Coordinators and Heads of school, November 2017

Sample size:	Student survey: N = 4,855
	School survey: N= 1,393 schools (including the 258 sampled schools)
Partial Response:	Student surveys:

* Wave 1: N=2,458, Wave 2: N=2,416, Wave 3: N=2,041
* Any response: N=3,565, All three waves N=1,063

School survey: N=1,393 schools,

Valid Response:	Student surveys (valid on subjective workload score):

* Wave 1: N=1,774
* Wave 2: N=1,527
* Wave 3: N=1,499
* Any wave: N=2,699; all three waves N=629

School survey: N=1,393 schools

N of observations:	Participants with observations N = 3,565, valid N= 2,699
Fieldwork agency:	IB-Research, Bethesda MD, USA
Fieldwork method:	Online questionnaires
Languages:	English, French, Spanish
Weights:	Not applicable

The complete data file (described in this documentation), will be archived at DANS [Data Archiving and Networked Services] in The Hague, Netherlands. The appropriate bibliographic reference to this file will be:

Schröder, H., Ganzeboom, H.B.G., & Nagel, I. (2020). Student Workload and Well-being Study in the International Baccalaureate 2016-2018 [machine-readable data file]. The Hague: DANS. To be archived.

2. General Information

The International Baccalaureate (IB)

- The **International Baccalaureate (IB)** is "a non-profit educational foundation, motivated by its mission to develop inquiring, knowledgeable and caring young people who help create a better and more peaceful world through its programmes of education that promotes intercultural understanding and respect. Founded in 1968, the IB currently works with more than **5,100 schools in 151 countries** to develop and offer four programmes to over one million students aged 3 to 19 years. The **Diploma Programme (DP)** is an academically challenging and balanced programme for students aged 16-19. It aims to develop students who:
  - have excellent breadth and depth of knowledge
  - flourish physically, intellectually, emotionally and ethically
  - study at least two languages
  - excel in traditional academic subjects
  - explore the nature of knowledge through the programme's unique theory of knowledge course."

- The Diploma Programme (DP) is composed of **three core elements and six subject groups**. The three DP core elements are theory of knowledge (TOK), the extended essay (EE), and creativity, activity, service (CAS). The six subject groups are: studies in language and literature, language acquisition, individuals and societies, sciences, mathematics and the arts. Subjects can be taken at High Level (HL) or Standard Level (SL).

The Student Workload and Well-being Study - 2016-2018

- This document describes the **Student Workload and Well-being Study 2016-2018**, conducted by IB Research in collaboration with researchers from the Department of Sociology of VU University Amsterdam. The study was designed as a three-wave panel survey among 4,855 students of the IB Diploma Programme in 36 countries, in order to examine the development of workload-induced stress among students in this rigorous pre-university international education. The study was a joint effort of IB Research in The Hague (Heike Schröder) and the Department of Sociology of VU University Amsterdam (Harry Ganzeboom and Ineke Nagel).

- The **main research questions** of the study were:
  - **How manageable is the DP workload and how is it related to DP students' experience of stress?**
  - **What are potential factors that can impact students' experiences of workload and stress in the DP?**

- More specifically, the study addresses the following **research questions** to elaborate on these two main questions:
  - Which **specific elements in the DP** (eg, type or combination of subjects), if any, **increase** students' **subjective workload**?
  - Is there a **causal relationship between workload and stress**, or can stress be explained by pre-existing conditions?
  - Which **risk and support factors** affect student's workload and well-being?
  - Which **factors help students cope** with their subjective workload?
  - To what extent do **workload and stress vary across students, schools and the course of the programme**?
  - How are the DP **workload and stress related to student outcomes** (eg, exam results, university admission)?
  - What do schools do to **help students cope** with potential stress? (eg, policies, support services) and how effective are these measures?

3. Questionnaire development and post-processing

Timing

- The research design was a **three-wave panel survey** of DP students supplemented with their **exam results** and a one-time **school survey**. Students were followed during their entire two-year DP career, with measurement taking place at **three** time points ('**waves**'). Three **online questionnaires** were sent to the students:
  - December 2016 (at the beginning of Year 1 of de DP)
  - June 2017 (at the end of Year 1 of the DP)
  - May 2018 (after completion of the exams)
- In addition, DP Coordinators and Heads of school of all 258 DP schools in the effective sample of students were sent an online **school survey** in the beginning of Year 2 (October 2017).
- The **results of the May 2018 exams** of the DP students in the sample were transferred from IB to the PI in the beginning of 2019.

Nature of the data

The complete data consist of the following components:

- **Sampling frame**: list of students invited to participate in the surveys. The sampling frame contains student names, birth date, contact information (email address) and school name. For students who took the final exam, the sampling frame also lists the exam code.
- **Student survey data**: obtained from the responding students in three waves.
- **Exam results**: obtained from IB's IBIS registration system. The data provided by students at the three waves were supplemented with IB-records on participation in the final exam and the results (grades) obtained for each topic at the exam.
- **School survey data**: obtained from a questionnaire send to the DP coordinators to all DP schools (including the participating schools) during the wave 2 data collection.

Survey questions

- **Questions** asked to the **students** included:
  - Workload as experienced by students: **subjective workload**.
  - Feelings of **stress**, using four standard scales taken from the psychological research literature.
    - 'Restlessness' = Quality of Life Scale[1]
    - 'Lack of Control' = Perceived Stress Scale[2]
    - 'Life Dissatisfaction' = Student Life Satisfaction Scale[3]
    - 'Physical Health Problems' = Short Form-12 Health Survey[4]
  - Participation in the DP: **number and kind of subjects** taken at higher level (HL) and standard level (SL) respectively.
  - Students' evaluation of the **difficulty and time burden** of the different elements of the DP.
  - **Time spent on school work**, extracurricular school activities and leisure time activities **Support and pressure** from schools (policy), teachers (pedagogues), parents and peers ('**P-factors'**).
  - **Suggestions** how to make the DP workload more manageable for students.

- **Questions** asked to **DP coordinators** and heads of school included:
  - **School characteristics**
  - Well-being-related school **policies**
  - Their perception of the **level of difficulty and time burden** of the different elements of the DP.

---

[1] Whoqol Group. (1994). The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In *Quality of life assessment: international perspectives,* 41-57. Springer, Berlin, Heidelberg.

[2] Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, *24*(4), 385–396.

[3] Huebner, E. S. (1991). Correlates of life satisfaction in children. *School Psychology Quarterly*, 6, 103–111.

[4] Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care*, *34*(3), 220-233.

o **Suggestions** how to make the DP workload more manageable for students.

4. Sample selection and response

- As of 3 September 2019, the **Diploma Programme is offered by 3,421 schools in 157 countries worldwide**. DP schools are differentiated by language of instruction (English, Spanish and French) and by exam period (Northern hemisphere schools have May exams, Southern hemisphere schools have November exams). The IB does not enroll students when they enter the DP, but only registers them when they enroll for the exam, by the end of Year 2. For this reason, the sample could not be selected from the IB administrative data (IBIS).

- The **sampling procedure** consisted of the following steps:
  o **Selection of Northern hemisphere countries with May exams and English, Spanish or French as working languages**. Southern hemisphere countries were selected into the sample by design if they have May schedules: Argentina, Australia, New Zealand, South Africa, Malawi, Lesotho, Mozambique. 367 schools had November session exams (despite some having May candidates) and 49 schools had a language of instruction other than English, Spanish or French. 139 countries with 2,599 schools remained in the sampling frame.
  o **Stratified random sampling of these schools, proportional to the number of DP schools in the 139 countries**. Countries were ordered with respect to the number of DP students in the previous year. 100 countries were then selected systematically at random. This selection procedure guarantees that the selected sample is representative with respect to the size of the student body in countries. The sample of schools from the selected countries was made proportional to the number of schools; each country is represented by a number of schools that is proportional to the number of DP schools in that country. If, however, only one school in a country was selected, this was upped to two schools, if available. This means that the minimum cluster size is two, which makes it possible to distinguish school differences from country differences. Yet, five countries are represented by one DP school only because they have only this one.
  o This resulted in a **sample of 258 schools** (very close to 10%) **from 50 countries.** Due to the adapted sampling procedure, the selected sample of schools is highly representative with respect to Continent, Language, Number of DP schools in country (with May schedules), and Status (private/state).
  o **DP Coordinators were invited by email to send complete lists of DP year 1 students in their schools, with contact information**. 98 DP Coordinators cooperated, but some selected only a random part of the students. This resulted in a list of **4,855 DP students from 98 schools**.
  o Students were asked for **explicit consent** to collect and use their information. A few dozen did not consent explicitly and did not take the interview. Other students did not start the interview at all and thus refused participation implicitly. They are counted as non-response.
  o Finally, students could **skip questions** during the interview and **drop out of the interview** at any time. This leads to item (partial) non-response that gradually increases during the interview.

- o **DP Coordinators and heads of school were also asked for explicit consent to collect and use their information.** Inadvertently, all DP schools were approached instead of only the schools in the study which provides our study with more information than intended. **2,894 DP coordinators and heads of school** were approached and around 2,200 filled out the survey, including 119 DP coordinators & head of schools from 91 schools where the student surveys were also distributed.

- **Table 1** gives an overview of the various types of **response and non-response at the level of students**. The multi-wave panel design and the occurrence of partial non-response make it hard to define the response by a single number. Some crucial summary numbers are:
  - o 3,565 students participated at some point in the project with some response,
  - o 1,063 students participated in all three waves and have provided more or less complete information. However, users have to calculate valid N for each analysis separately.

Table 1: Response in student surveys and exam records

	Students wave 1	Students wave 2	Students wave 3	Students Exams	School: Students
Sample (approached)	4,855	4,855	4,855	4,855	4,855
No response	2,375	2,439	2,813	NA	NA
Collaboration	2,480	2,416	2,042	NA	NA
No consent	20	80	43	NA	NA
Refusal (completely empty)	477	612	500	NA	NA
Some response	1,983	1,724	1,499	NA	NA
Effective sample: full response on Subjective	1,774	1,527	1,499	3,984	4,794
Response rate	36.5%	31.5%	30.9%	82.0%	98.7%

The data for the schools referred to the number of students in the sampled schools covered by the school coordinators data.

5. Data management protocol

Ethics

- **IRB approval** was obtained from the **Advarra Institutional Review Board (IRB)**, an independent provider of IRB services based in Canada. Advarra **evaluated the research ethics and approved** each of the following throughout the research period:
  - o Research plan
  - o Questionnaires
  - o Consent forms
  - o Parent information

- Data have been acquired with **informed consent of students, parents and DP coordinators**. All were informed about the goal of the data collection and were assured that the data would be handled confidentially.

- **Confidentiality** means that directly identifying information was immediately separated from the data and preserved only as long as needed to prepare data collection and analysis. During data-collection respondents were identified by their email-address used to invite them to participate. Their identities were checked at each wave by asking for gender and exact date of birth, which was compared to the sampling frame. The email addresses were then replaced by a non-informative RESPNR. Names of school were replaced by a non-informative SCHOOLNR. Indirectly identifying data (in particular exact birth dates and school number) was subsequently recoded to be non-identifying. The resulting anonymized data, along with the documentation will be stored in the Dutch national data-archive DANS. Access will be made contingent on permission granted by the IB for third party use.

- The collected **information was classified** into three groups:
  - **Secret:** This concerns directly identifying information, in particular students' email address, exact date of birth, and the school name. This information is part of the Sampling Frame, which can only be matched with the other data only using the newly created RESPNR (respondent number) and SCHOOLNR. The secret information is encrypted and held password-protected and is never to be shared with any other user of the data than one of the Principal Investigators. The secret information is not archived but destroyed after the finalization of the project.
  - **Confidential**: This concerns answers to open questions to be coded, such as parental occupations, aspired education and occupation, as well open comments solicited at various times in the online interviews. This information is in its verbatim format separated from the main data file and organized in coding files. These coding files were used to categorize the open information into numerical codes, either matching with standard classifications (such as the International Standard Classification of Occupations) or coding schemes that were specific to the project. This confidential information was only shared with data coders and will be available for professional use at special request (eg, to check the coding).
  - **Public:** The remaining information has been anonymized. Students can be identified by country number and year of birth, but schools are not recognizable, nor are specific birth dates.

- Both confidential and public-use information is archived at the **DANS Archive in The Hague** and available for professional use, but permission for use needs to be obtained from the IB Organization. The secret information was destroyed at the end of the project.

Anonymization

- All **directly identifying information** (name, email address, exact birth date) was removed from the confidential and public data and replaced by an uninformative number that was generated at VU University. School names were replaced by an uninformative school number that was generated at VU University (this is not connected to the internal school code in the IBIS administration).

- **Indirectly identifying information** (such as answers to open questions) is stored in the confidential data file.

Variables

- Administrative variables

    o **RESPNR** is the fixed respondent number, built by $CASEID after sorting on a randomly generated uniform number (Set SEED 180453).
    o **SCHOOLNR** derived from SCHOOLCODE (an internal IB identifier), random ordered before recode.
    o **aSEQNR** is the internal response number for wave 1, as provided by IB Research.
    o **bSEQNR** is the internal response number for wave 2, as provided by IB Research.
    o **cSEQNR** is the internal response number for wave 3, as provided by IB Research.
    o **GENDER** is the gender according to the sampling frame.

- **Variable naming** generally followed the organization and ordering of the questionnaires:
    o Wave1: a01 – a103
    o Wave2: b01 – b77
    o Wave3: c01 – c60
    o Schools: q01 – q74

- All **string variables** were removed from the data files and stored in confidential files. They consist of occupations, subjects and invited comments.
    o **Occupations** were coded according to ISCO-88 (International Standard Classification of Occupations).
    o **Subjects** were coded into a standard hierarchical code: DP subject group + subject code + subject level (HL/SL).
    o **Invited comments** were only stored.

Files and directories

- All **data files with confidential information** were stored in a subdirectory Confidential (password protected). They were kept on only one medium, Co-PI1's external hard-drive, which is not connected to the internet.

- All other project members and collaborators deleted files with confidential information from their drives (and email boxes).

- The questionnaires (as shown on-screen) did not contain question numbers, but **question numbers were added in post-processing to the pdf prints.** All variables and numerical values were labeled in close correspondence with the questionnaires. Codes for the open questions were added in post-processing and labeled to obtain optimal transparency.

- The final **data-file** contains the variables in order of data collection. For data-analysis, a correspondence table between the three waves has been prepared. In this **variables' correspondence table**, the variables have been organized and reordered by topic.

#	File name	Records	Columns	Contents
1	Main_quex_schools_screen_nrs.pdf	23 pp		Questionnaire
2	Main_quex_wave1_screen_nrs.pdf	22 pp		Questionnaire
3	Main_quex_wave2_screen_nrs.pdf	22 pp		Questionnaire
4	Main_quex_wave3_screen_nrs.pdf	22 pp		Questionnaire
5	Main_wave123_varlist.xlsx	450		Variable corres-pondence table
6	Main_School_Data_NN.sav	N=2894	NV=231	Numerical data
7	Main_School_Constructed_vars.sav			
8	Main_Student_Data_wave123_NN.sav	N=4858	NV=633	Numerical data
9	Main_Student_Constructed_vars.sav	N=4858	NV=19	Numerical data
10	Main_Student_SES_vars.sav	N=4858	NV=13	Numerical data
11	Exam_results_NN.sav	N=32220	NV=11	Numerical data