INTEGRITY

D2.2 Final report on questionnaire and qualitative interviews

Deliverable Code	D2.2
Distribution Level	Public
Agreed upon by General Assembly	Date: 01 31 2021
Responsible Partner	UCPH
Checked by WP leader	Peter Sandøe Date: 16/1 2021
Approved by Project Coordinator	Mariëtte van den Hoven (UU) Date: 25/1 2021

This project received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 824586

Contents

List of Abbreviations

HUM: Humanities including design and theology

SOC SCI: Social science including law

STEM: Here used to denote the natural and medical sciences as well as engineering, mathematics, and computer science

1.Preface

This report presents results from task 2.2 and task 2.3 within work package (WP) 2 (Examine: survey tool) in the H2020 project **INTEGRITY**.

The overall goal of WP 2 is to construct and apply a survey tool that can provide evidence-based input to the development of an effective and innovative tool for teaching research integrity (undertaken in WP4 and WP5). Specifically, the aims of the WP are to:

1. **identify, map and evaluate** the existing understanding of students and early career researchers in terms of academic and research integrity and key concepts in relation to questionable research practice and research misconduct, including fabrication, falsification, and plagiarism via a new survey tool developed for this purpose;
2. **map and categorise** the grey area issues relating to good scientific practice that students and early career researchers are most likely to face in their daily practice, and the extent to which these vary across cultural and disciplinary backgrounds.

To this effect, WP2 first embarked (in task 2.1) on an explorative qualitative investigation of the integrity dilemmas students and early career researchers face and how they think about them. Some results from this investigation were presented in D2.1 (2019). Following up on this, in tasks 2.2 and 2.3, a survey tool for mapping issues and the understanding of good scientific practice was developed and applied to students at three different levels and across disciplines in nine different countries. The main results of the survey are reported here. Where relevant, insights from the interview study are included in the report, but the reader is referred to D2.1 for a detailed presentation of the qualitative study (see also Goddiksen et al. (2020)).

Due to the COVID-19 outbreak, data collection was delayed. In the original plan, all quantitative data was supposed to be collected from January to June 2020. To allow for a second round of data collection in autumn 2020, the deadline for data collection was extended to 10 December 2020. This allowed the vast majority of the planned data to be collected, and more than 6,000 students answered the questionnaire. However, even with the extended deadline, not all countries managed to obtain large enough datasets for all study levels, and we did not obtain full datasets for any study levels in Germany.

To ensure a meaningful analysis, only completed datasets (for a given country and a given study level) are included in this report. In Denmark, Ireland, Portugal and Switzerland, data collection was completed

for all study levels. In Slovenia and Lithuania, the survey was completed for high school and bachelor students, and in Hungary, it was completed for PhD and bachelor students. In the Netherlands, the target was not reached for any study levels, but sufficient data for PhD students were gathered to allow for a meaningful analysis and thus this dataset is also included in the analysis. Detailed information on participants and datasets can be found in Chapter 3.

Fortunately, the completed datasets are rich enough to allow the desired analysis of the possible differences across study levels, study directions and cultures. In this report, we will mainly focus on the differences across study levels, as these differences have the most direct consequences on the work in INTEGRITY. We will, however, comment on other differences where relevant.

The report begins with a summary of the major results (Chapter 2), including some implications for teaching. Next, it lays out data and materials (Chapter 3), and then presents some main findings covering central concepts. These include self-reported textbook understanding and practical challenges experienced (Chapter 4), conceptions about questionable research practice and misconduct (Chapter 5), perceptions about misconduct among peers compared with own behaviour (Chapter 6), and what the students have learned about academic and research integrity through courses, etc., as well as what topics they want to learn more about (Chapter 7). Finally (Chapter 8), we present some results on the potential effect of research integrity training on the likelihood that PhD students will engage in questionable practice.

The focus of this report is to present results that may serve to inspire ongoing and subsequent work in INTEGRITY. We are confident that the findings presented here will be of great help in identifying knowledge gaps, motivations (or lack thereof) and pertinent issues at the three levels and in the different fields of study that the tools resulting from INTEGRITY must address.

The report is the joint effort of the following teams and team members:

University of Copenhagen: Mads Paludan Goddiksen, Mikkel Willum Johansen, Marcus Tang Merit, Thomas Bøker Lund & Peter Sandøe.

University of Utrecht: Mariëtte v.d. Hoven, André Krom

Trinity College Dublin: PJ Wall, Una Quinn, Linda Hogan

University of Debrecen: Nóra Kovács, Anna Catharina Vieira Armond, Orsolya Varga

University of Ljubljana: Vojko Strahovnik, Mateja Centa, Roman Globokar

University of Geneva: Céline Schöpfer, Christine Clavien

University of Zürich: Roberto Andorno, Sandra Rossi

University of Vilnius: Margarita Poškutė, Eugenijus Gefenas

University of Münster: Maren Behrensen

University of Porto: Rita Santos, Júlio Borlido Santos, Anna Olsson

The authors would like to thank all the participants who took the time to fill in our survey or participate in our interviews, as well as all the teachers, administrators and others who helped with the recruitment. Special thanks to Hanne Andersen for her input and to Sarah Layhe for language editing.

The report should be referenced in the following way:

Mads Paludan Goddiksen, Mikkel Willum Johansen, Marcus Tang Merit, Thomas Bøker Lund, Roberto Andorno, Anna Catharina Vieira Armond, Maren Behrensen, Mateja Centa, Christine Clavien, Eugenijus Gefenas, Roman Globokar, Linda Hogan, Mariëtte v.d. Hoven, Nóra Kovács, André Krom, Anna Olsson, Margarita Poškutė, Una Quinn, Sandra Rossi, Rita Santos, Júlio Borlido Santos, Céline Schöpfer, Vojko Strahovnik, Orsolya Varga, PJ Wall & Peter Sandøe: *INTEGRITY, Final report on questionnaire and qualitative interviews*. Copenhagen, 2021, available at the following link: http://h2020integrity.eu/

The teams plan to turn the main findings from the report into a number of papers that will be submitted to relevant international journals. These papers will be made available on the project website (http://h2020integrity.eu/) as they appear.

Copenhagen 31 January 2021, on behalf of the authors

Peter Sandøe

2. Executive summary and implications for teaching

We will begin with a summary of the results that we find most relevant for teachers and others working with research and academic integrity training. We start with some general results and then turn specifically to each of the three levels (PhD, bachelor and high school) in focus in the different streams of WP4 of INTEGRITY.

2.1. Across levels

First, our study shows that students generally believe that they know how to act ethically in relation to our three main themes: 1) drawing on the work of others, 2) collaboration and authorship, and 3) collecting, analysing and presenting data (section 4.1, Figure 1). This result has several implications for teaching. Most importantly, it can pose a challenge to motivation, as there are few things less motivating than expecting to already know what you are about to be taught. On the other hand, it seems that many students had been in situations of doubt within the past 12 months (section 4.2, Figure 2). To increase motivation and to show that students are not alone with their doubts, it can therefore be relevant include specific situations in teaching, and to show the students how widespread doubts about academic integrity issues are, and the specific issues about which students are typically in doubt (Chapter 6).

Second, although students most frequently experienced doubt in situations related to plagiarism (section 4.2, Figure 2), the qualitative interviews showed (as highlighted in D2.1) that they were less confident when faced with situations involving academic integrity issues concerning collaboration, and data collection and analysis than issues about plagiarism. This is especially true at bachelor and high school level. There is therefore a need to bring these topics into focus for students at these early levels.

Third, students' beliefs that they know how to behave in an ethically correct way may, in some cases, be false. For most of the questions and cases used to test the students' knowledge and ability to handle specific ethical situations, we see a clear gradient following study level – in that students at higher levels were generally more knowledgeable and more competent in relation to issues of research and academic integrity. False beliefs about one's own knowledge of the field were therefore most clearly present in the high school student populations.

Despite their high level of confidence, high school students were generally not able to discern the subtle difference between permissible and inadmissible paraphrasing in a concrete example (section 5.1.2, Figure 6). It is also alarming to see that only around 60% of the high school students in Ireland (63%) and Lithuania (61%) considered a clear example of plagiarism (copying an entire page) to be plagiarism (section 4.1.2, Figure 5). Similarly, large minorities (about 25%) of high school students in Switzerland, Lithuania and Portugal did not consider paying someone to write an assignment for them to be against the official rules (section 5.2.1, Figure 8). It is therefore clear that students' beliefs are not always in line with what is officially considered good practice. These results pose a particular challenge to academic integrity teaching, as such false beliefs will have to be challenged.

Fourth, false beliefs may also play out in other ways. When faced with more ambiguous scenarios, such as getting help from friends and family or comparing results with other students for individual assignments (section 5.2.1, Figure 8), a large majority of bachelor students believed there to be clear rules that either forbid or allow the action, although they were also given the answer options: "It depends on the situation" and "The rules are unclear". From the qualitative interviews, we know that such actions are handled very differently in different situations and by different teachers. The false belief that there are clear rules when in fact there are not can be just as detrimental as a direct misunderstanding of the rules, as this kind of false belief may stop students from asking their teacher about how they are expected to behave in the specific situation. It is particularly challenging to teach such issues as they involve grey zones or even situations where teachers expect actions that – strictly speaking – may go against widely recognised norms.

Fifth, as another general observation, there was (especially for bachelor and PhD students) a discrepancy between what the students subjectively perceived to be important issues to teach (Chapter 7, Figure 16) and the types of transgressions they reported as common among their peers. This may also pose a real challenge to academic integrity teaching. In general, students are mainly interested in plagiarism. This is perhaps unsurprising as plagiarism is the type of questionable practice that is likely to be most clearly and harshly enforced by institutions. Students who wish to learn about plagiarism may thus be motivated by the simple desire to avoid getting into trouble themselves. However, from the point of view of research integrity, it may be at least as important to address other issues.

Sixth, we analysed the gender distribution of answers to questions related to perception of (questionable) peer behaviour and own questionable behaviour. The answers to a number of these questions differed significantly across gender identities, especially for the high school population (see

section 6.5, Figure 14a-g). In particular, male-identifying high school students were somewhat more likely than female-identifying high school students to add other students as authors of a group assignment, even when the other students did not contribute to the work. It is clearly relevant to take these gender-related differences into account in integrity teaching.

Seventh, although there is a clear gradient in knowledge and competence, such that students at higher study levels have a better understanding of certain integrity issues, this does not mean that students at higher study levels should not learn about integrity. As the qualitative interviews highlighted, issues may present themselves in different ways at different study levels, and students at higher levels are required to understand different aspects and more nuances of the issues (see D2.1).

Finally, the results presented in Chapter 8 indicate that there is no discernible correlation between time spent on RI training and a propensity to conduct questionable practices. However, there is some correlation between the degree to which students believed their peers engage in questionable practices and their own propensity to do the same. The positive take on this finding is that there is a lot of room for improvement when it comes to developing academic integrity training. A central aim of such teaching should be to provide students with the knowledge and tools to influence the research culture they are in, and thus help to develop not only the students' own academic integrity, but also a culture of academic integrity.

2.2. PhD level

The questionable practices that the PhD students perceived to be most common among their peers are: "…keeping inadequate records of parts of research/work that should be documented"; "…allowing research group leaders, supervisors or others in power to become co-authors of papers, even though they did not make a significant contribution to them", and "…citing sources that are not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour". Among these three forms of conduct, students were most inclined to report that they themselves had "…allowed [people in power] to become co-authors" (21% in Hungary to 38% in Portugal), and "…kept inadequate records" (across countries: 10% in Portugal to 29% in Ireland).

All these issues are to some extent related to the structure/culture the student is working within. It is extremely important to be aware of this when preparing research integrity training for PhD students. Simply teaching the students what is right and wrong in relation to these issues may not be enough to

address these problems, since the students may not be in a position where they can choose how to behave. Students should also be given empowering tools to cope with or change a detrimental culture or structure. This is underlined by the results presented in Chapter 8, which show a positive correlation between the students' own propensity to behave unethically and the perceived level of questionable practice among peers, whereas RI courses did not seem to have a discernible effect. This underlines both the need to develop better RI teaching and the need to address the culture and not just the conduct of the individual in the teaching.

In relation to authorship, it is also important to be aware of (and make the students and supervisors aware of) the marked differences between the various scientific fields (see section 5.2.2).

2.3. Bachelor level

Bachelor students perceived "...receiving help from other students or family members on assignments you were supposed to complete on your own" and "...adding students as co-authors of group assignments, even though they did not contribute" to be the most common questionable behaviours among their peers. Across countries, 22% (Switzerland) to 62% (Portugal) of the students admitted to having "...added students as co-authors of group assignments, even though they did not contribute", and 40% (Denmark) to 62% (Portugal) admitted to having "..received help from other students or family members on assignments you were supposed to complete on your own".

This suggests that although students are aware of plagiarism as an important issue, the grey zones and possible ethical transgressions related to collaboration and receiving help must also be addressed in academic integrity training.

Furthermore, although bachelor students generally believe that they know how to behave ethically when drawing on the work of others, their judgement in actual cases is not always aligned with that of more experienced academics (Section 5.1.2). When faced with a specific scenario, more than 50% of Hungarian and Portuguese bachelor students, and more than 45% of Danish students judged that direct copying of a short passage without reference was not unacceptable.

2.4. High school level

Similar to bachelor students, high school students generally believe that they know how to behave ethically when drawing on the work of others, but their judgement in specific cases is not always aligned with that of more experienced academics (section 5.1.2).

Even the most fundamental academic integrity issues related to data analysis seem to be a blind spot for students in high school. Roughly 25% (21-35% across countries) of the respondents answered "I don't know" to a question about whether it was permissible to replace outlying data points with constructed data without being transparent about it (section 6.4). Combined with the general observation presented above that academic integrity in relation to data analysis is under-addressed, this suggests that there is a need to introduce high school students (and probably also their teachers) to this topic.

The high school students saw receiving "...help from other students or family members on assignments you were supposed to complete on your own" as the most common questionable practice among their peers. This is also the most frequently reported behaviour of the students themselves across countries, as 68% (Lithuania) to 77% (Slovenia) admitted to having done this. It is important to discuss this topic, but also to understand the culture at the specific school, as it may well be unclear to both students and teachers when an assignment is supposed to be conducted individually.

3.Data and materials

3.1. Countries and study levels included

Nine countries were originally included in the study: Denmark, Germany, Hungary, Ireland, Lithuania, Netherlands, Portugal, Slovenia, and Switzerland. We assume that these countries cover much of the variation among European countries with respect to didactic and academic tradition.

A key notion of INTEGRITY is that academic and research integrity is gained through a gradual learning process that should be introduced and taught as soon as possible – i.e. as early as high school – and continue through all levels of the educational system, so as to cover issues specific to the different study levels. For this reason, we aimed to include multiple study levels in different areas of the educational spectrum: high school, bachelor, and PhD. We defined these levels largely according to the definitions laid down by the Bologna process (https://ehea.info/). Thus, we defined a *PhD student* as a student in the third cycle defined in the Bologna declaration. Students at this level are preparing a dissertation based on one or more research projects guided by one or more senior researchers (in some countries, students taking special research master's programmes fit this definition, in which case they were included and treated as PhD students). A *bachelor student* was defined as a student in the first cycle defined in the Bologna declaration, enrolled at an institution that also offers PhD level education, and having earned less than 180 ECTS points. Furthermore, we decided to focus on students who had earned at least 60 ECTS points. Finally, a *high school student* was defined as a student in a programme that is a sufficient requirement for entering one or more bachelor programmes, but that is not itself a bachelor programme.

Consequently, the general procedure defined in the data collection plan was to obtain responses from students at each level across all countries. Unfortunately, in some countries, practical challenges made it difficult to collect a sufficient amount of data across all study levels. Furthermore, we did not succeed in collecting enough data for any of the three study levels in one country. This will be detailed in the sections below.

3.2. Sampling and recruitment

3.2.1 Planned sampling and recruitment procedures

In our data collection plan, we opted for a random selection of schools or study programmes. we planned to have a list of institutions compiled for each study level . More specifically, at the high school level, the aim was to construct a database with a complete list of high schools. At the bachelor level, the aim was to construct a database with a list of all study programmes, and where possible, distinguish between programmes within three main study areas:

- STEM: including natural and medical sciences, engineering, as well as mathematics and computer science
- Social science (SOC SCI): including classic social sciences like economics and political science as well as law
- Humanities (HUM): including the arts, as well as theology and design

At the PhD level, the aim was to construct a database with a list of study programmes/schools, and where possible, distinguish between study programmes within the above-mentioned three main study areas STEM, SOC SCI, and HUM.

Using these databases, we then planned to randomly select high schools, study programmes (at the bachelor level), and PhD schools that would be invited to participate in the study. All students in the selected bachelor study programmes and PhD Schools were chosen for inclusion, except for bachelor students who had completed less than one year of their programme. A flexible inclusion procedure was chosen for high schools, depending on what was possible for the specific targeted school – i.e. the number of students invited from the selected high schools could range from all students in a particular class (randomly chosen by an institutional administrator) to all students in a study year.

Multi-stage random sampling procedures require that random sampling is conducted in several stages, and not only at the aggregate level, e.g. the school/study programme. It soon became apparent, however, that it was not possible to carry out a multi-stage sampling procedure at the within-institution level, i.e. where students (within the randomly selected institutions) were randomly invited to participate. This was because for most countries it was not possible to procure individual-level contact information for students within the institutions. For this reason, we opted to include entire classes or study years (at high school level), entire study programmes (at bachelor level), and entire PhD schools. An advantage

of including all students at the aggregate unit (in particular entire high school classes or study years) is that it then became possible to study differences in teaching culture, as this requires that an entire class or pool of students respond to the questionnaire.

In some cases, a total population study was carried out, i.e. all institutions were invited. This was the case where practically necessary (and feasible). For instance, we decided to contact all HUM PhD schools in Denmark because there are so few of them.

3.2.2 Country-specific recruitment procedures, sample sizes, and response rates

In this section, we describe the recruitment procedure in the nine countries originally selected to take part in the study. We highlight deviations from the procedure laid out in 3.2.1, e.g. where databases could not be compiled, or where different institutional set-ups demanded other sampling procedures, as well as other challenges.

Denmark

In Denmark, the procedure described in 3.2.1 was largely followed.

At high school level, 30 of the 190 high schools in Denmark were randomly chosen to participate. The school administration was contacted and invited to participate in one of two ways: either by distributing links to the survey to students who were 18 years or older, or by allowing members of the project to visit the school and give a short seminar where students would first answer the questionnaire, followed by a short presentation and discussion. Seven high schools agreed to participate (we visited four schools, three distributed the link). The estimated student population size at these seven high schools (only including students who were 18 years or older) was approximately 1,170. We received 389 complete responses, giving a 33% response rate.

At the bachelor level, there are 345 study programmes at nine university institutions in Denmark. We divided these programmes into three study directions: STEM, SOC SCI, and HUM, and randomly selected programmes within each of them. For each selected programme, the study programme leader was contacted, and if the leader agreed to participate, we invited bachelor students in the second and/or third year of study, depending on what the study leader agreed. In some cases, teachers on the programmes selected were also contacted and asked to distribute the questionnaire in class. In total, 66 programmes were invited and 38 agreed to participate (17 STEM, 10 SOC SCI, and 11 HUM programmes). The estimated student population size at these 38 institutions was approx. 7,200. We received completed questionnaires from 218 students, giving a 3% response rate.

All PhD students in Denmark are affiliated with a PhD school. There are 49 such schools: 32 within STEM, 12 within SOC SCI, and 7 within HUM. We randomly selected ten of the STEM PhD schools to participate in the study. Due to the low number of SOC SCI and HUM schools, we invited all of them to participate. In total, 29 PhD schools were contacted and 23 agreed to participate. The estimated student population size at these 23 institutions was approx. 4,200. We received completed questionnaires from 427 students, giving a 10% response rate.

Germany

In Germany, convenience sampling was employed. Heads of relevant schools and programmes were contacted and asked to distribute the link to the questionnaire. We do not provide details, as the recruitment did not result in sufficient data to be included in the analysis below.

Hungary

In Hungary, the procedure described in 3.2.1 was largely followed at bachelor and PhD level. The major deviation was that the unit of sample was universities rather than programmes. Details on the recruitment of high school students is not presented here, as the recruitment did not result in sufficient data to be included in the analysis.

At the bachelor level, five universities per study area were randomly chosen (five universities offering bachelor programmes in STEM, five universities with SOC SCI programmes, and five with HUM programmes). All faculties of the selected universities were then identified. At each selected university, five faculties offering bachelor programmes for the relevant study area were randomly selected. If there were fewer than five faculties per discipline at a selected university, they were all contacted.

All students from the selected faculties who had completed at least their first year of bachelor studies were contacted via the education offices (following permission by the dean of the faculty). The actual number of students were given by the education offices.

A total of 51 faculties were contacted (22 STEM, 19 SOC SCI, 10 HUM), 23 of which agreed to participate (10 STEM, 9 SOC SCI, 4 HUM). The estimated population at these 23 institutions was approx. 15,694. We received completed questionnaires from 292, giving a 2% response rate.

At the PhD level, doctoral schools were selected using a stratified sampling method. Each university may have several doctoral schools. Based on a public list of universities in Hungary, five universities per study area (five with PhD programmes in STEM, five with PhD programmes in SOC SCI, five with PhD programmes in HUM) were randomly selected. All doctoral schools at the selected universities were then

listed, and five doctoral schools for each university offering doctoral programmes in the given study area were randomly selected. If there were fewer than 5 doctoral schools within the given study area, all were invited. All students from the selected doctoral schools who had completed at least the first year of their PhD programme were contacted by the PhD office of their university (permission from the head of the doctoral school was necessary). A total of 92 doctoral schools were contacted (40 STEM, 29 SOC SCI, 23 HUM), and 50 agreed to participate (22 STEM, 17 SOC SCI, 11 HUM). The estimated population size at these 50 institutions was approx. 2,082. We received completed questionnaires from 221, giving an 11% response rate.

Ireland

In Ireland, the procedure described in 3.2.1 was largely followed, except that a full population recruitment was attempted at PhD level.

At high school level, 39 of 724 high schools were randomly selected. The directors were contacted and asked to distribute the questionnaire links to the transition year students. Fourteen schools agreed to participate. The estimated population size at these institutions was 932. We received completed questionnaires from 292, giving a response rate of 31%.

At bachelor level, 63 of 623 programmes were randomly selected (19 STEM, 23 SOC SCI, 21 HUM). Programme administrators were contacted and invited to distribute the link to the questionnaire via email. In total, 27 agreed to participate (8 STEM, 9 SOC SCI, 10 HUM). The estimated population size at these institutions was 2,500. We received completed questionnaires from 231, giving a response rate of 9%.

At PhD level, all Deans of Postgraduate Studies, postgraduate student unions and student councils, as well as various other PhD networks at the nine Irish Universities were contacted. Two universities declined to participate. In the remaining seven institutions, the estimated population size was 6,518. We received completed questionnaires from 245, giving a response rate of 4%.

Lithuania

In Lithuania, high school students were recruited largely as described in 3.2.1. At bachelor level, we decided to use a convenience sampling approach. We do not report details on the recruitment at PhD level, as this did not result in sufficient data to be included in the analysis below.

At high school level, 112 of the 401 high schools in the country were randomly selected. The directors were contacted and asked to distribute the questionnaire links to the final year students (over 18 years

old). In total, 109 schools agreed to participate. The estimated population size at these institutions was 8,561. We received completed questionnaires from 215, giving a response rate of 3%.

At bachelor level, administrators of 118 relevant programmes were contacted (37 STEM, 35 SOC SCI, 46 HUM) and asked to distribute the link to the survey. All agreed to participate. The estimated population size at these institutions was 10,500. We received completed questionnaires from 204, giving a response rate of 2%.

The Netherlands

In the Netherlands, convenience sampling was used to recruit PhD students. Details on the recruitment of bachelor and high school students are not presented here as data were not included in the analysis below due to the low number of responses.

All 2,400 PhD students at the University of Utrecht were invited to participate through their graduate school. Furthermore, PhD students from other Dutch universities were recruited through personal networks. It has not been possible to estimate the PhD population size, and a response rate could therefore not be estimated.

Portugal

In Portugal, the procedure described in 3.2.1 was largely followed. Of 667 high schools in mainland Portugal, 23 were randomly selected and invited to participate by either distributing the survey in class or by having a project member visit and distribute the survey as described above for Denmark. Five schools agreed to participate. Of these, four distributed the questionnaire in class and one was visited by a project member. Two additional high schools were contacted through personal contacts late in the process to obtain a sufficient number of responses. The estimated population size at these institutions was 250. We received completed questionnaires from 219, giving a response rate of 88%.

At bachelor level, 288 out of 825 programmes were randomly selected (52 STEM, 89 SOC SCI, 147 HUM). The director was contacted and invited to distribute the link to the survey. In total, 117 agreed to participate (17 STEM, 45 SOC SCI, 55 HUM). The estimated population size at these institutions was 11,700. We received completed questionnaires from 274, giving a response rate of 2%.

At PhD level, 294 out of 702 PhD programmes were randomly selected (65 STEM, 87 SOCSCI, 142 HUM). The head of the programmes were contacted and invited to distribute the link to the survey. In total, 111 programmes agreed to participate (23 STEM, 40 SOCSCI, 48 HUM). The estimated population size at these institutions was 4,995. We received completed questionnaires from 241, giving a response rate

of 5%. An additional round of recruitment was conducted towards the end of the data collection to obtain a sufficient number of responses from HUM PhD students. In this round, personal contacts at the teacher or administrative level were approached to obtain the last few responses needed to reach the target of 60.

Slovenia

In Slovenia, convenience sampling was employed at both bachelor and high school level. Details on recruitment for PhD level are not presented, as the data are not included in the analysis due to an insufficient number of responses.

At high school level, 12 out of 149 schools were approached. Most of these schools were selected through personal contacts at the teacher or administrative level. The link to the survey was distributed to students via email. The estimated population size at these institutions was 1,100. We received completed questionnaires from 250, giving a response rate of 23%.

At bachelor level, 15 out of 251 programmes were contacted (7 STEM, 4 SOC SCI, 4 HUM). These programmes were randomly selected, and the link was distributed to students via email. In addition, invitations to the survey were circulated to students in the selected programmes on Facebook and via personal contacts. The estimated population size was 9,000. We received completed questionnaires from 221, giving a response rate of 3%.

Switzerland

Recruitment in Switzerland was achieved via two parallel efforts: one by project members from the University of Geneva (mainly targeting the French-speaking population) and one by project members from the University of Zürich (mainly targeting the German-speaking population). Although we present the combined data in the analysis below, the recruitment strategies differed somewhat, and we will describe each in turn.

The recruitment conducted by the University of Geneva targeted almost the entire population. All 34 high schools in the French-speaking part of Switzerland were contacted and invited to distribute the link to the questionnaire via email. Of these, 19 agreed to participate. The estimated population size (students over 18 years of age) was 2,860. We received completed questionnaires from 360, giving a response rate of 13%.

At bachelor level, the heads of all of the 181 bachelor programmes in the five French-speaking universities were contacted and invited to distribute the link to the questionnaire via email. Of these, 71

(24 STEM, 12 SOC SCI, 35 HUM) agreed to participate. The estimated population size was 11,600. We received completed questionnaires from 199, giving a response rate of 2%.

At PhD level, the management of all five French-speaking universities were contacted and invited to distribute the link to the questionnaire via email. Subsequently, teachers and deans on specific programmes were contacted to send out reminders. The estimated population size was 3,300. We received completed questionnaires from 202, giving a response rate of 6%.

The recruitment conducted by the University of Zürich at high school level followed a random sampling procedure. At bachelor and PhD levels, convenience sampling was largely used.

At high school level, we decided to focus only on high schools within the five most populous German-speaking cantons (Zurich, Bern, Saint Gallen, Aargau and Lucerne). Of the 76 high schools, 35 were selected at random, and the school director was approached and asked to distribute the link to the questionnaire via email. Ten schools agreed to participate. The estimated population size at these institutions was 1,330. We received completed questionnaires from 76, giving a response rate of 6%.

At bachelor level, 30 programmes were initially drawn from a complete list of 150 programmes. Heads of studies were contacted and invited to distribute the questionnaire to the students via email. Only two programmes agreed to participate (1 SOC SCI, 1 HUM). Personal contacts were therefore used to invite third year students from a STEM programme. The estimated population size at these institutions was 430. We received completed questionnaires from 11, giving a response rate of 3%.

At PhD level, personal contacts for four PhD programmes (2 STEM, 1 SOC SCI, 1 HUM) were used to distribute the survey. The estimated population size at these institutions was 290. We received completed questionnaires from 59, giving a response rate of 20%.

3.3. Ethical approval and parental consent

Before embarking on the recruitment of respondents as outlined above, ethical approval was sought and granted in all relevant countries (in some countries, e.g. Hungary, this type of research does not require ethical approval). Of particular concern was the fact that it cannot be assumed that high school students in the target groups are over 18 years of age, and that the questionnaire included questions about the students own questionable behaviour. The latter made it even more important to protect the privacy of the respondents, as well as ensuring and explaining the voluntary nature of the study. This was achieved through standard means of informed consent and anonymous questionnaire set-up.

Furthermore, data about individual institutions will not be made publicly available to eliminate the small risk that it, combined with the remaining demographic data, could enable reidentification of individual respondents.

The fact that not all high school students are over 18 meant that a choice had to be made in each country about whether to allow persons under the age of 18 to participate. In most countries, students under 18 were excluded from the questionnaire before the data were stored. However, in Switzerland, Ireland and Portugal it was decided that this would exclude too many of the target population, and they were therefore allowed to participate provided that parental consent had been given.

3.4. The questionnaire

In accordance with the timeline of WP2, the interviews were conducted, translated, and (preliminary) analysis was done during the period from January 2019 to June 2019. The questionnaire for the pilot survey was created and translated from June 2019 to September 2019. The pilot survey was conducted in October and November 2019, and the questionnaire was revised and the final version translated, set up on the internet platform and tested for technical issues in December 2019 and January 2020. The questionnaire used in this study was developed based on a literature study, an explorative qualitative study, and a pilot questionnaire. Questionnaire development began by formulating 15 general research questions based on the general aims of the study (see Preface):

Primary Research questions:

- RQ1: What are the most important differences in the self-reported understanding of good student/research practice across educational levels?
- RQ2: What are the most important differences in the self-reported understanding of good student/research practice across fields of study?
- RQ3: What are the most important differences in the adequacy of respondent understanding of good practice across educational levels?
- RQ4: What are the most important differences in the adequacy of respondent understanding of good practice across fields of study?
- RQ5: How does the perception of causes of deviations from good practice vary across levels of study?

- RQ6: How does the perception of causes of deviations from good practice vary across nationalities?
- RQ7: What are the most important differences in respondents' experiences of grey area issues across levels of study?
- RQ8: What are the most important differences in respondents' experiences of grey area issues across fields of study?
- RQ9: What are the most important national differences in respondents' experiences of grey area issues?

Secondary research questions:

- RQ10: What are the most important differences in perceived causes of deviations from good practice across levels of study?
- RQ11: What are the most important differences in perceived causes of deviations from good practice across nationalities?
- RQ12: How does the perception of causes of deviations from good practice vary across fields of study?
- RQ13: To what extent are there differences across levels and fields of study in the grey area issues that respondents have faced during their studies?
- RQ14: What are the most important differences in the conception of good practice across gender, nationality, amount of RI training, and type of RI training?
- RQ15: What are the most important differences in experiences of grey area issues across gender, amount of RI training, and type of RI training?

The questionnaire was not intended to answer all of these questions, as it was more appropriate to address some of them through the interview study.

The questions in the pilot questionnaire were developed based on these research questions, and took inspiration from earlier studies – particularly Martinson et al. (2005), Johansen & Christiansen (2020), Roig (1997), but also McCabe (1999), McCabe et al. (2001), Jensen et al. (2002), Fanelli (2009) and De Vries (2006). In addition, the results from the qualitative study informed the design of the questionnaire. In particular, the interviews gave us important information about the kinds of grey area issues that students commonly face, and the causes of deviations from good practice that they perceive (reported in Chapter 6). This input was very important, as the final survey contains no open text fields due to anonymity concerns and because open text fields would be extremely demanding to analyse given the

large number of respondents. It was therefore important that the lists of possible issues and causes that the respondents could select were realistic and relevant.

Drafts of the pilot questionnaire were circulated among all partners for feedback and to ensure that the questions were applicable and understandable across the different didactic traditions.

Once a final English version of the pilot questionnaire was available, it was translated into the languages of the three countries that participated in the pilot study: The Netherlands, Portugal and Lithuania. To ensure comparability, we used a translation/back-translation approach. For each language, the questionnaire was first translated into the intended language, then back-translated into English by an independent translator without access to the original. Any discrepancies between the original English version and the back-translation were discussed and adjustments to the translated versions made.

Two types of pilot tests were conducted. One in which participants were observed by a project member as they thought aloud while answering the questionnaire followed by a debriefing on what the participants found difficult, irrelevant or unclear, and one where the pilot survey was simply distributed electronically to participants. In total, 550 students from the three countries participated in the pilot tests.

The main issues identified in the pilot study were that the questionnaire was too long and the language too technical. The final questionnaire was therefore constructed from the pilot by taking out the least relevant questions, merging others, and generally simplifying the language. In addition, to enable more valid statistical analysis, we added a question about the institution with which the students were affiliated.

Once the questionnaire had been revised, it was translated into the languages not included in the pilot test, following the same translation/back-translation paradigm as described above. It was then set up on an internet survey platform so that participants in any country could choose between an English version of the questionnaire and a version translated into the national language of that country (for Switzerland, the second language was either French or German depending on what part of the country the participants came from). Native-speaking project members thoroughly and systematically tested each version of the electronic questionnaire before they were sent to respondents. The final questionnaire was ready and data collection commenced mid-January 2020.

3.5. The study data

A minimum sample size of n=160 was chosen for each study level (within each country) before inclusion in the present report. Only some of the countries reached this minimum sample size for all study levels (Denmark, Ireland, Portugal, Switzerland). In Hungary, Lithuania, the Netherlands, and Slovenia, the minimum sample size was only reached for some study levels. In Germany, the minimum requirement was not met for any of the study levels, following which Germany was excluded entirely from the results reported here. **Table 1** gives an overview of number of completed responses at each study level within the nine countries. The grey shaded cells highlight the segments that are included in the report.

Table 1. Number of completed questionnaire responses by country and study level.

	High School	Bachelor level	Research master's/ PhD	TOTAL
Denmark	389	218	427	1034
Germany	17	94	85	196
Hungary	60	292	221	573
Ireland	292	231	245	768
Lithuania	215	204	64	483
Netherlands	30	96	171	297
Portugal	219	274	241	734
Switzerland	436	210	261	907
Slovenia	250	221	48	519
TOTAL	1908	1840	1763	5511
TOTAL (only including the grey shaded cells)	1801	1650	1566	5017

In total, we received responses from 5,511 students at various study levels and from different countries (6,061 including the responses to the pilot survey). After excluding students from countries where the minimum sample size was not reached, our analysis included responses from 5,017 students.

Proceeding to demographic and study programme characteristics of the sample, Table 2 shows the distribution of three different gender identities among the 5,017 students included in the analyses. There was a higher proportion of female students across all study levels and countries (from 45.7% of high school students in Portugal to 69.7% of bachelor students in Ireland), followed by male students (from 22.9% of Swiss high school students to 41.5% of PhD students in Denmark). In most countries, a small

share (0% to 3.2%) stated that they were neither male nor female ("none of the above"). Quite a high proportion of respondents across many countries and study levels preferred not to answer the gender identity question (pooled proportion=5.8%). We are not able to determine the reasons for this response.

Table 2. Gender identity of the sample[a]

	Male	Female	None of the above	I prefer not to answer
DEN – High school	30.8%	60.4%	2.8%	5.9%
DEN – Bach level	32.1%	63.8%		4.1%
DEN – PhD level	41.5%	53.9%		4.7%
HUN – Bach level	27.1%	67.8%	1,0%	4.1%
HUN – PhD level	37.6%	52.9%	1.4%	8.1%
IRE – High school	37.7%	57.5%	1.4%	3.4%
IRE – Bach level	26.8%	69.7%	0.4%	3.0%
IRE – PhD level	34.7%	59.6%	0.8%	4.9%
LITH – High school	28.4%	58.1%	1.9%	11.6%
LITH – Bach level	17.6%	74.5%	2.5%	5.4%
NL – PhD level	36.8%	50.9%	2.3%	9.9%
POR – High school	35.6%	45.7%	3.2%	15.5%
POR – Bach level	27.0%	68.2%	0.4%	4.4%
POR – PhD level	36.1%	55.6%	0.8%	7.5%
SWI – High school	28.9%	61.2%	2.5%	7.3%
SWI – Bach level	22.9%	69,0%	2.9%	5.2%
SWI – PhD level	37.2%	59,0%	0.8%	3.1%
SLO – High school	28.8%	65.2%	1.6%	4.4%
SLO – Bach level	28.5%	68.8%	1.8%	0.9%
Total	31.7%	61.0%	1.5%	5.8%

[a]The question was: "Which gender do you primarily identify with?"

The age of the sample is quite varied and incudes very young students (<21 years), students in their 20s, and students 31 years or older (see Table 3). Unsurprisingly, high school students are youngest, followed by bachelor students, and finally PhD students.

Table 3. Age of the study sample

	15-17 years	18-20 years	21-25 years	26-30 years	31 years or more
DEN – High school	0%	97.9%	1.8%	0%	0.3%
DEN – Bach level	0%	7.3%	81.2%	6.9%	4.6%
DEN – PhD level	0%	0%	6.3%	52.5%	41.2%
HUN – Bach level	0%	25.0%	54.1%	5.8%	15.1%
HUN – PhD level	0%	0%	14.9%	49.3%	35.7%
IRE – High school	99.0%	1.00%	0%	0%	0%
IRE – Bach level	0%	51.5%	40.7%	2.6%	5.2%
IRE – PhD level	0%	0%	19.2%	31.4%	49.4%
LITH – High school	0%	100.0%	0%	0%	0%
LITH – Bach level	0%	37.7%	60.8%	0.5%	1.00%
NL – PhD level	0%	0%	50.9%	35.1%	14.00%
POR – High school	74.9%	24.7%	0.5%	0%	0%
POR – Bach level	0%	53.6%	35.8%	5.5%	5.1%
POR – PhD level	0%	0%	9.5%	27.4%	63.1%
SWI – High school	4.1%	91.1%	4.8%	0%	0%
SWI – Bach level	0%	26.7%	61.4%	5.7%	6.2%
SWI – PhD level	0%	0%	12.6%	54.4%	33.0%
SLO – High school	0%	99.6%	0.4%	0%	0%
SLO – Bach level	0%	26.2%	62.9%	5.9%	5.0%
Total	9.4%	36.8%	23.9%	15.1%	14.8%

At the bachelor and PhD study levels, the study direction (STEM, SOC SCI, HUM) and type of empirical data used (quantitative, qualitative, historical/arts artifacts) arguably constitute important distinctions that will shape the educational culture, and research culture. Our data collection plan aimed to ensure that all relevant directions and data types were included sufficiently.

In **Table 4**, we report the study direction and type of data used for all PhD students who completed the questionnaire. Looking at the totals from the table, it is clear that there is a high number of completed questionnaires across all study directions and data types used, except for law (where there were only 65 completed questionnaires). We can also see that there is a high number of completed responses in all common combinations of study direction and data types (e.g. Medical science and Quantitative data).

Table 4. Number of completed questionnaire responses at PhD level, stratified by study type and type of data used

	Quantitative data	Qualitative data	Historical sources or works of art/craft	Other type of data or no data used	Total
Natural science/engineering	396	27	7	78	508
Medical science	289	18	3	9	319
Social science	187	154	26	13	380
Humanities/arts	33	64	137	38	272
Law	3	10	12	40	65
Other	8	13	0	1	22
Total	916	286	185	179	1566

The study direction and type of data used for all bachelor students who completed the questionnaire are reported in **Table 5**. This table also reveals a high number of completed questionnaires across all study directions and data types used, except for law (where there only were 74 completed questionnaires). We can also see that there is a high number of completed responses in all common combinations of study direction and data types (e.g. Humanities/arts and Historical sources or works of art/craft).

Table 5. Number of completed questionnaire responses at the bachelor level, stratified by study type and type of data used

	Quantitative data	Qualitative data	Historical sources or works of art/craft	Other type of or no data used	Total
Natural science/engineering	331	11	51	90	483
Medical science	120	8	3	42	173
Social science	256	117	47	38	458
Humanities/arts	54	76	209	67	406
Law	12	1	22	39	74
Other	24	14	7	11	56
Total	797	227	339	287	1650

3.6. Analyses

The following chapters that report results are divided into thematic blocks of relevance. Across these chapters, the main results outline differences across study levels and countries. This is possible when students at all study levels (high school, bachelor, and PhD) receive similar questions. In some areas, we report mean or proportional differences across study levels and countries, while in other areas we supplement this by outlining study level and country differences in the relationship between different measures of interest: e.g. self-reported knowledge of questionable practice and instances of doubt about how to behave.

When students at different study levels received different questions, these are typically reported in separate figures.

Finally, at the PhD level, we also report differences between students from different study directions and students using different types of data (cf. Table 2). We do this in section 5.2.2. Country-specific differences are not included in this analysis. In Chapter 8, we also report associations between variables without distinguishing between countries.

In general, we report descriptive results (percentages within countries and study levels) or associations (either in box plots or as correlation coefficients) and avoid significance testing. In some places, however, results from significance testing are reported. For one section where we study gender identity differences (section 6.5), we conducted multivariable regression analysis to rule out possible confounding from age, country, study direction, and type of data used.

4. Self-reported understanding and practical challenges

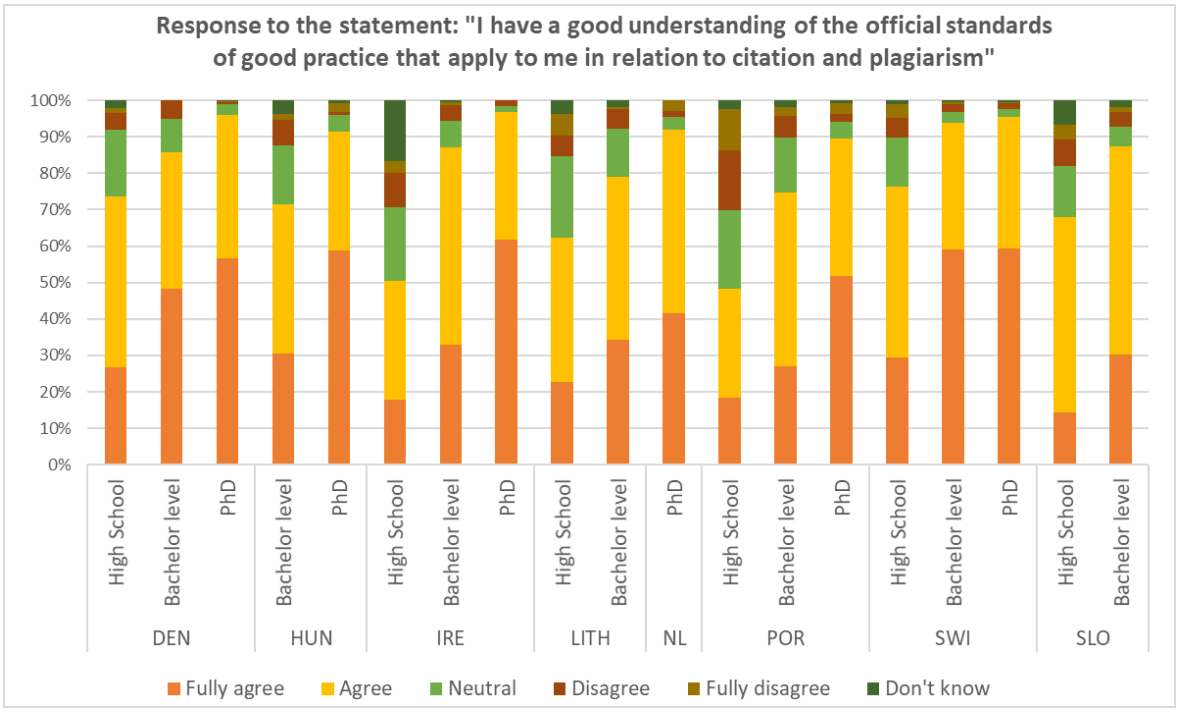
4.1. Students' understanding of what constitutes cheating and misconduct

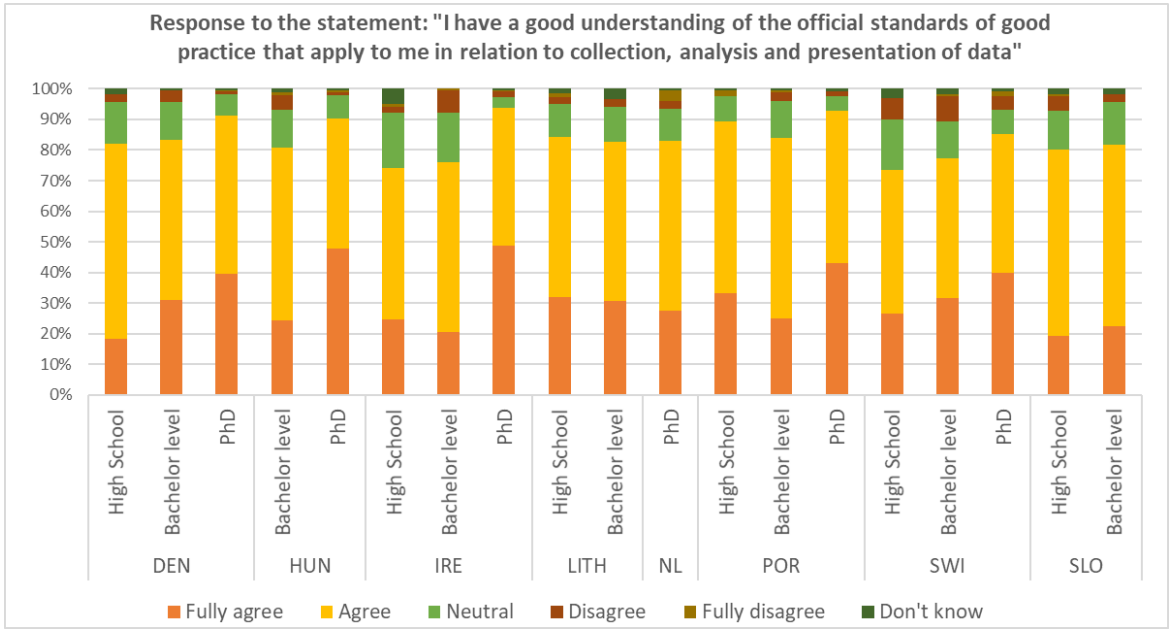
The questionnaire included three questions about the students' self-reported understanding of the official standards of good practice that apply to them regarding our three main themes: 1) *drawing on the work of others*, 2) *collaboration and assigning authorship*, and 3) *collecting, interpreting and presenting data* (including data handling). The term "official standards" was not defined in the questionnaire, but exemplified as follows: "Examples of "official standards of good practice" could be rules and regulations stating what is prohibited, but it could also include guidelines and codes of conduct describing how to behave correctly." The term was thus used as a more general term than "rules and regulations", as it covers both rules and regulations that are enforced, as well as other standards such as recommendations and codes of conduct that may not be enforced very strictly. In three further questions, the questionnaire asked whether the students knew how to behave in an ethically correct manner regarding the same three topics.

For all six questions, students were asked to respond on a gradient scale ranging from "Fully disagree" to "Fully agree" (there was also an "I don't know" option). **Figure 1** reports detailed frequencies from the questions probing the students' self-reported knowledge regarding the three themes, i.e. whether the students had "a good understanding of the official standards of good practice" that applied to them in relation to the three themes. The students generally reported high levels of knowledge, such that a clear majority agreed (i.e. "fully agree" or "agree") with this statement. For many of the reported study levels, 90% or more of the students agreed. There are, however, some noticeable exceptions. In particular, high school students in Ireland and Portugal reported relatively low levels of knowledge regarding citation and plagiarism (only about half agreed or fully agreed that they have a good understanding of the official standards). As the pattern is not repeated in the answers given by the Irish and Portuguese students for the other two themes, these differences are difficult to explain.

Furthermore, it is worth noting that there is a clear gradient following study level (such that PhD students declared better knowledge than bachelor students, and bachelor students declared better knowledge than high school students) in terms of citation and plagiarism, but not the two other general themes. However, self-reported knowledge may in part be based on false beliefs. The self-reported knowledge thus partly contradicts the findings of the qualitative interviews (reported in D2.1), in which high school students showed little or no understanding of the rules pertaining to data handling and analysis. A possible explanation for this discrepancy and the pattern observed in Figure 1 may be that the rules concerning citation and plagiarism are very clearly enforced at high school and bachelor levels, whereas the rules concerning authorship and data are not as clearly enforced – if at all. Consequently, the high school and bachelor students' lack of knowledge about the rules on plagiarism may be perceived as a graver deficit – and something that they are therefore more aware of – than a lack of knowledge about the rules on authorship and data collection. This analysis is consistent with the students' preferences concerning what topics to include in teaching on academic integrity (see Chapter 7).

The three behaviour questions, which asked whether the students "know how to behave in an ethically correct manner in relation to" the three main themes show a similar pattern to the questions concerned with self-reported knowledge. For this reason, we do not to report details of the responses to these three questions.


Response to the statement: "I have a good understanding of the official standards of good practice that apply to me in relation to citation and plagiarism"
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
High School
Bachelor level
PhD
DEN
HUN
IRE
LITH
NL
POR
SWI
SLO
Fully agree
Agree
Neutral
Disagree
Fully disagree
Don't know


Response to the statement: "I have a good understanding of the official standards of good practice that apply to me in relation to collection, analysis and presentation of data"
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
High School
Bachelor level
PhD
DEN
HUN
IRE
LITH
NL
POR
SWI
SLO
Fully agree
Agree
Neutral
Disagree
Fully disagree
Don't know

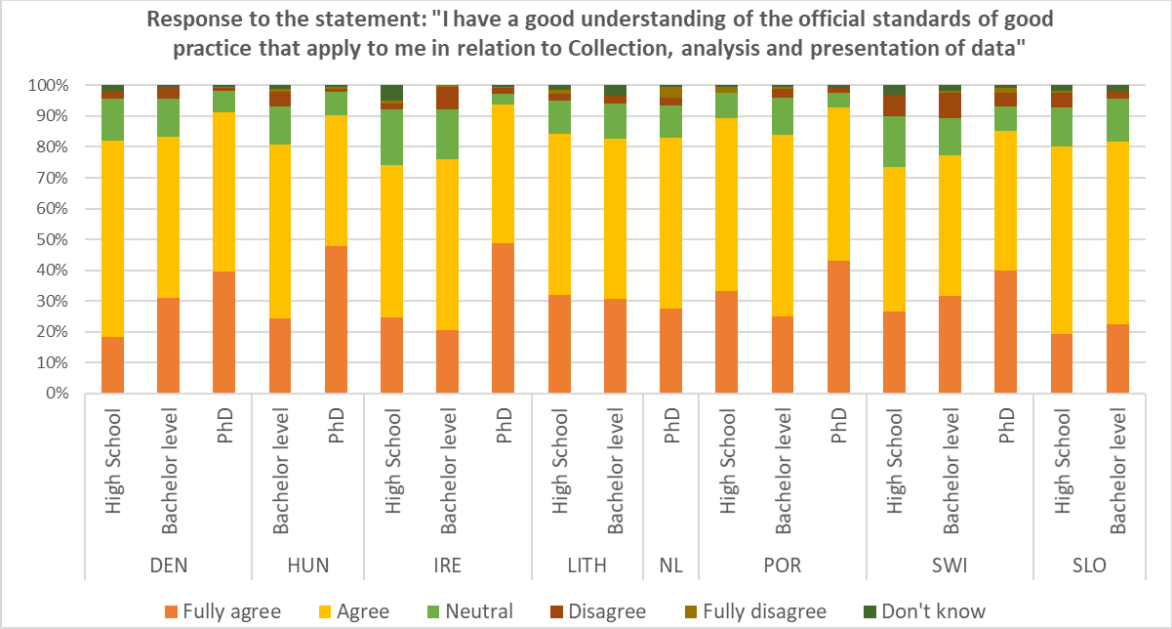


Figure 1. Self-reported understanding of what is good practice in relation to our three themes (Stacked bars; responses sum to 100% for each bar), stratified by country and study level.

To supplement the theme-specific results, it is also possible to construct a composite measure where responses to all six questions on self-reported knowledge and behaviour are summed together to construct an index indicating *self-reported knowledge about rules and behaviour of academic integrity*. This decision is statistically justified by the fact that the measure has very good internal validity with Cronbach's alpha, ranging between 0.79 and 0.87 in the six countries (pooled Cronbach's alpha 0.83). The index ranges from 0 to 1, where 0 indicates a very low, 1 indicates a very high, and 0.5 indicates a medium level of self-reported knowledge. This type of index has the advantage that it reduces multiple single measures into one measure that conveys most of the common variation in the single items. The obvious disadvantage is that we cannot identify particular themes where students across the educational levels differ in their belief that they have limited knowledge (e.g. the differences observed in the answers from Irish and Portuguese high school students to the first and the last two themes as seen in the top and the bottom two graphs in Figure 1).

We report country- and study-level-specific variation on this index in the form of a box plot (**Figure 2**). This shows that the majority of students across all countries and study levels had high self-reported knowledge, with median scores of 0.70 (Portuguese high school students) or higher. There are some country-specific differences, and while these differences are statistically significant (at the 0.05 level), they are arguably of minor practical significance as the average student across all countries believed

themselves to be quite knowledgeable regarding official rules and how to behave. Still, it should be noted that a minority of students in some countries and at certain study levels did feel less confident in their knowledge. This is primarily seen among students in high school or bachelor level programmes. Most strikingly, among high school students in Portugal, the bottom 25% of students have a score of 0.58 or lower on the index, and an almost equally low level was found for the bottom 25% of high school students in Ireland and Slovenia. Similarly, although PhD students generally scored very high with medians well above 0.8 on the index, PhD students from the Netherlands apparently felt less confident, reaching a median of only 0.75.

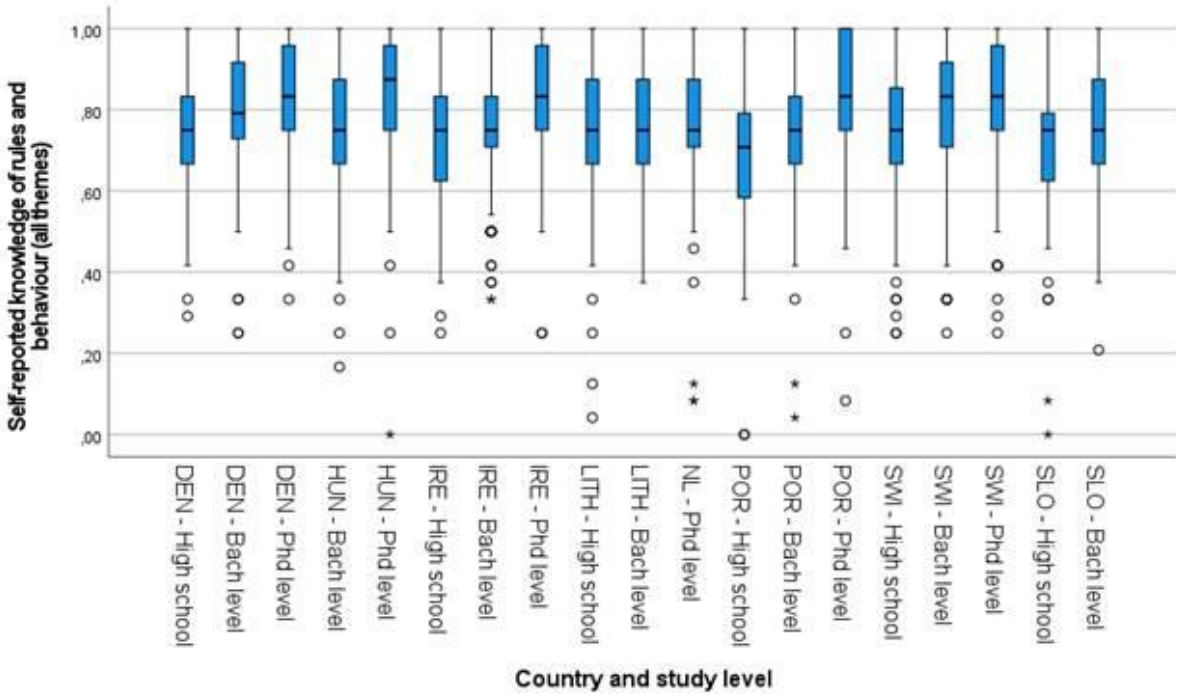


Figure 2. Scores on the index *self-reported knowledge about rules and behaviour of academic misconduct* (Box plot – where the top and bottom of the boxes are 1st and 3rd quartiles and the lines in the middle of the boxes are medians), stratified by country and study level.

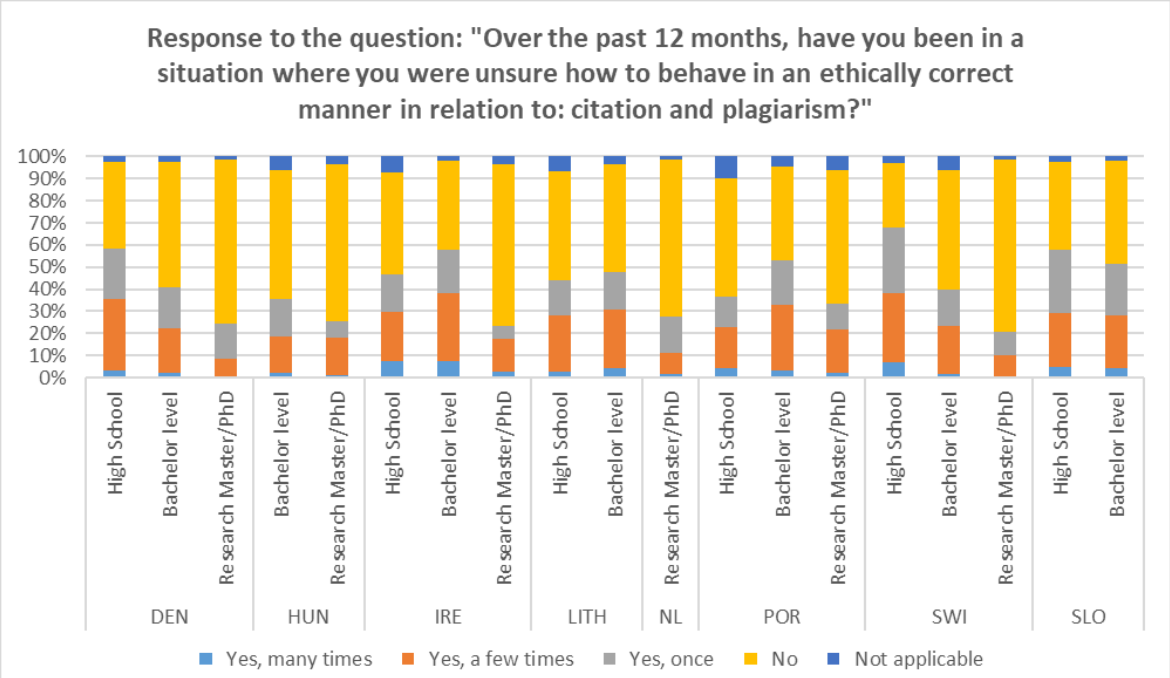
4.2. Are students confident about how to behave in practical situations?

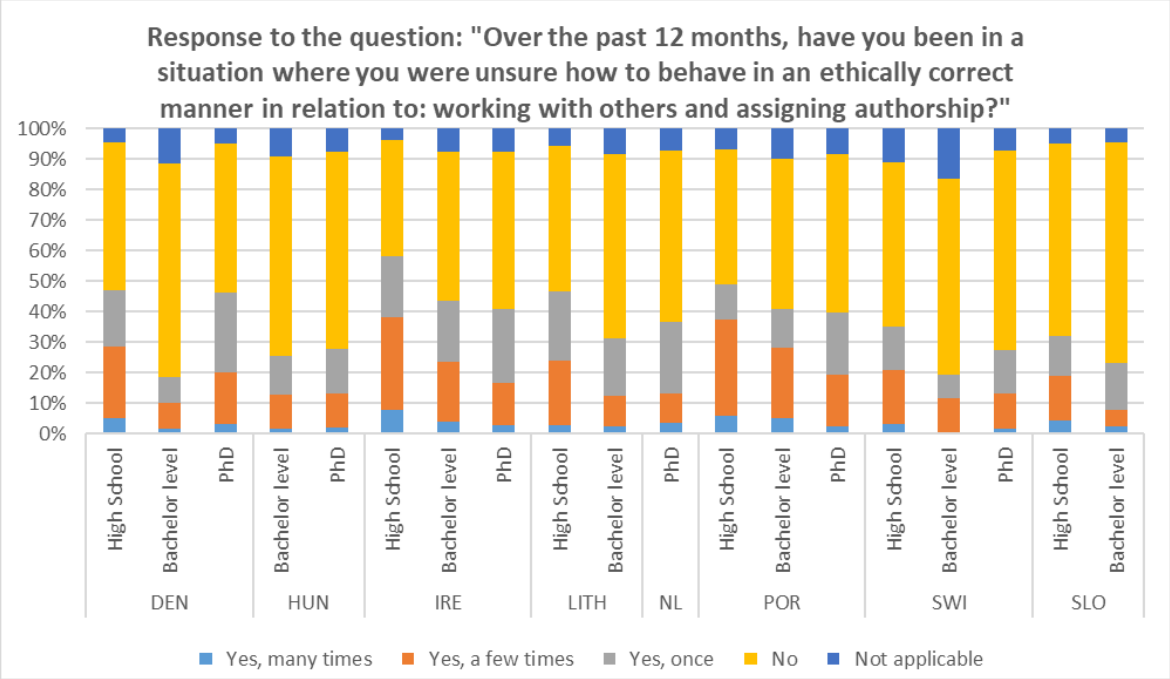
As laid out in section 4.1, most students felt that they had a good understanding of the official standards and rules of good behaviour when asked about them in general terms. To come a step closer to practical reality, we also asked respondents whether they had been in a situation where they "were unsure how to behave in an ethically correct manner in relation to [the three general themes (citation and plagiarism, authorship and data collection and analysis)]" over the past 12 months. Students were asked to respond on a scale ranging from "Yes, many times", "Yes, a few times", and "Yes, once" to "No" (there was also an "Not applicable" option). **Figure 3** reports detailed frequencies across all countries, study levels and ethical themes.

Interestingly, many students, regardless of their general confidence in their own knowledge, often found themselves in situations where they were in doubt about how to behave correctly in practice – and this was the case across all countries and study levels, even though the share of students who had doubts varied considerably across study levels and the three general topics considered.

The top graph of Figure 3 shows that a considerable share of students across all study levels had been in at least one situation where they "were unsure how to behave in an ethically correct manner in relation to *citation and plagiarism*". Although this doubt exists across all study levels, there is a strong gradient, as PhD students reported doubt less frequently (21-36% across countries experienced this once or more frequently; $Pooled_{Countries}$: 26%) than bachelor students (38-60% across countries; $Pooled_{Countries}$: 48%) and high school students (41-70% across countries; $Pooled_{Countries}$: 57%).

The second graph of Figure 3 also shows that many students across study levels and countries were at least on one occasion (between 21% and 60%; $Pooled_{CountriesxStudylevels}$: 40%) unsure about how to behave correctly with respect to "*…working with others and assigning authorship*". Here, the gradient reported above is less clear; in Ireland and Portugal the trend holds, but in Denmark and Switzerland, bachelor students reported less doubt than PhD students, while high school students still reported the highest level of doubt. Across countries, 27-46% of PhD students ($Pooled_{Countries}$: 40%) reported being in doubt at least once, while 19-44% of the bachelor students ($Pooled_{Countries}$: 32%) and 32-58% of the high school students ($Pooled_{Countries}$: 47%) reported being in doubt. However, as we will show in section 4.3, there is a qualitative difference between the doubts of PhD students and bachelor and high school students when it comes to collaboration and authorship.


Response to the question: "Over the past 12 months, have you been in a situation where you were unsure how to behave in an ethically correct manner in relation to: citation and plagiarism?"
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
High School
Bachelor level
Research Master/PhD
DEN
HUN
IRE
LITH
NL
POR
SWI
SLO
Yes, many times
Yes, a few times
Yes, once
No
Not applicable


Response to the question: "Over the past 12 months, have you been in a situation where you were unsure how to behave in an ethically correct manner in relation to: working with others and assigning authorship?"
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
High School
Bachelor level
PhD
DEN
HUN
IRE
LITH
NL
POR
SWI
SLO
Yes, many times
Yes, a few times
Yes, once
No
Not applicable

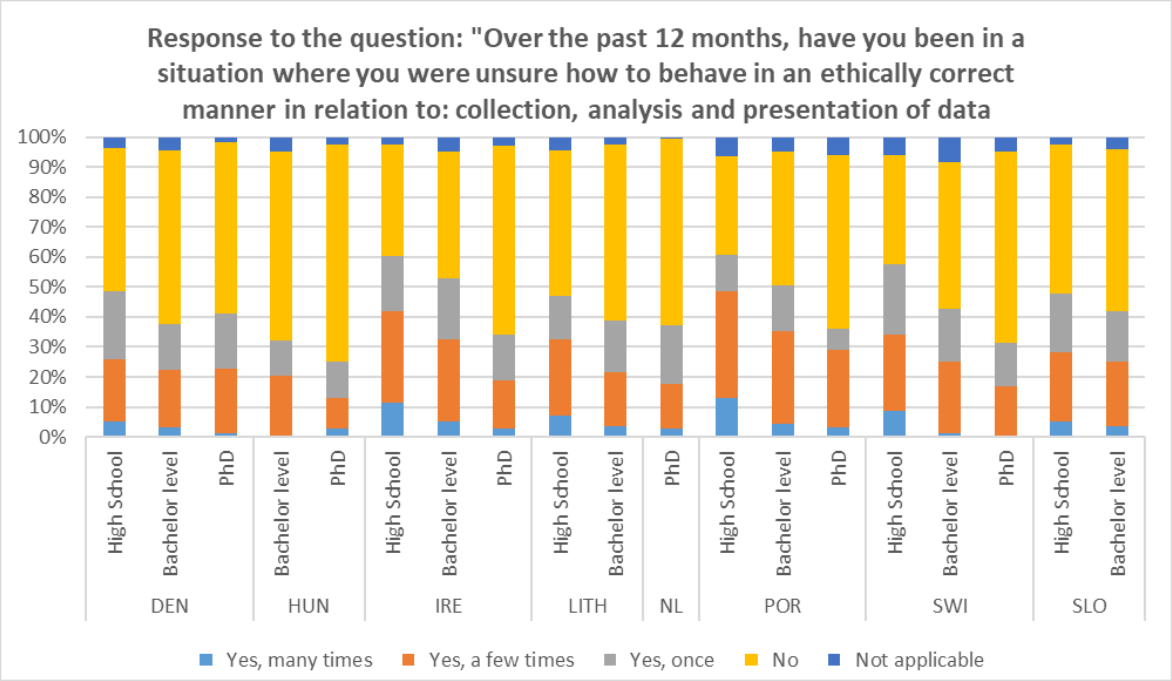


Figure 3. Frequency with which students were in doubt about how to behave in an ethically correct manner over the past 12 months for three themes (Stacked bars; responses sum to 100% for each bar), stratified by country and study level.

Finally, the bottom graph in **Figure 3** shows that quite a large share of students across all study levels and countries were unsure about how to behave ethically with respect to "*…collection, analysis, and presentation of data*" over the past 12 months. This doubt occurred at least once for 26-42% of PhD students across countries ($Pooled_{Countries}$: 36%), 34-56% of bachelor students across countries ($Pooled_{Countries}$: 47%), and 49-65% of high school students across countries ($Pooled_{Countries}$: 56%). Once again, we see a clear downward gradient with PhD students reporting less doubt than students from the other two levels. The only exception is Denmark, where slightly more PhD students (41%) than bachelor students (38%) reported to have been in doubt at least once.

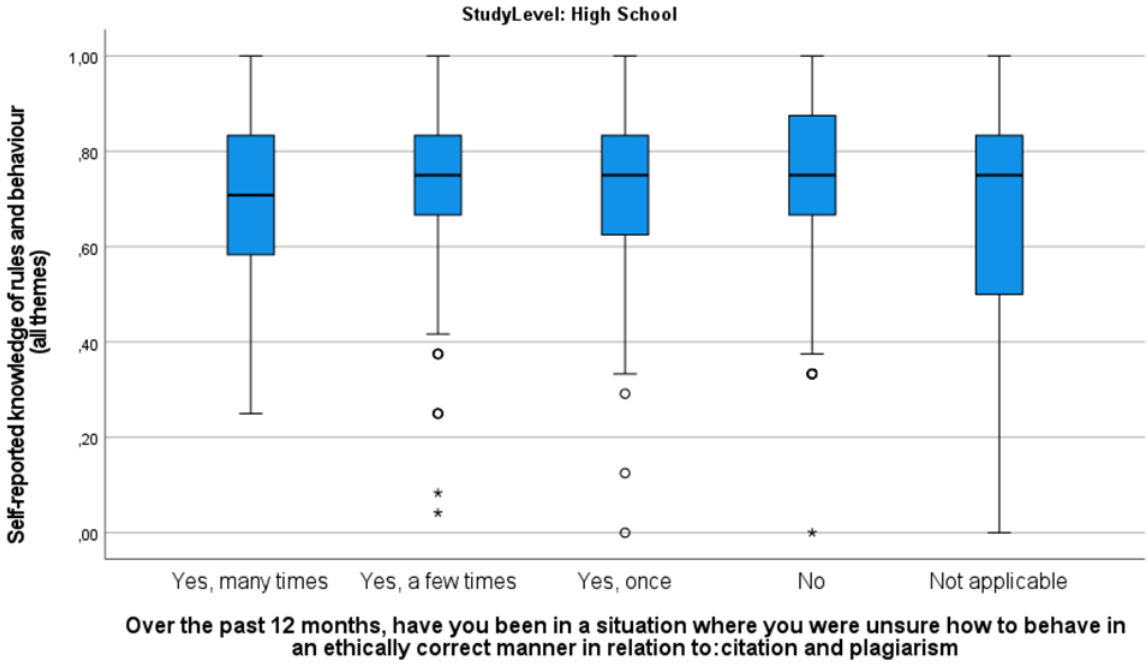
4.3. Possible connection between self-reported knowledge and experienced doubts

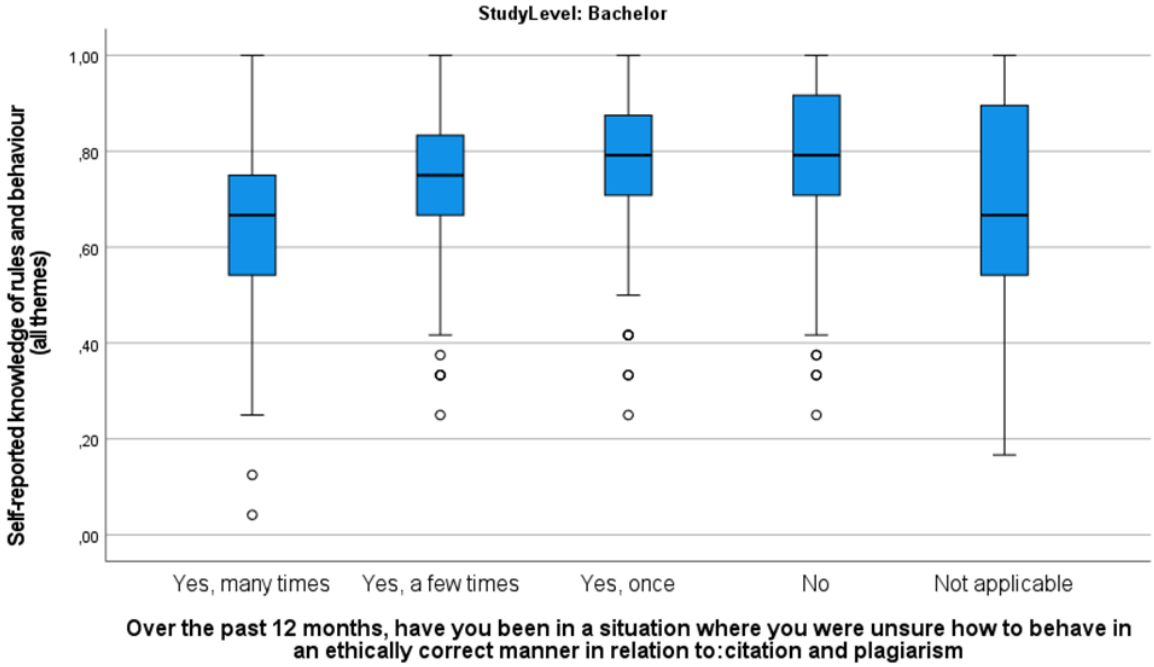
It is tempting to believe that individuals who are more knowledgeable would be better equipped to handle issues connected to data handling, plagiarism, and authorship correctly. To evaluate this hypothesis, we analysed the association between *self-reported knowledge about rules and behaviour of*

academic misconduct and *doubt as to how to behave in an ethically correct manner*. We did find a statistically significant association between self-reported knowledge and doubt about how to behave for all three topics. We illustrate this in **Figure 4**, using the case of *citation and plagiarism* (a roughly similar pattern would have emerged if illustrating the association using one of the other two integrity topics) and stratifying by study level.

The three graphs show a gradient relationship where higher self-reported knowledge is associated with fewer situations of doubt. This association is strongest among bachelor (Spearman's r=0.192) and PhD students (Spearman's r=0.214), and more limited among high school students (Spearman's r=0.081). While the association is low to modest (correlation coefficients below <0.300), it does appear to increase in tandem with study progression. This can be seen as an indication that students develop a more reflective understanding of their own knowledge of academic integrity as they progress, or that students who have a more reflective attitude are also more likely to continue higher education beyond high school.

Furthermore, it is important to mention that self-reported knowledge is quite high, irrespective of the number of situations where the student was in doubt. Among high school students, even those who reported being in doubt "many times" have a median score over 0.7 on the knowledge index, and among bachelor and PhD students, the same median score is a little under 0.7.


StudyLevel: High School
Self-reported knowledge of rules and behaviour (all themes)
1,00
,80
,60
,40
,20
,00
Yes, many times
Yes, a few times
Yes, once
No
Not applicable
Over the past 12 months, have you been in a situation where you were unsure how to behave in an ethically correct manner in relation to:citation and plagiarism


StudyLevel: Bachelor
Self-reported knowledge of rules and behaviour (all themes)
1,00
,80
,60
,40
,20
,00
Yes, many times
Yes, a few times
Yes, once
No
Not applicable
Over the past 12 months, have you been in a situation where you were unsure how to behave in an ethically correct manner in relation to:citation and plagiarism

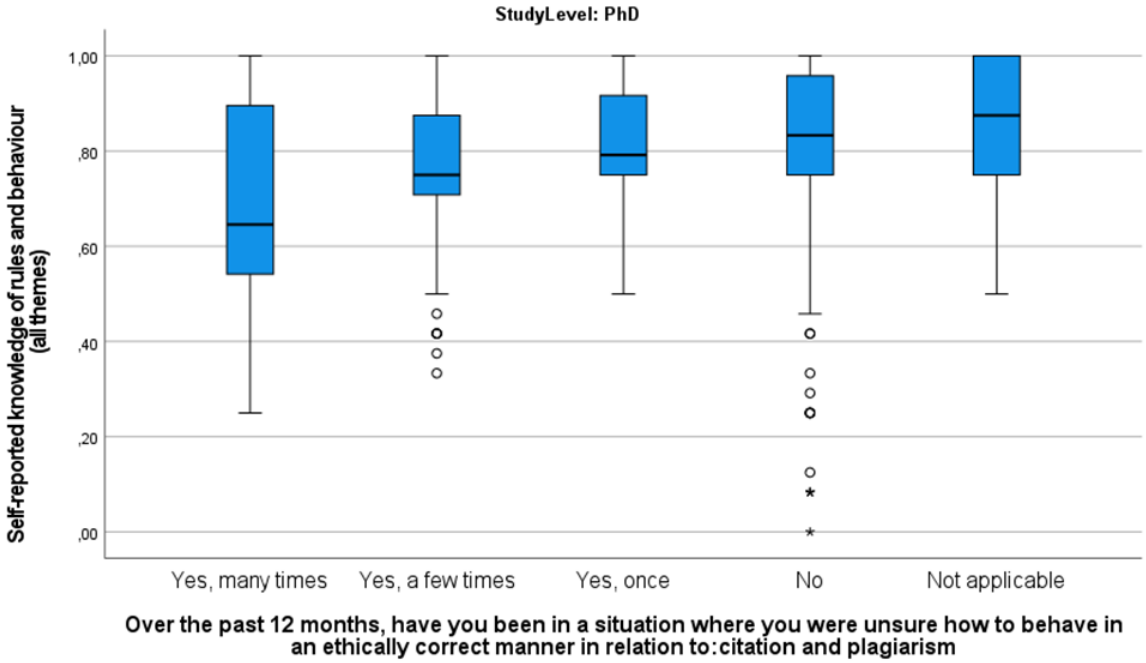


Figure 4. Scores on the index *self-reported knowledge about rules and behaviour relating to academic misconduct* (Box plot – where the top and bottom of the boxes are 1st and 3rd quartiles and the lines in the middle of the boxes are medians), stratified by the number of times the student reports being unsure about how to behave ethically in relation to citation and plagiarism across study levels.

5.Perceptions of rules and good practice

This chapter focuses on student perceptions of rules and regulations as well as good practice with respect to our three themes. The term "rules and regulations" was not defined in the questionnaire, but unlike "standards of good practice", it was used in contexts where it was clear that they could be violated and that violations could come at a cost. Some of the results do not cover all study levels at the same time because the issues that students face, particularly those focusing on authorship issues, will vary across levels.

5.1. Perception of rules about drawing on the work of others

This section includes perceptions about what constitutes acceptable citation practice in general terms (section 5.1.1) and in relation to an actual case (section 5.1.2).

5.1.1 Perceived rules in general terms

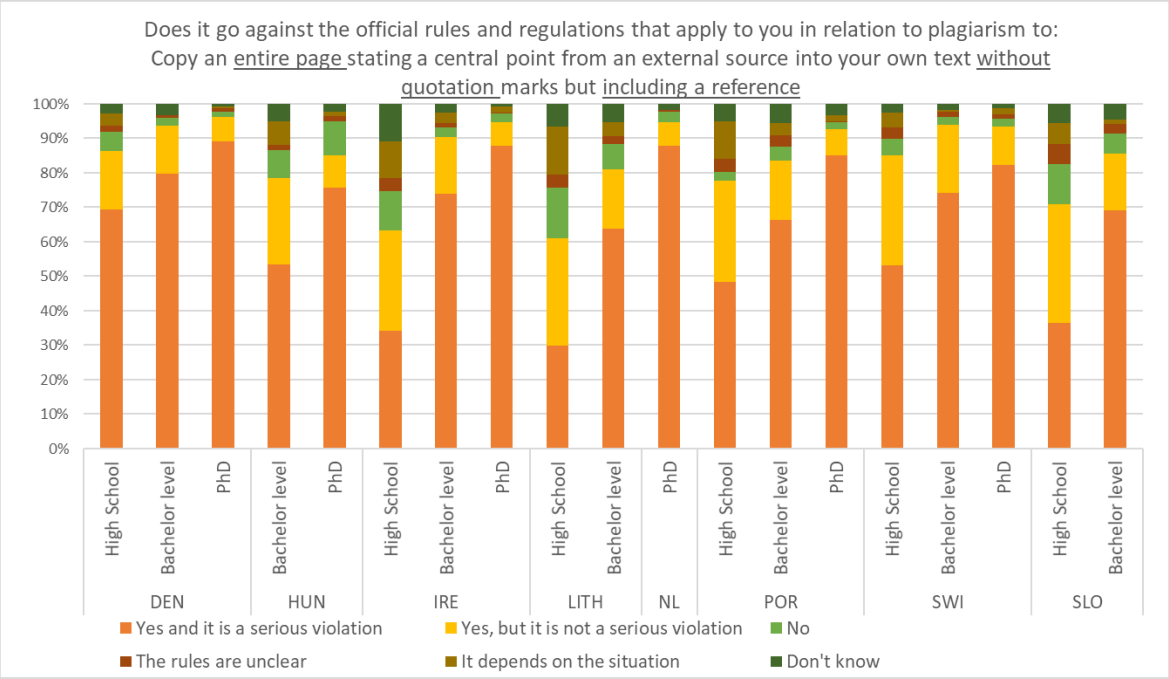
In **Figure 5**, comprised of the 4 graphs below, we report results from student perceptions of what constitutes a violation of rules that apply to them with respect to plagiarism. We asked the respondents to evaluate the seriousness of four cases in which an original source was used in one's own text with different degrees and types of text copying. Response options were: "Yes, it is a serious violation", "Yes, but it is not a serious violation", "No, it is not against the rules", "The rules are unclear", "It depends on the situation", and "I don't know".

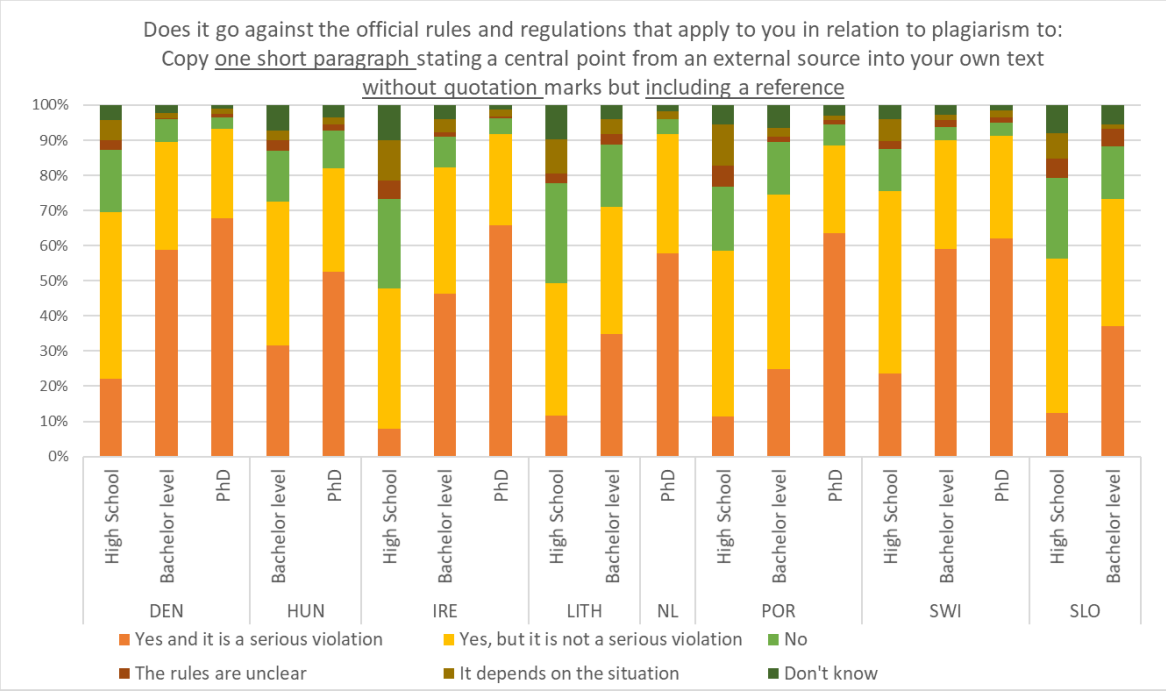
In general, when comparing the four cases, we note that there is a clear gradient in the perception of possible violations of the rules, such that the first case is considered to be a more serious violation than the second and so on. There are also interesting differences in the evaluation of cases across study levels and countries as we will note below.

In the first case (**Figure 5**, first graph), students were asked to evaluate whether they believed it to be a violation of rules to "Copy an entire page stating a central point from an external source into your own text without quotation marks but including a reference" (emphasis added). Unsurprisingly, a very large share of students across all study levels and countries found this to be a "serious violation". In the four countries where all study levels were included (Denmark, Portugal, and most notably, Ireland and Switzerland), a clear gradient can be observed where the case is perceived to be a more serious violation

at higher study levels. It is, however, also worth noting that just over 60% of the high school students in Ireland (63%) and Lithuania (61%) saw it as a violation (either serious or not serious).

The second scenario, to "Copy one short paragraph stating a central point from an external source into your own text without quotation marks but including a reference" (**Figure 5,** second graph) was also perceived as a case of plagiarism by most students. In general, however, it was deemed to be less serious than the first case. However, the majority of PhD students reported it to be a "serious violation", while high school students generally believed it to be "not serious", and less than 50% of high school students in Ireland and Lithuania actually saw it as a violation at all. Similar to the first case (but to a lesser extent), for the four countries where all study levels are included (Denmark, Portugal, Ireland, and Switzerland), a gradient can be observed where the case was perceived as a more serious violation at higher study levels.




Does it go against the official rules and regulations that apply to you in relation to plagiarism to:
Copy one short paragraph stating a central point from an external source into your own text without quotation marks but including a reference
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
High School
Bachelor level
PhD
DEN
HUN
IRE
LITH
NL
POR
SWI
SLO
Yes and it is a serious violation
Yes, but it is not a serious violation
No
The rules are unclear
It depends on the situation
Don't know

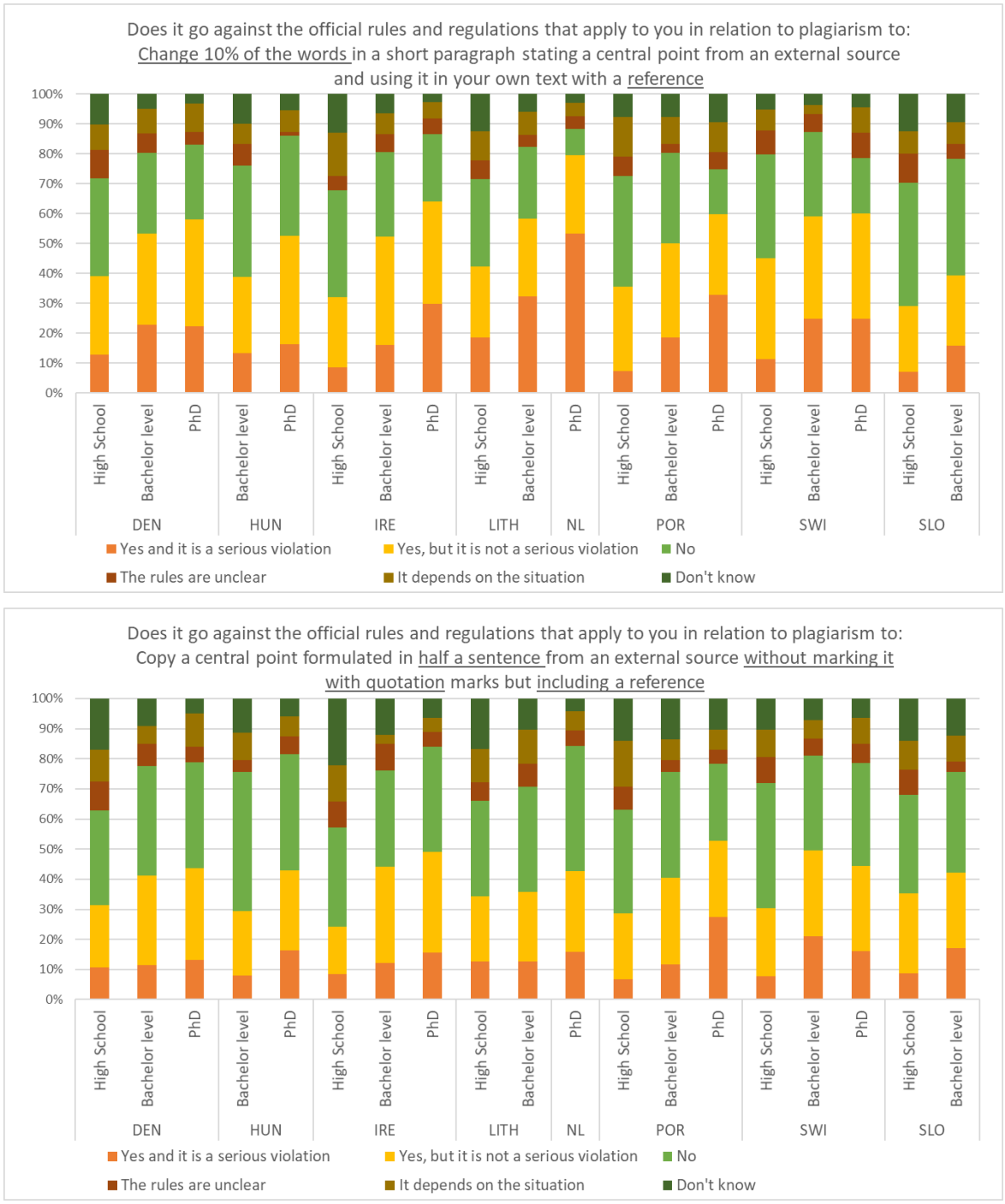


Figure 5. Perceptions about what constitutes a violation of rules and regulations regarding plagiarism (Stacked bars; responses sum to 100% for each bar), stratified by country and study level.

Proceeding to the third case, "Changing 10% of the words in a short paragraph stating a central point from an external source and using it in your own text with a reference" (**Figure 5,** third graph), we can see that a larger proportion of students did not report this as a violation at all (i.e. they respond "no"). Approximately one third of students across most countries responded "no", one third responded that it is a violation "but not [a] serious" one, and one third believed it to be a "serious violation". PhD students in the Netherlands stand out here, as the case is regarded as much more serious by this group. In the four countries where all study levels are included (Denmark, Portugal, Ireland, and Switzerland), a gradient similar to the first two cases can be observed, such that the case is generally perceived as a more serious violation at higher study levels. It should also be noted that a relatively large share of students could not give a clear answer to this case, and responded either "The rules are unclear", "It depends on the situation", or "I don't know". This share is highest among high school students (across countries).

The fourth case (**Figure 5,** last graph), "Copying a central point formulated in half a sentence from an external source without marking it with quotation marks but including a reference", was perceived as unproblematic by an even greater proportion of students, as 30% to 45% across study levels and countries responded "no". Approximately one in four students across study levels and countries responded "yes, but not a serious violation". Finally, at most study levels, 10% to 20% found the case to be a "serious violation". Portuguese PhD students deviate from the trend here, as approx. 30% perceived it as a "serious violation". Similar to the third case, a relatively large share of students could not give a clear answer, and responded either "The rules are unclear", "It depends on the situation", or "I don't know". This share is highest among high school students (across countries).

5.1.2 Perceived rules in a practical case

In addition to the questions discussed in section 5.1.1, respondents were also presented with four versions of a scenario, where a student paraphrased the following passage to different degrees:

> Process Y was first discovered in 1931 by the German scientist H. Neumann, and was refined in the late 50s. Its full importance was, however, only fully understood when it was described in the seminal paper by D. Dirksen in 1971.

The different versions were paraphrased in the following way:

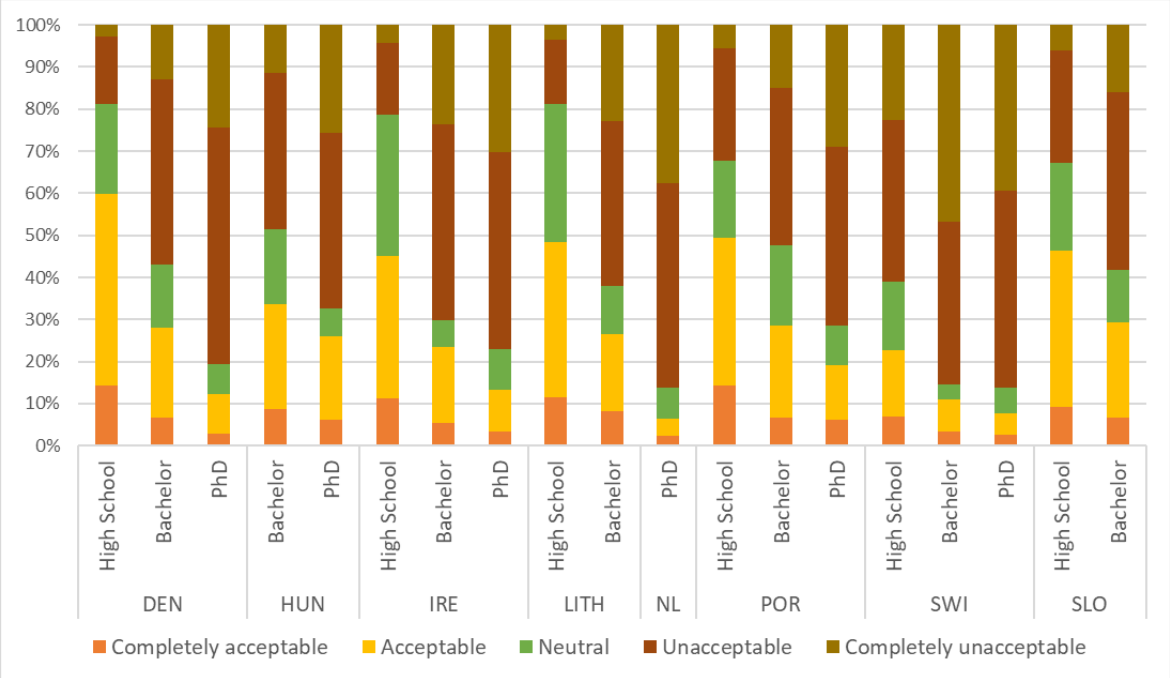
- Paraphrase 1: a direct copy without quotation marks and no reference.
- Paraphrase 2: a few insignificant words had been deleted or changed to synonyms. There was no reference to the original.
- Paraphrase 3: identical to paraphrase 2, except that a reference to the original had been added.
- Paraphrase 4: a more substantial rewriting with a reference to the original.

For each version, the students were asked to rate the behaviour on a five-point scale: "Completely acceptable", "Acceptable", "Neutral", "Unacceptable", "Completely unacceptable", and "I don't know".

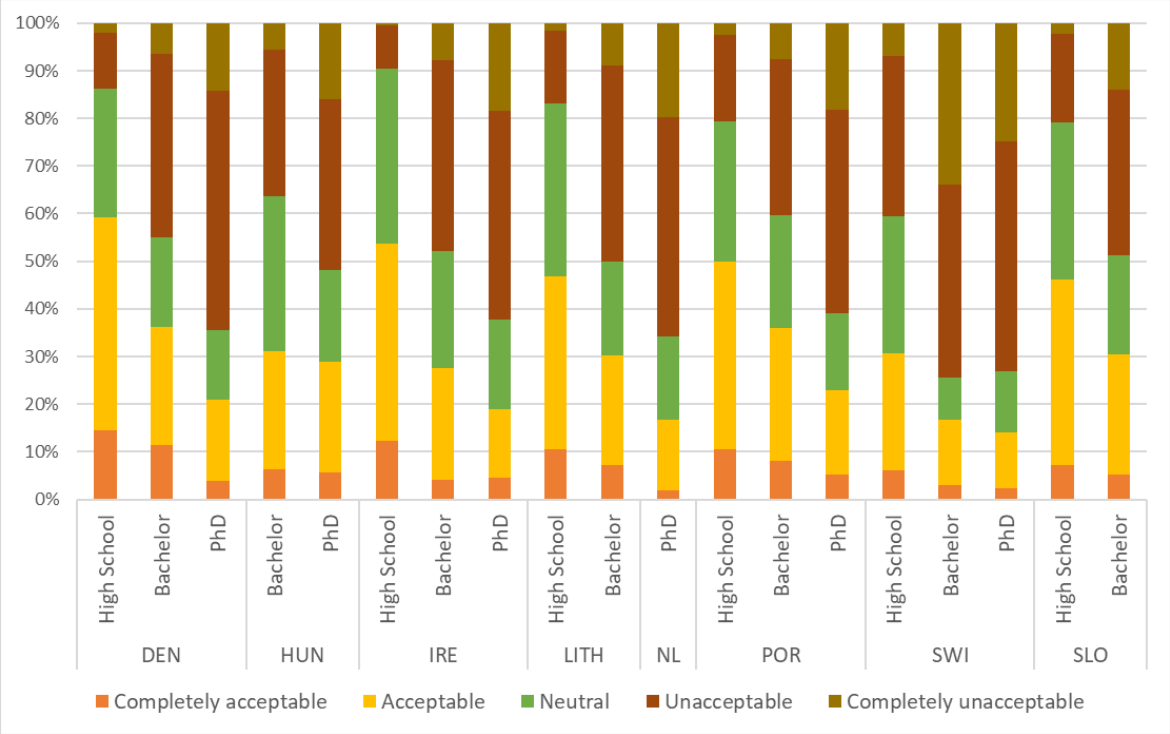
According to the European Code of Conduct, plagiarism "is using other people's work and ideas without giving proper credit to the original source, thus violating the rights of the original author(s) to their intellectual outputs" (ALLEA 2017, p. 8). From this perspective, paraphrase 1 and 2 are unacceptable, whereas paraphrase 3 is perhaps more of a grey zone. Paraphrase 4 was designed in to be reasonably acceptable, at least in the authors' view.

The students' views on the paraphrases are shown in **Figure 6** for different countries and study levels. The response "I don't know" was omitted from the graphs in the figure, as a limited number of students (<10% in all cases) used the option.

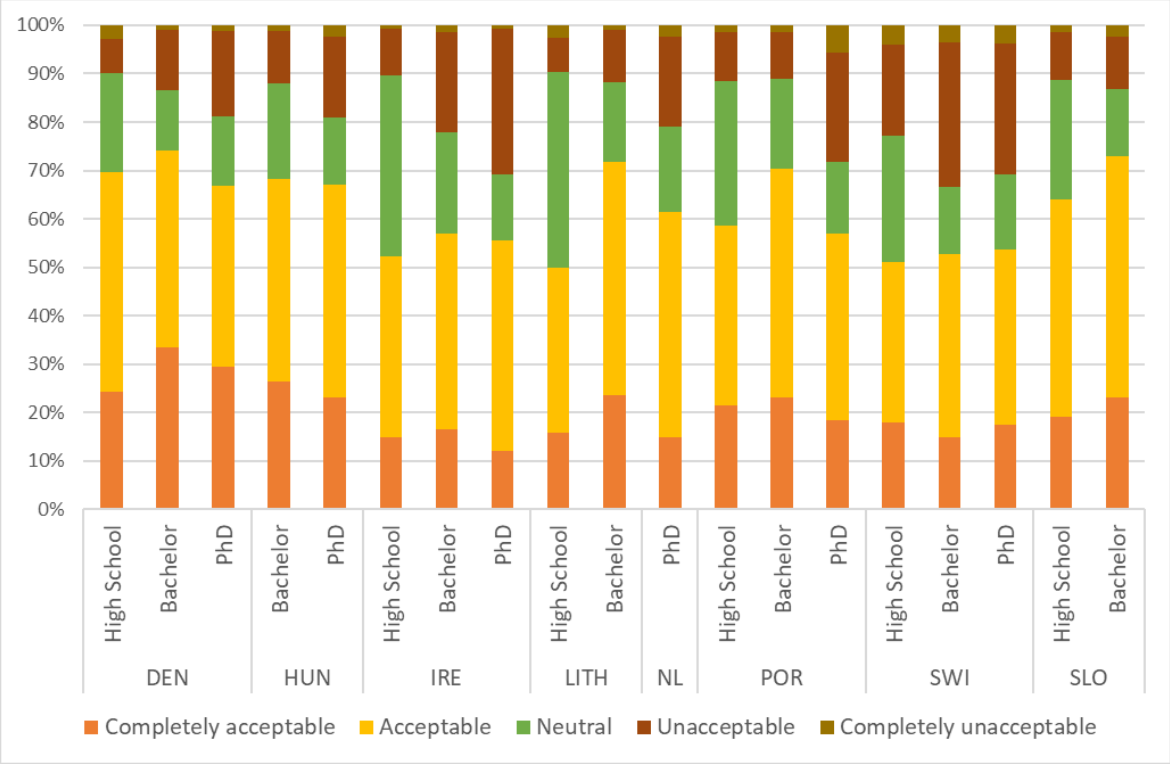
Paraphrase 1: Direct copy, no reference



Paraphrase 2: Insignificant changes, no reference



Paraphrase 3: Identical to paraphrase 2, but with a reference



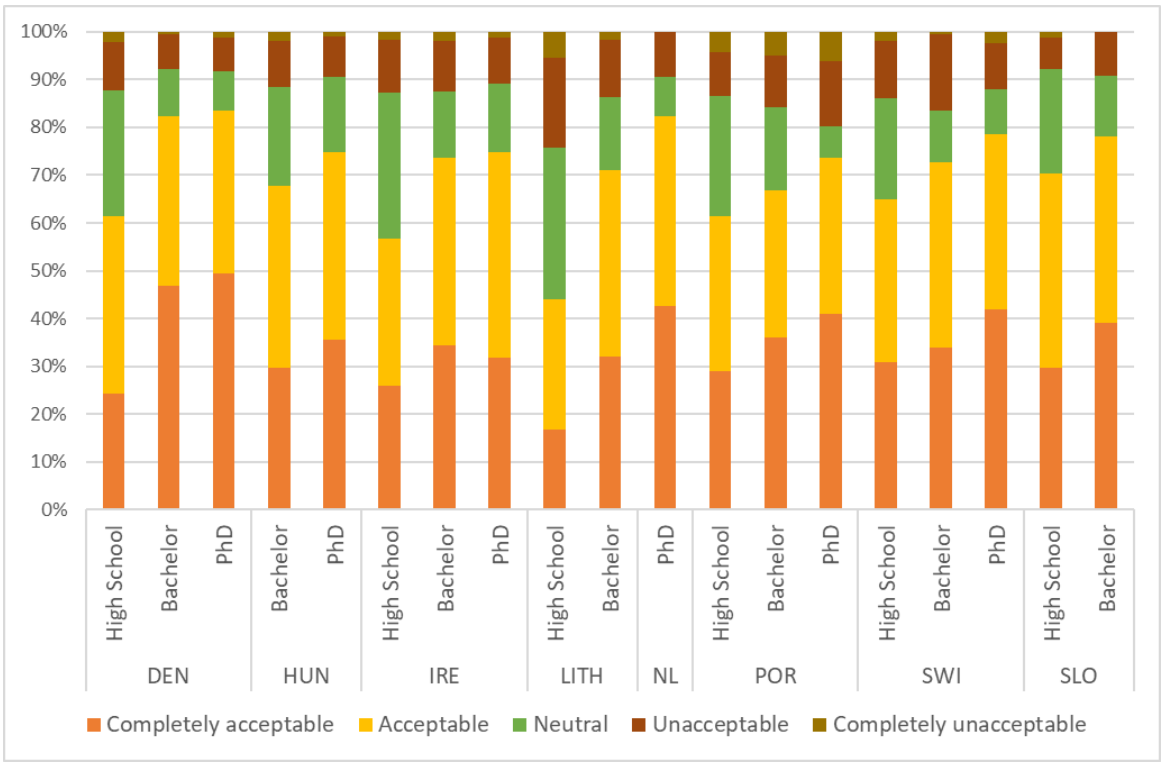


Figure 6. Perceptions about plagiarism – four practical situations (Stacked bars; responses sum to 100% for each bar), stratified by country and study level.

The first important observation is that there is little difference in the answers given by high school students across the four cases, with Switzerland as the only exception (disregarding Switzerland, 40-55% of the high school students across countries answered "Acceptable" or "Completely acceptable" to Paraphrase 1, 40-55% to Paraphrase 2, 40-60% to Paraphrase 3, and 40-65% to Paraphrase 4). This could indicate that the high school students generally did not perceive any real difference in the acceptability of the four paraphrases, but it might also be seen as an indication that the high school students did not understand the scenarios they were presented with. At any rate, the results indicate that although high school students may have – or may believe that they have – a good grasp of the rules in theory, they have real difficulty in applying these to concrete cases. This result is consistent with the results of the qualitative interviews, in which high school students generally found it difficult to handle scenarios and concrete situations.

To our minds, Paraphrase 1 is a clear instance of unacceptable practice. Figure 6 shows that this is not necessarily so for the students. While the majority of PhD students marked Paraphrase 1 as either "Completely unacceptable" or "Unacceptable", there was a minority of more than 10% (23% in Ireland)

who did not find this paraphrase unacceptable. Major national differences can be observed at bachelor level. In Switzerland, 87% of bachelor students found Paraphrase 1 either unacceptable or completely unacceptable, which is similar to the Swiss PhD students. In Denmark, Hungary and Portugal on the other hand, there was a very substantial minority of bachelor students (45%, 52%, and 51%, respectively) who did not find Paraphrase 1 unacceptable. In Denmark, Portugal and to a lesser extent Ireland, we also see a substantial shift from bachelor to PhD level towards considering Paraphrase 1 unacceptable. This generates a gradient where the case is perceived as more unacceptable at higher study levels.

Paraphrase 2 is no less problematic than Paraphrase 1. From the students' perspective, we see that there is a slight shift towards seeing Paraphrase 2 as more acceptable than Paraphrase 1 (McNemar test, $p<0.001$), indicating that some students find that changing insignificant words can make paraphrasing more acceptable. However, the difference is very modest, with 27% of students finding the first paraphrase acceptable and 31% finding the second paraphrase acceptable.

We see a major shift towards acceptability from Paraphrase 2 to Paraphrase 3, and even though the authors still view it as problematic, it is clearly less problematic than the previous two paraphrases. The majority of students across levels and countries found Paraphrase 3 acceptable (with the exception of high school students in Lithuania, where approx. 50% found it acceptable). Notice also that Ireland and Switzerland generally had a lower acceptance rate than the other countries, indicating a possible cultural factor at play. The overall result is in line with results obtained by Roig (1997) from American undergraduate students. Roig, however, also found that university teachers did not always agree that paraphrases that include a reference, but where only small insignificant changes have been made are acceptable. Sometimes such paraphrases are perceived to be problematic, though not necessarily serious (ibid.).

As expected, the majority of students across all levels and nationalities found Paraphrase 4 acceptable. However, the minorities that disagreed are in some cases rather substantial (9-24%), indicating that doubt and disagreement can exist even when a passage has been rewritten thoroughly enough to be acceptable. Such doubts and disagreements were also observed in the interviews, where a commonly described doubt among bachelor students in relation to plagiarism was when a passage had been sufficiently reworked to be acceptable (see D2.1). It is also worth noting that in this case, PhD students were more inclined to consider the paraphrase to be acceptable or completely acceptable compared to students at the other levels. Compared to Paraphrase 1 and 2, where we saw the opposite trend, this indicates that the PhD students were better at recognising the differences between cases than students

from the other two levels, and thus had a better practical understanding of how to incorporate the works of others into their own text.

5.1.3 Are students who find plagiarism acceptable confident about their knowledge on plagiarism?

In section 4.1, we saw that students are generally rather confident that they know the rules covering our three topics. The previous section indicates that this belief is likely to be false for many students. Specifically, the students who believed that Paraphrase 1 above was acceptable and at the same time believed that they had a good understanding of the rules and regulations on citation seem to have false beliefs about their own knowledge. To investigate the prevalence of such false beliefs, we compared the index measuring confidence about knowledge of rules and ethical behaviour described in section 4.1 with the answers for Paraphrase 1. The comparison shows a slight tendency for students who found Paraphrase 1 acceptable, completely acceptable, or who were neutral or didn't know to be significantly less likely to claim to be knowledgeable about rules and ethically correct behaviour ($F=38.04(df5)$; $p<0.001$). The pattern is similar across the three study levels (data not shown). However, as shown in **Figure 7a**, students who found Paraphrase 1 acceptable were also rather confident about their knowledge of rules and ethical behaviour, as they have a median score of approx. 0.75 on the index, with students who found the paraphrase unacceptable only having a slightly higher median score (0.81).

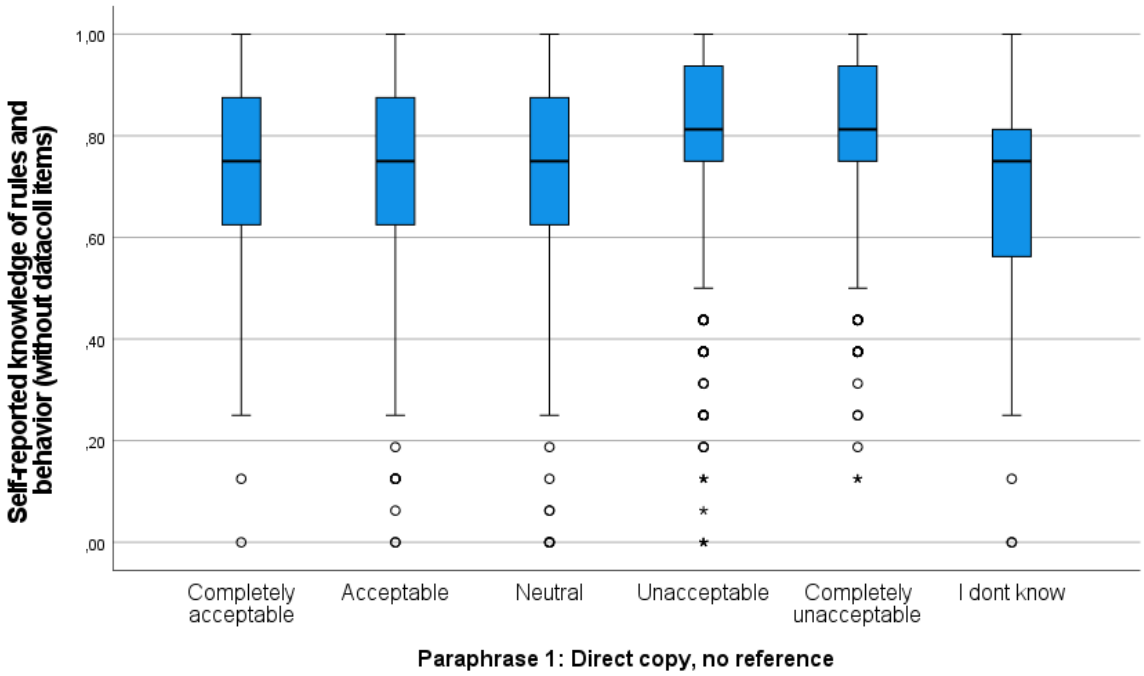


Figure 7a: *Scores on the index* self-reported knowledge about rules and behaviour related to academic misconduct in general*, but not including data collection items (Box plot – where the top and bottom of the boxes are 1st and 3rd quartiles and the lines in the middle of the boxes are medians), stratified by how acceptable the students find Paraphrase 1.*

A somewhat stronger pattern emerges if we focus exclusively on perceived knowledge of ethical behaviour concerning citation and plagiarism (see Figure 7b) ($F=76.99(df6)$; $p<0.001(Eta^2: 0.271)$). As expected, students who found Paraphrase 1 unacceptable were more confident (Median scores approx. 0.75), and students who found it acceptable were less confident about their knowledge of rules and ethical behaviour concerning citation and plagiarism (Median scores approx. 0.90). However, the difference between the two groups is small, so a substantial proportion of those who found paraphrase 1 acceptable believe that they know the rules and how to behave ethically in relation to citation and plagiarism. Importantly though, the bottom 25% of students who responded "I don't know" scored very low on the knowledge index.

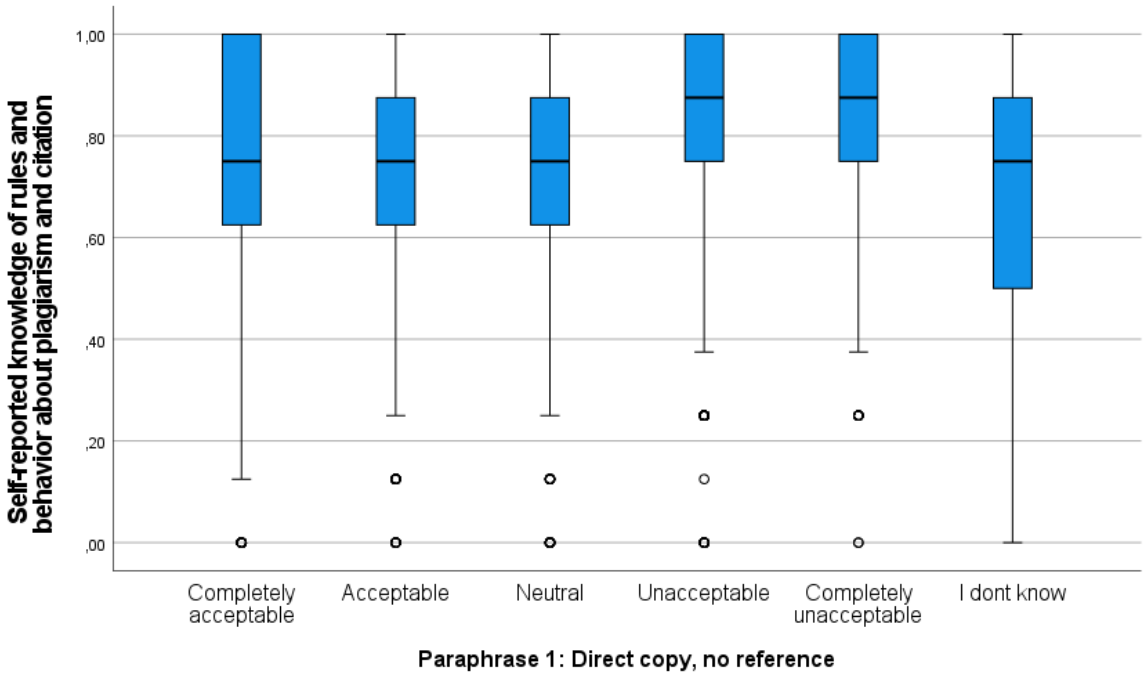


Figure 7b: Scores on the index *self-reported knowledge about rules and behaviour related to academic misconduct regarding plagiarism and citation* (Box plot – where the top and bottom of the boxes are 1st and 3rd quartiles and the lines in the middle of the boxes are medians), stratified by how acceptable the students find Paraphrase 1.

5.2. Perceived rules and regulations regarding working with others and assigning authorship

Since the contexts surrounding the production of texts and assigning authorship are markedly different across the various study levels, the examination in this section is divided into issues relevant to high school and bachelor level students (section 5.2.1), and PhD students (section 5.2.2).

5.2.1 High school and bachelor students

Figure 8, consisting of the four graphs below, reports results from four different cases where high school and bachelor students were asked whether specific cases – all involving assistance when completing an assignment – violated the rules that apply to them. Students were asked whether they "believe the following actions are against the official rules and regulations that apply to you in relation to working with others and assigning authorship", and were asked to respond using the same options as those described in section 5.1.1 (Figure 5).

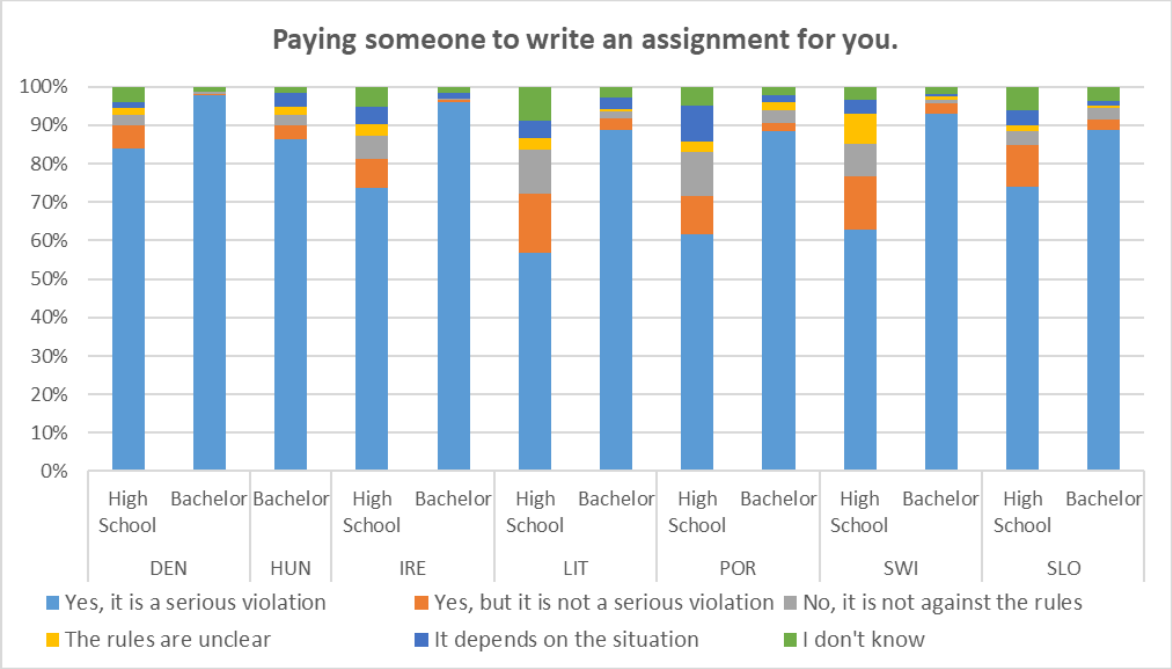
When comparing the four cases, we note that there are clear differences in the extent to which they are deemed to violate rules regarding working with others and assigning authorship. Furthermore, there is a high level of variation in responses to three of the cases, which indicates that there is much confusion, or perhaps loose and undefined rules. Indeed, for three of the cases, many (typically 10-20%) responded that "it depends on the situation" whether the case was against the rules, and approximately 10% said that "…the rules are unclear". Another important and clear finding to highlight is that high school students were less inclined to flag that all four of the cases violated the rules that apply to them.

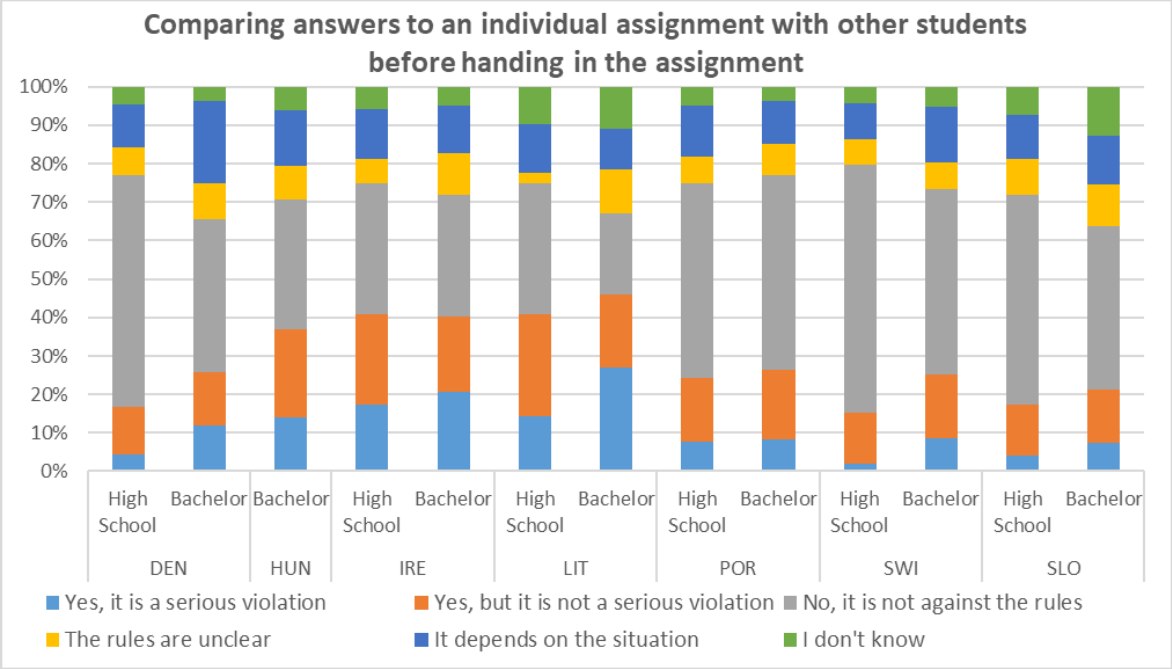
Most students reported that it is clearly against the rules to pay "…someone to write an assignment for you" (**Figure 8**, first graph). Across countries, a large majority of bachelor students in general, and practically all bachelor students in Denmark and Switzerland, deemed this a "serious violation". However, only about 60% of high school students in Switzerland, Lithuania and Portugal believed this to be a "serious violation" (63%, 57% and 62%, respectively), while only about 75% of the high school students in these three countries (77%, 72% and 75%, respectively) believed it to be a violation at all (serious or not serious). These numbers are worth noting given that paying someone to write an assignment for you is a serious and clear transgression of the rules by any standards. They indicate that large minorities

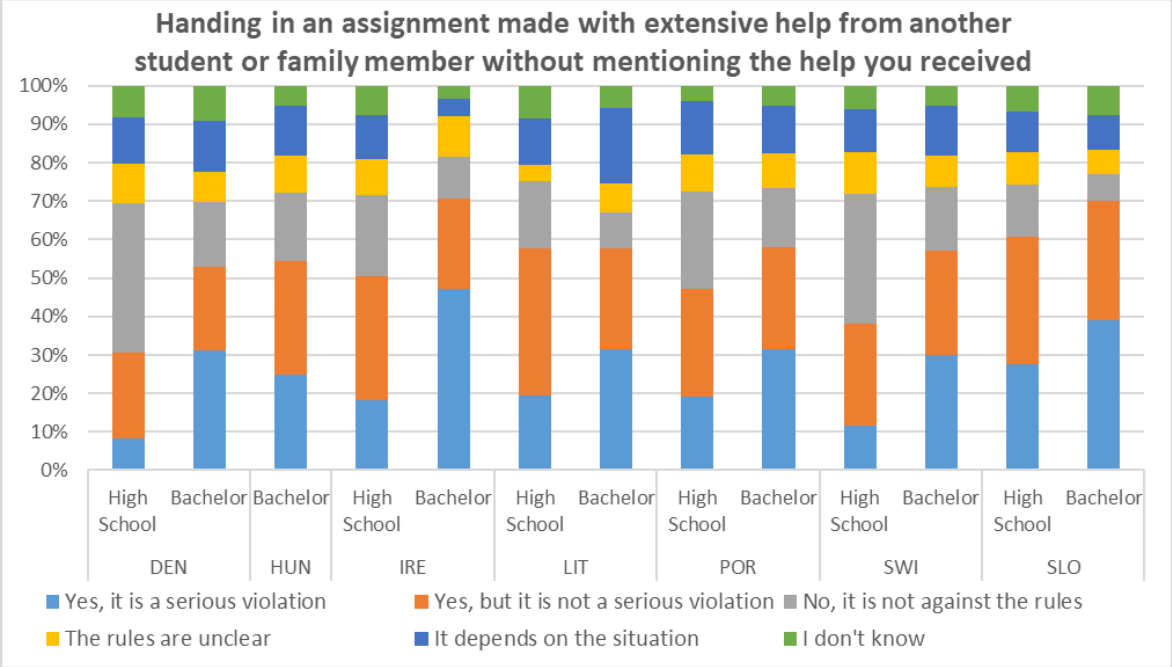
(about 25%) of high school students in these countries have serious deficits in their understanding of rules and regulations for authorship.

If we turn to the second scenario, many students responded "No, it is not against the rules" to compare "…answers to an individual assignment with other students before handing in the assignment" (**Figure 8**, second graph). There are some country-specific variations, as only one third of bachelor students in Denmark and Portugal responded "No, this is not against the rules". However, on the whole, there was no common understanding of whether "comparing answers… before handling in the assignment" is against the rules. In this case, one could argue that "It depends on the situation" or "The rules are unclear" would be the most appropriate answers at least in some countries, as there is a considerable difference between comparing answers to, say, a take-home exam and a weekly assignment. In countries where this is the case, students who believed there to be clear and universal rules that either allow or rule out the conduct have a false belief. This false belief is worth noting as it may deter students from asking about the rules that apply in specific situations.

"Please indicate whether you believe the following actions are against the official rules and regulations that apply to you in relation to working with others and assigning authorship"




Comparing answers to an individual assignment with other students before handing in the assignment
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
High School
Bachelor
DEN
HUN
IRE
LIT
POR
SWI
SLO
Yes, it is a serious violation
Yes, but it is not a serious violation
No, it is not against the rules
The rules are unclear
It depends on the situation
I don't know


Handing in an assignment made with extensive help from another student or family member without mentioning the help you received
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
High School
Bachelor
DEN
HUN
IRE
LIT
POR
SWI
SLO
Yes, it is a serious violation
Yes, but it is not a serious violation
No, it is not against the rules
The rules are unclear
It depends on the situation
I don't know

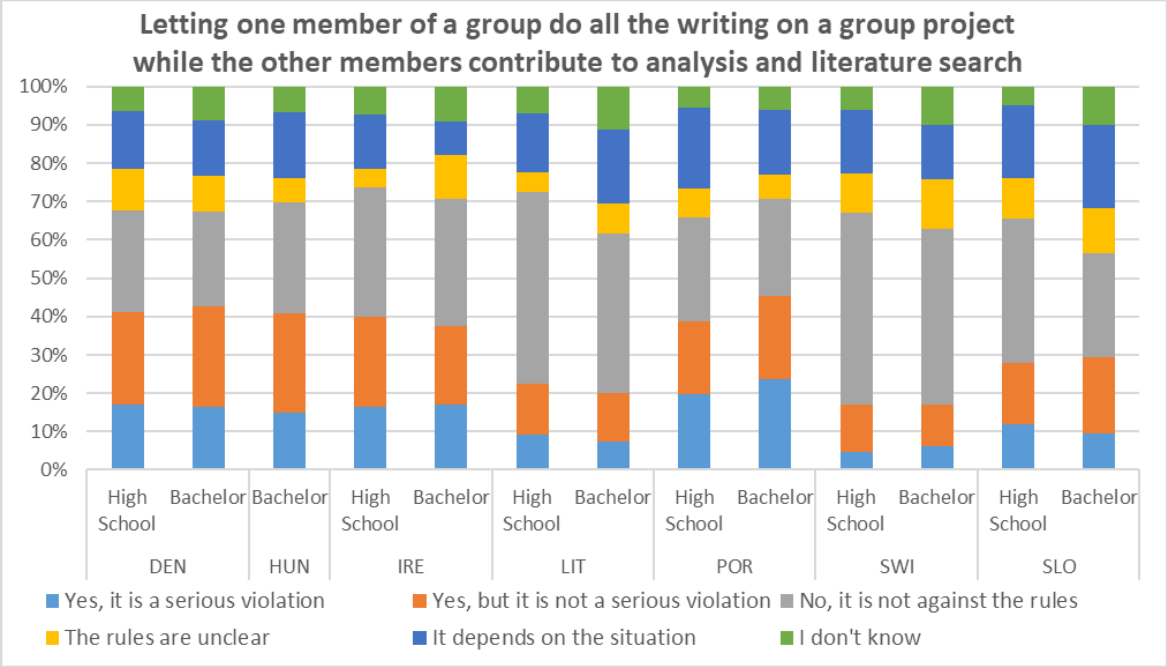


Figure 8. Perceptions of what constitutes a violation of rules and regulations when working with others and assigning authorship (Stacked bars; responses sum to 100% for each bar) for high school students and bachelor students across countries.

In the third case, "Handing in an assignment done with extensive help from another student or family member without mentioning the help you received" (**Figure 8**, third graph), the students were also divided about whether this would violate the rules. As in the previous case, there was no clear common understanding of the rules, although it should be mentioned that approximately 50% of bachelor students found this to be a violation ("serious" or "not… serious"), and one could also question whether there are clear and universal rules governing this conduct (where "The rules are unclear" or "It depends on the situation" might be the most appropriate answers).

The fourth case focused on the division of labour when writing an assignment and asked whether it is against the rules to "Let one member of a group do all the writing on a group project while the other members contribute to analysis and literature search" (**Figure 8**, bottom graph). As with three of the other cases, there was no common understanding of the rules among students – responses to the question varied greatly. The clearest difference between countries can be seen in this fourth case: Swiss students at both the bachelor and high school level were much less inclined to report this case as a

violation (approximately 15% flagged this as a "serious" or "not... serious" violation) compared to bachelor students in Denmark, Hungary, and Portugal.

The qualitative interviews conducted prior to developing the survey and reported in D2.1 also indicate that there is substantial uncertainty about these topics, and that the confusion sometimes stems from unclear rules and mixed signals from institutions and teachers. For instance, some interviewees mentioned that while there may be an official rule saying that they must complete a specific assignment, e.g. a take-home exam, without help from others, their teachers may more or less directly encourage them to collaborate, or at least have someone read their assignments before handing them in. Unsurprisingly, such practice can create confusion, especially if there are various deviations from the official rules on different courses and if the specific rules that students are expected to follow in the concrete situation are not communicated well.

5.2.2 PhD students

Turning to PhD students, the initial examinations (not reported) revealed that the perception of rules regarding production of texts and assigning authorship depends on study direction (natural science, medical science, social science, humanities/art), and the type of data used. There are also national differences, which will not be reported in this publication. Therefore, we focus on differences observed among study directions and data type used in the four cases reported in the four graphs of **Figure 9**. We omitted PhD students who reported that they study subjects within law or non-classifiable study directions (other), as these two segments had limited representation in the data (n <30; see Table 4 in the method section).

In all four cases, the student was asked to consider a situation in which they are finalising a research paper within their field. They were then asked whether they believe it is acceptable that a particular person should be included as co-author on the paper. Response options ranged from 1 "completely unacceptable" to 5 "completely acceptable" (and "I don't know"). We report the share of respondents who found it at least "acceptable" to include the person in question. The first person, Dr. Doe (upper graph of **Figure 9**), is the students' primary supervisor who suggested and designed the idea/hypothesis for the paper, gave critical comments on the first draft, but did not write anything (the question formulation varied depending on the type of data and study direction – see details above each graph). Dr. Doe fulfils both the first and second of the four Vancouver criteria for authorship by having made a substantial contribution to the conception and design of the study (first criterion, which can also be

fulfilled by making a substantial contribution to "he acquisition, analysis, or interpretation of data") and also providing critical comments (second criterium, which can also be fulfilled by drafting the manuscript) (ICMJE 2017). The second person, Dr. Jones (second graph in **Figure 9**), is the students' secondary supervisor whom the student visited during the study. Data or material for the study were collected during this visit. Furthermore, Dr. Jones also read the manuscript and gave critical comments (the question formulations varied depending on the type of data used by the PhD student – see details above the graph). Dr. Jones fulfils the second, but not the first Vancouver criterion.

Looking at the two graphs, major differences can be seen between STEM students (comprising natural science, engineering, and medical students) and humanities/arts students, as 57-90% of the former group found it acceptable to include Dr. Doe and Dr. Jones, compared to 27-58% of the latter group. The social science students occupy a middle position, but also exhibit the highest variation, which is very much a function of the type of data they use. Social science students who use quantitative data were much more likely to accept both individuals as authors (approximately 55% (Dr. Jones) and 67% (Dr. Doe) found it acceptable), while those who use qualitative data were less likely to accept (44% for Dr. Doe and 41% for Dr. Jones), and those who use historical data sources or works of art were even less likely to accept (27% for Dr. Doe and 35% for Dr. Jones).

The third case (third graph in **Figure 9**) also asked PhD students to consider whether a particular person should be included in a paper that the student is finalising. However, in this case (and the next), the person is not a supervisor. It is instead Dr. Santos, who made an analytical contribution and provided early input to the method section. However, Dr. Santos has not read the full manuscript (the question varied depending on the type of data used by the PhD student – see details above the graph). Dr. Santos fulfils the first Vancouver criterion, and depending on the interpretation, also the second. Since responses to the first two cases could be influenced by the asymmetric relationship between supervisors and students, it is important to see whether the observed differences in study direction and type of data used also appear in non-supervisor situations. In fact, we see the same patterns as in the two first examples, i.e. students who work with quantitative data were more inclined to accept Dr. Santos as an author. In addition, there is a similar gradient among social science students when comparing those who work with quantitative data, qualitative data, and historical sources/works of art.

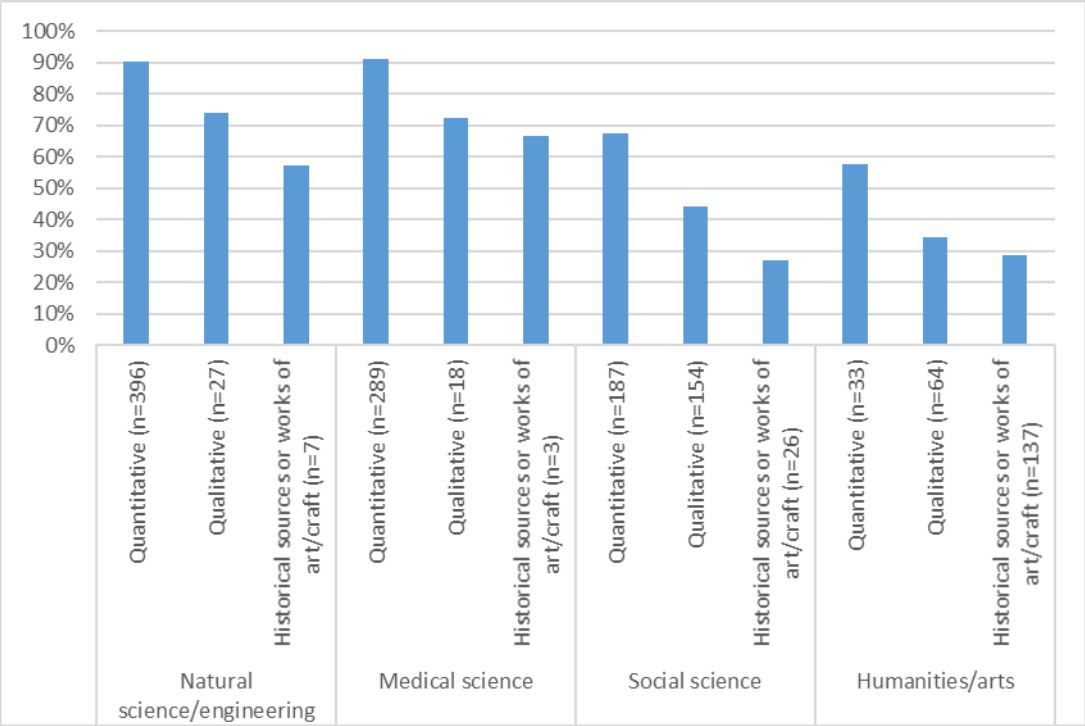
"You are finalising a research paper reporting on a study that you were in charge of. The study tests a novel hypothesis using data from two different sources. An additional four people were involved in the study in various ways. For each person, please indicate whether you believe it would be acceptable to add him or her as a co-author of the paper."

"Dr. Doe: Is your primary supervisor. He did not write anything. He.."

QUAN DATA: …suggested the hypothesis to be tested, helped you design the study, and gave critical comments on the first draft of the publication, but did not write anything.

QUAL DATA: …suggested the research topic, helped you design the study, and gave critical comments on the first draft of the publication, but did not write anything.

HISTORICAL SOURCES/WORKS OF ART: …suggested the research topic, helped you design the study and gave critical comments on the first draft of the publication, but did not write anything.



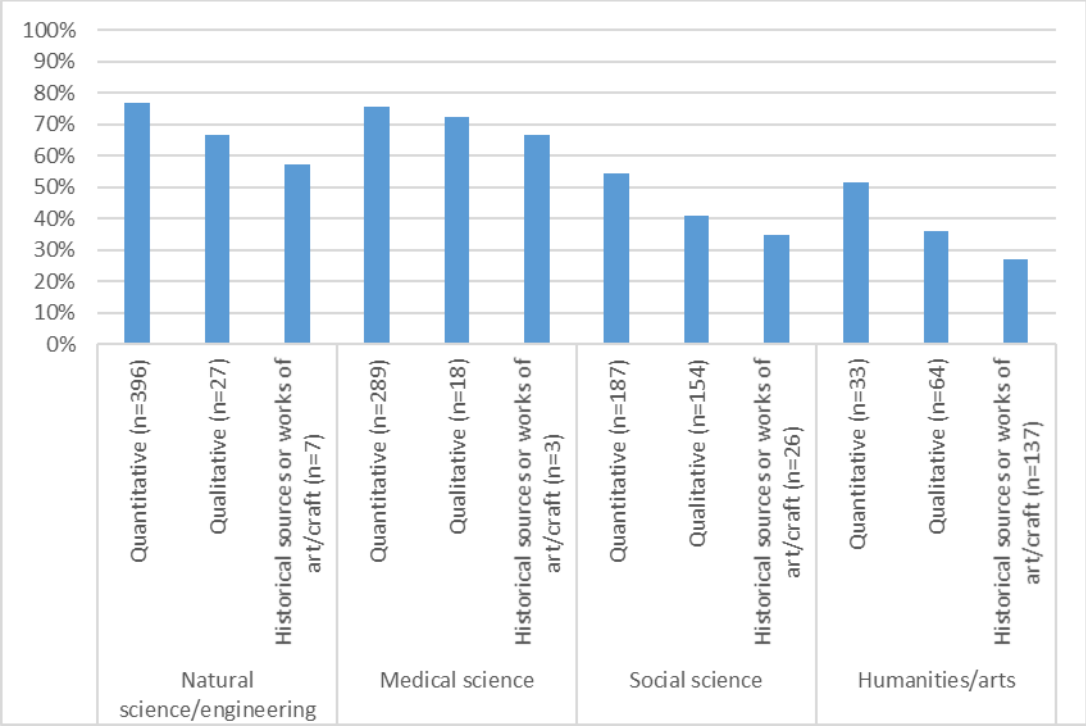
"Dr. Jones: Is your secondary supervisor. You visited Dr. Jones' research group during the study/project. You collected some of the …"

QUAN DATA: …data you used in the study during your visit. Dr. Jones has read the manuscript and provided critical comments.

QUAL DATA: …material you used in the study during your visit. Dr. Jones has read the manuscript and provided critical comments.

HISTORICAL SOURCES: …material you used in the study during your visit. Dr. Jones has read the manuscript and provided critical comments.

WORKS OF ART: …works of art you used in the study during your visit. Dr. Jones has read the manuscript and provided critical comments.

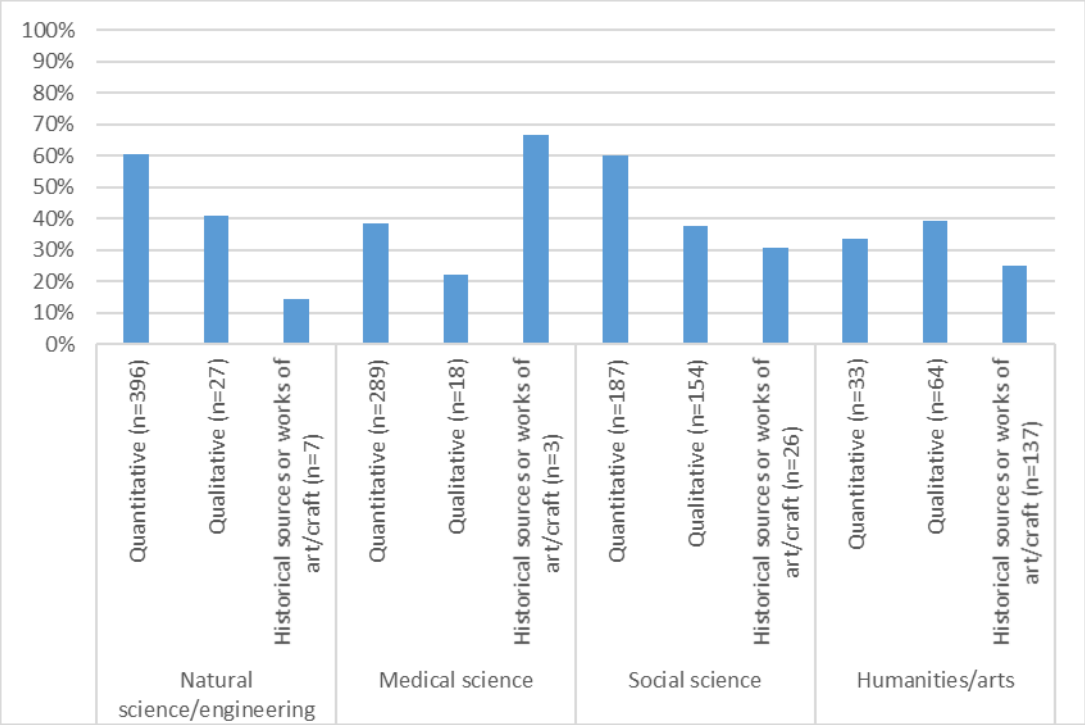


"Dr. Santos:

QUAN DATA: …performed part of the statistical analysis of the data. She provided some early input for the methods section, but has not read the full manuscript.

QUAL DATA: …coded some of the qualitative data. She provided some early input for the methods section, but has not read the full manuscript.

HISTORICAL SOURCES/WORKS OF ART: …helped you to compile background information from the archives. She provided some early input for the methods section, but has not read the full manuscript.

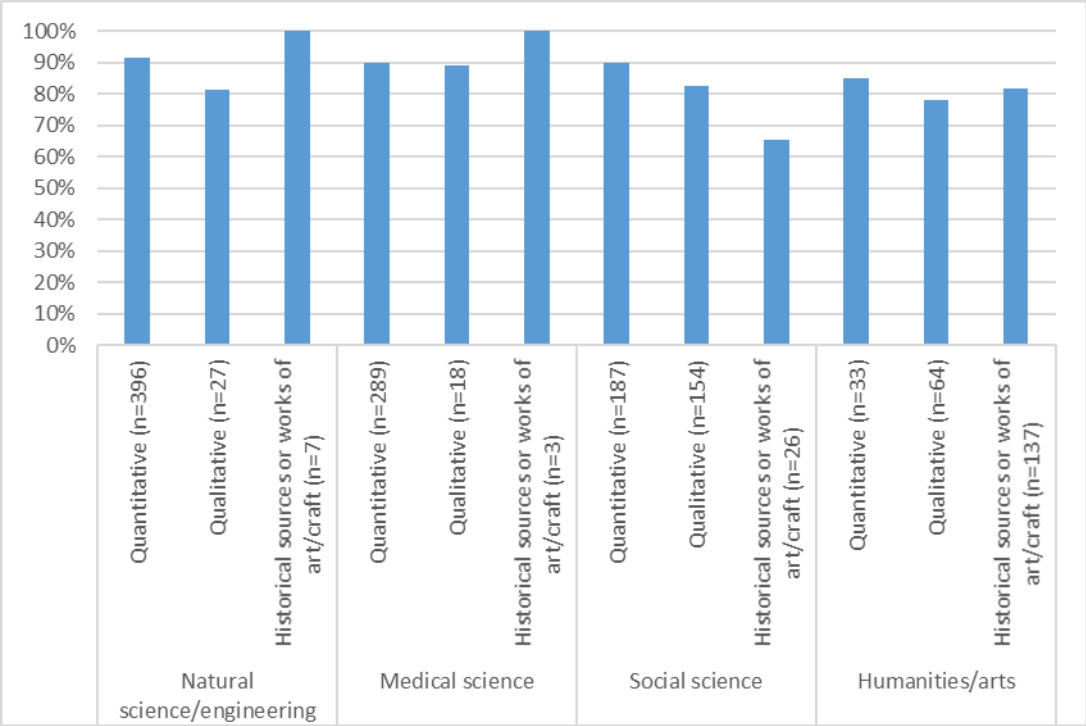


"Ms. Olson: Works as a technician/librarian at your institution. She taught you…

QUAN DATA: …one of the key methods you used in connection with your data collection and contributed with an important adjustment of the method. She wrote part of the methods section and has read the full manuscript.

QUAL DATA: …how to use a new computer program to support your analysis and contributed with an important adjustment to your method. She wrote part of the methods section and has read the full manuscript.

HISTORICAL SOURCES/WORKS OF ART: …how to use a new computer program to support your analysis and contributed with an important adjustment to your method. She wrote part of the methods section and has read the full manuscript.



We identify roughly the same difference between study directions, and types of data used in acceptability to include a non-supervising person compared to a supervisor. However, there is one marked difference – PhD students in medicine (irrespective of the data used) found it much less acceptable to include Dr. Santos than PhD students from natural and social science. This difference is difficult to interpret. Perhaps it can be ascribed to the division of labour within medicine, where it is quite common to hire a data analysis specialist (typically a statistician), who would not be part of the author team, to carry out analyses. Another explanation could be that Dr. Santos had not read the manuscript, which may be seen as the minimum requirement for authorship within large areas of medicine.

The fourth case explores whether a person's position in the academic hierarchy will influence the students' decision to grant authorship. In this case, the helper – Ms. Olson – is positioned low in the academic hierarchy (as a technician/ librarian) and is not designated with the academic title "Dr.". Nevertheless, Ms. Olson has contributed to the design of the study and has written part of and read the final paper. Ms. Olson is thus the only one of the four helpers to fulfil all of the first three Vancouver criteria, and should be granted authorship (if she accepts it). Across study directions and data types, students found it "Acceptable" or "Fully acceptable" to add Ms. Olson as an author (with acceptance rates generally over 80%). The only exception is social science students working with historical sources or works of art, where the acceptance rate was 65%. It is especially worth noting that the acceptance rate for Ms. Olson is well beyond that of any of the other helpers, including the primary supervisor Dr. Doe. This suggests that the students are able to base the decision to grant authorship on merits and not (only) on perceived position within the academic hierarchy.

Furthermore, it should be noted that the helper in the first case (Dr. Doe) was gendered using the masculine pronoun ("he"), the second helper (Dr. Jones) was not gendered, the third helper (Dr. Santos) was gendered using the feminine pronoun "she" and the fourth helper (Ms. Olson) was gendered using both a feminine pronoun ("she") and a feminine title ("Ms."). It is well-know from the literature (eg. Monroe et al. 2008; Moss-Racusin 2012, Maliniak et al 2013, Knobloch-Westerwick et al. 2013, Reuben et al. 2014) that there is a heavy gender-bias in academia in the sense the female researchers and students are perceived as less competent in teaching and research. To mention just a few of the ways this biased perception is visible in academic life female teachers get less favourable student evaluations than their male colleagues, female researchers are seen as lees eligible for academic positions, female researchers are cited less than their male colleagues, and female researchers have more difficulties

getting funding than their male colleagues. Due to the length of the questionnaire we had to conflate the gender and the academic hierarchy dimensions. Thus, we chose to gender Ms. Olson as a female to test the combined effect of gender bias and bias towards a person perceived as being low in the academic hierarchy. Had the acceptance rate of Ms. Olson been disproportionately low, we would have known that either gender or academic status or both are a source of unjust bias in authorship decisions. On the positive side, the acceptance rate we see for Ms. Olson is not in itself a cause of immediate alarm. However, the relatively high acceptance rate of the two supervisors (Dr. Doe and Dr. Jones), who strictly speaking do not qualify as authors, may in part be due to their gendering, and given the construction of the questionnaire, we certainly cannot rule out that a different gender distribution among the helpers would have led to a different distribution of acceptance rates.

"Please indicate whether you believe the following actions go against the rules and regulations that apply to you regarding authorship."

"Adding a supervisor who only added critical comments on the manuscript as a co-author of the research publication."

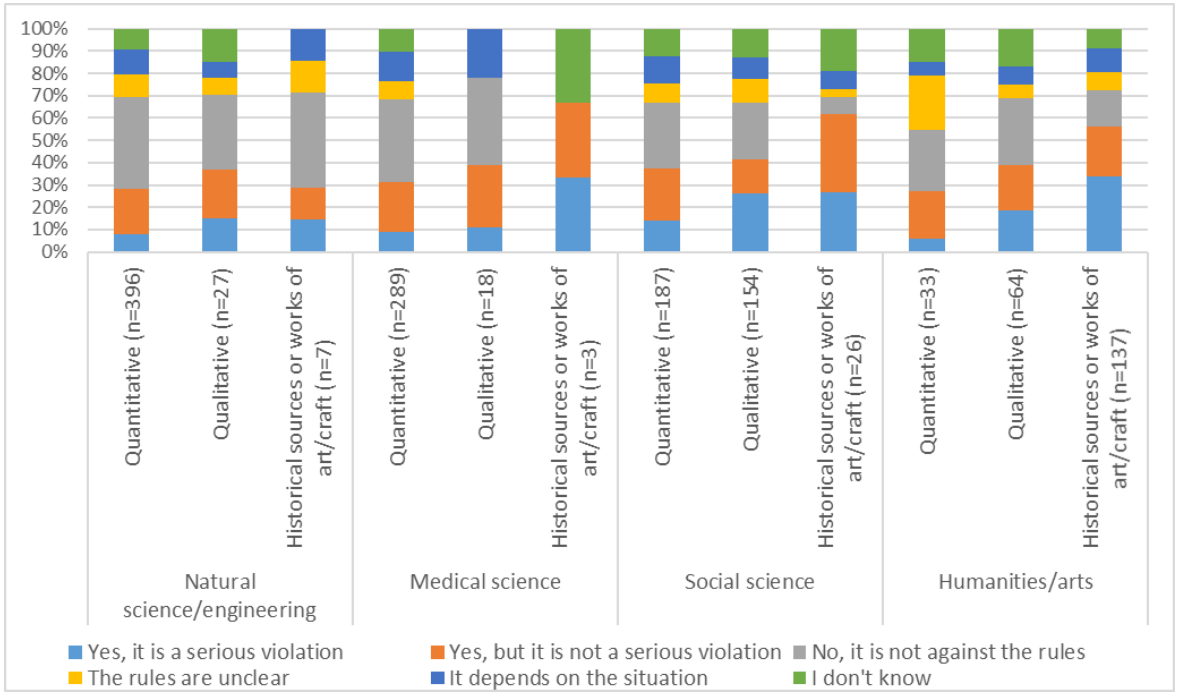


Figure 9. Perceptions of questionable conduct regarding assigning authorship (First four graphs report shares within groups. Bottom graph is a stacked bar where responses sum to 100%) – PhD students, stratified by study direction and primary data source employed.

Figure 9 illustrates an aspect of how PhD students perceived the legal side (rather than the ethical acceptability) of granting co-authorship to someone who only fulfils the second Vancouver criterion (similar to Dr. Jones in Figure 8). Again, students working with quantitative data were less prone to think that this is a violation of the rules and regulations that apply to them. Across study directions, the majority of students answered either yes or no, indicating that they believe that the rules and regulations that apply to them provide a clear answer to the question. Of those unable to clearly state whether or not the action is a violation of the rules and regulations that apply to them, there is some variation across fields as to whether they indicated that this uncertainty stems from the rules and regulations being unclear or from their own lack of knowledge of existing rules.

5.3. Perceived rules and regulations regarding handling of data

In **Figure 10**, we report student perceptions of what constitutes violations of the rules that apply to them with respect to the handling and presentation of data. We presented three different situations. Similar to many other cases reported in this chapter, the response options were "Yes, this is a serious violation", "Yes, but not a serious violation", "No, it is not against the rules", "The rules are unclear", "It depends on the situation", and "I don't know". Students from all study levels were given the three questions, which varied depending on the type of data used and the type of text produced (see details above each graph).

We found quite large differences across study levels, where students at higher academic levels were inclined to report the cases as violating the rules. Furthermore, bachelor students and particularly high school students were more inclined to answer "I don't know", suggesting more uncertainty of what constitutes good practice.

The first case (top graph of **Figure 10**) asked whether it would violate the rules and regulations that apply to the student in relation to data collection, analysis, and presentation to remove data or material that does not give similar results compared with the remaining data or material when the reason for the lack of similar results is unknown. On the whole, we found large differences across study levels, and a clear gradient such that students from higher study levels were more inclined to see the action as a violation of the rules. PhD students were therefore much more inclined to state that the action would violate the rules (70-80% of PhD students, except in Portugal, perceived this to be a violation – in most cases a "serious violation") than bachelor students, 60-75% of whom saw this as a violation (however,

many bachelor students did not see this not a serious violation). High school students were clearly less inclined to think of the action as a violation (35-53% of high school students stated that it is a violation or a serious violation).

The second case (middle graph of **Figure 10**) asked whether it would violate the rules and regulations that apply to the student in relation to data collection, analysis, and presentation to remove data or material that does not give similar results compared with the remaining data or material when the reason for the lack of similar results is known. Across countries and study levels, the overall perception of this case was similar to the previous case, since a similar share of students believed it to be a violation. However, the overall share reporting this as a violation dropped slightly, and fewer students believed it to be a "serious violation".

The third case (bottom graph of **Figure 10**) asked whether it would violate the rules and regulations that apply to the student in relation to data collection, analysis, and presentation to insert constructed data based on the general findings of the remaining data or material. Compared to the two former cases, an even greater number of students perceived this type of data handling to violate the rules. Similar to the other two cases, a clear gradient following study level can be observed, such that PhD students were more inclined to report this as a violation (PhD students in Portugal are, again, an exception here), bachelor students were slightly less inclined to report it as a violation (except in Portugal), and high school students were less inclined to report it as a violation.

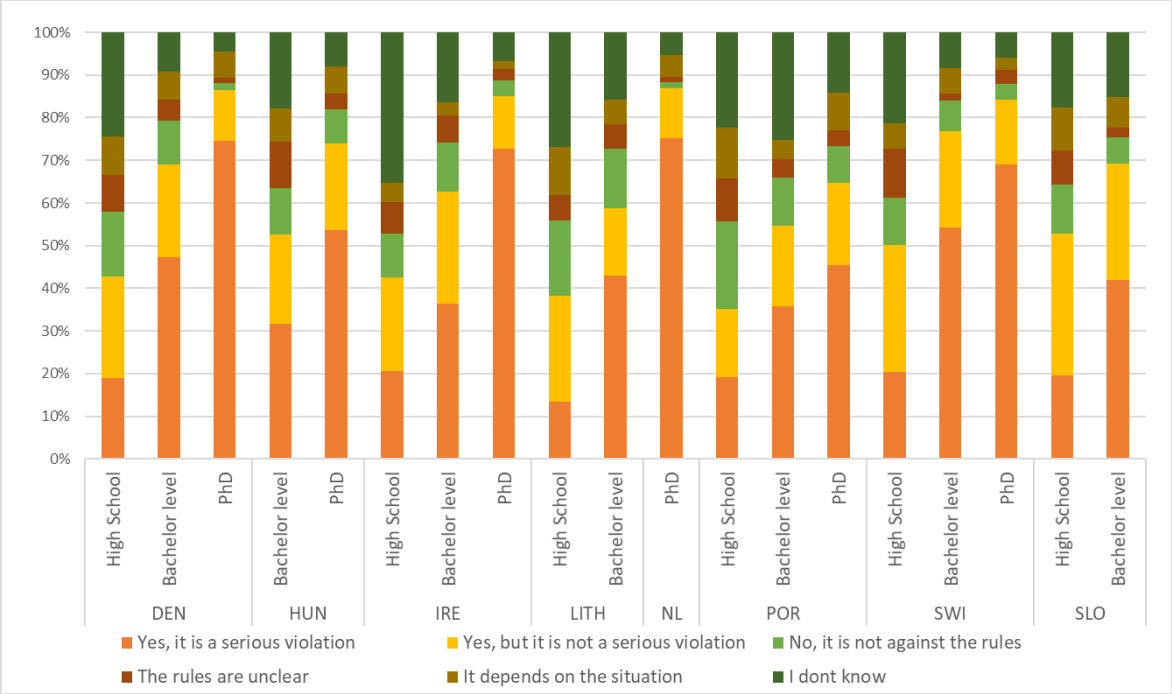
"Please indicate whether you believe the following actions go against the rules and regulations that apply to you in relation to data collection, analysis and presentation."

Not mentioning in **[an assignment/a research publication]** that you…

…[*Quantitative data]*: …removed some deviating data points from a data set when the cause of the deviation was unknown.

…[*Qualitative data]*: …discarded one or more interviews that did not fit well with the rest of your interviews when the cause of the deviation was unknown.

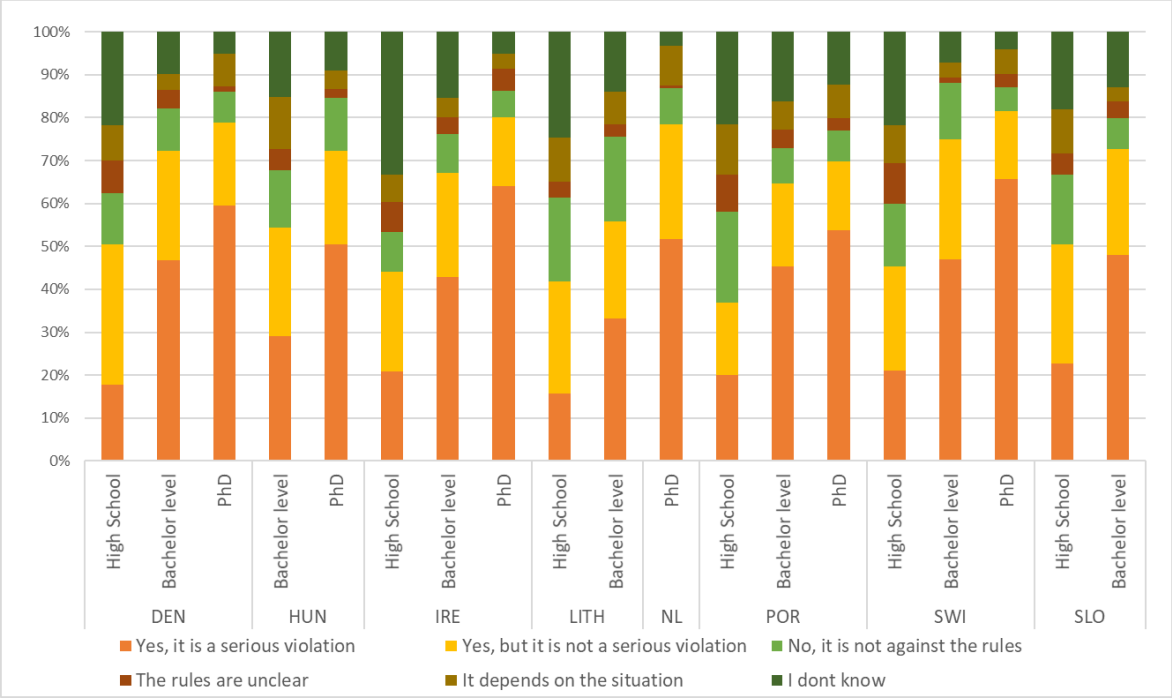
…[*Artifacts/Historical]*: …discarded some of your material that did not fit well with the rest of your material when the cause of the deviation was unknown.



Not mentioning in **[an assignment/a research publication]** that you…

…[*Quantitative data]*: …removed some deviating data points from a data set when the cause of the deviation was known.
…[*Qualitative data]*: …discarded one or more interviews that did not fit well with the rest of your interviews when the cause of the deviation was known.

…[*Artifacts/Historical]*: …discarded some of your material that did not fit well with the rest of your material when the cause of the deviation was known.



Not mentioning in **[an assignment/a research publication]** that you…

…[*Quantitative data]*: …replaced some outliers in a data set with data points obtained through estimates based on the remaining data points.
…[*Qualitative data]*: …made up a number of direct quotes from an interview based on your general impression of the interviewee's viewpoints.

…[*Artifacts/Historical]*: …made up a number of quotes based on your general impression of the viewpoints expressed in a source.

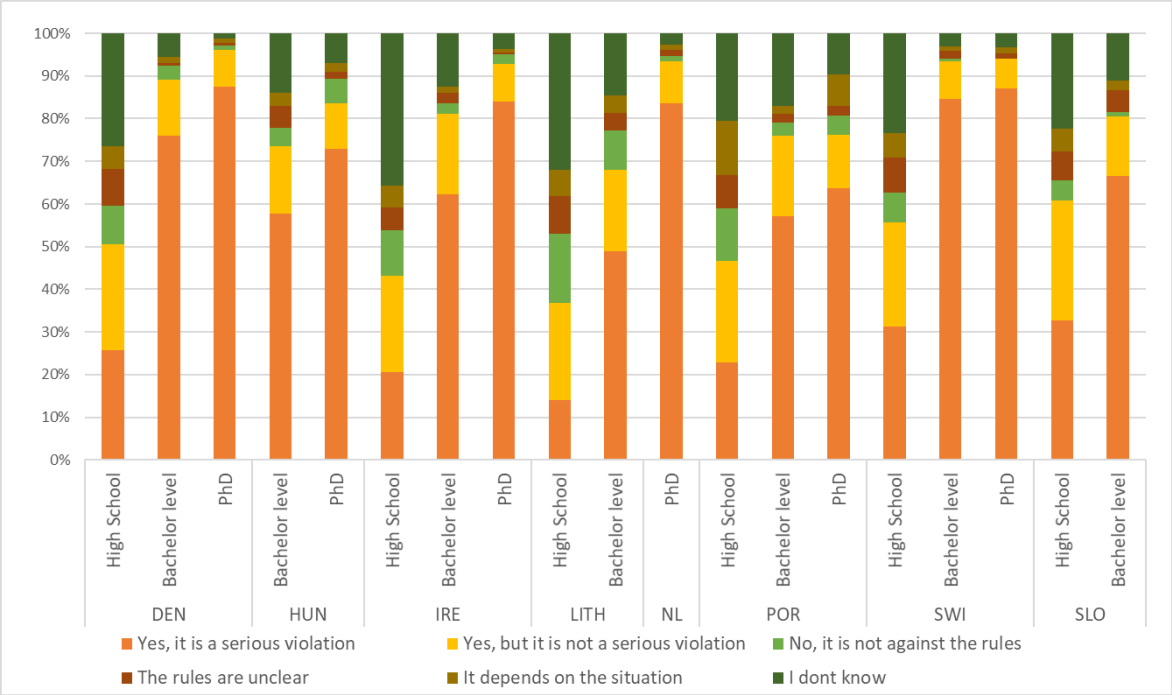


Figure 10. Perceptions of questionable conduct regarding the collection, analysis, and presentation of data (Stacked bars where responses sum to 100%), stratified by country and study level.

6. Perceptions of questionable practice among peers, and own practice

This chapter examines the extent to which students believe that their peers (i.e. classmates and, for PhD students, other researchers in the research field) behave in a way that would usually be considered cheating or questionable practice. We compare this with the students' responses about their own behaviour regarding the same.

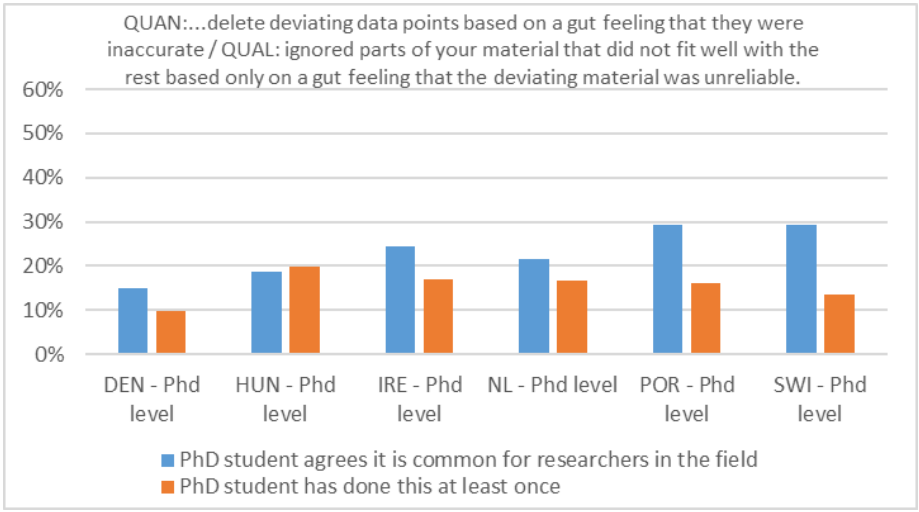
Students at different study levels were not given the same number of questions, as some questions were only relevant to PhD-level students. Although there are a few overlapping questions, we present the results in three separate figures to avoid a very lengthy figure.

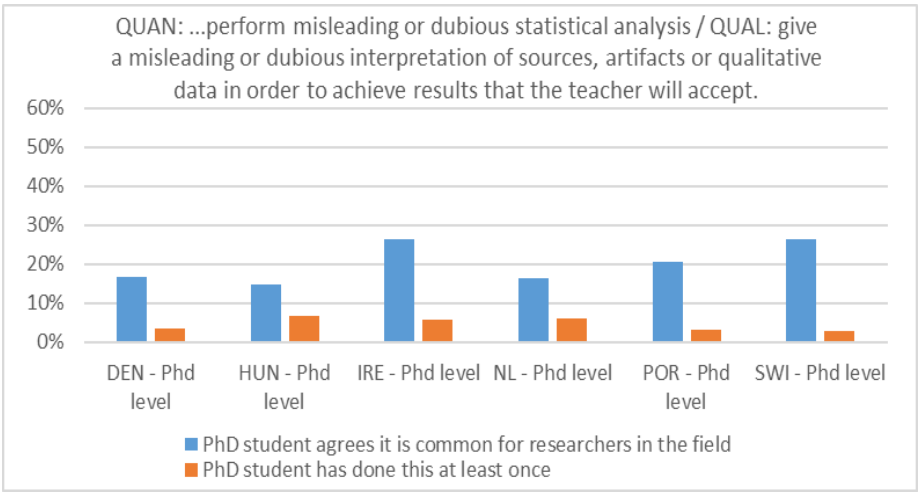
We asked the students a number of questions about whether they agreed or disagreed that it is common for their peers to perform particular actions. In the figures, we report the share of students who agreed that it is common. We asked the students whether they themselves had carried out these actions during their current education. We report the share of students who responded "Yes, once", "Yes a few times", and "Yes, many times".

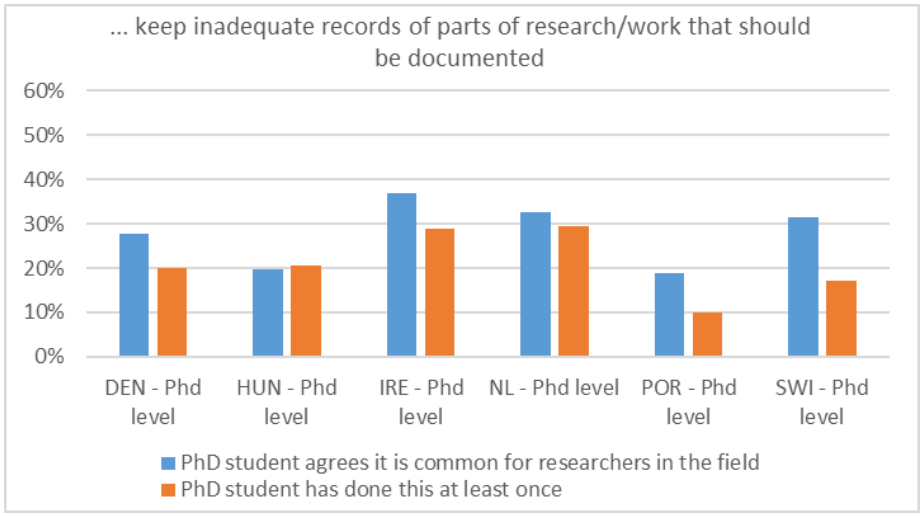
6.1. Perceived peer practice and own practice among PhD students

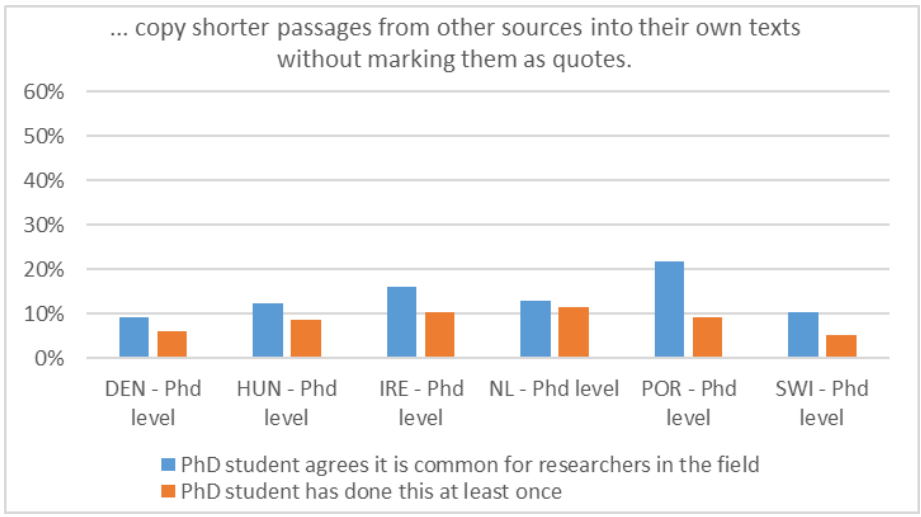
The questionable research practices: "...keeping inadequate records of parts of research/work that should be documented", "...allowing research group leaders, supervisors or others in power to become co-authors of papers, even though they did not make a significant contribution to them", and "...citing sources that are not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour" were perceived as being most common among PhD students. Of these three practices, students were most inclined to report that they themselves had "...allowed [people in power] to become co-authors" (21% in Hungary to 38% in Portugal), and "...kept inadequate records" (across countries: 10% in Portugal to 29% in Ireland). Note that the stated fraction of PhD students who allowed people in power to have an undeserved authorship mention is from the total number of PhD

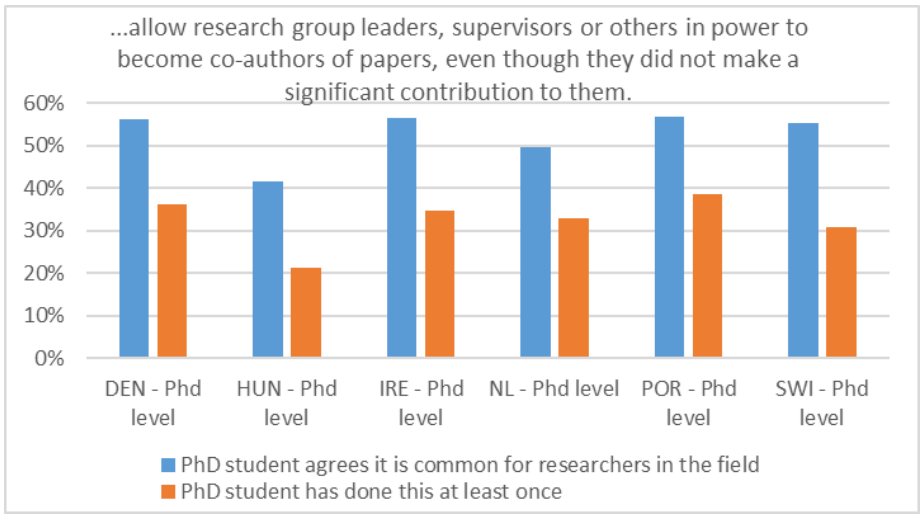
students participating – including those who had not yet published anything. The questionnaire did not include questions about whether the participating PhD students had co-authored any publications, so we cannot estimate how many of the students who had actually published had allowed people in power to become co-authors on publications even when it was not deserved.






... keep inadequate records of parts of research/work that should be documented
60%
50%
40%
30%
20%
10%
0%
DEN - Phd level
HUN - Phd level
IRE - Phd level
NL - Phd level
POR - Phd level
SWI - Phd level
PhD student agrees it is common for researchers in the field
PhD student has done this at least once


... copy shorter passages from other sources into their own texts without marking them as quotes.
60%
50%
40%
30%
20%
10%
0%
DEN - Phd level
HUN - Phd level
IRE - Phd level
NL - Phd level
POR - Phd level
SWI - Phd level
PhD student agrees it is common for researchers in the field
PhD student has done this at least once


...allow research group leaders, supervisors or others in power to become co-authors of papers, even though they did not make a significant contribution to them.
60%
50%
40%
30%
20%
10%
0%
DEN - Phd level
HUN - Phd level
IRE - Phd level
NL - Phd level
POR - Phd level
SWI - Phd level
PhD student agrees it is common for researchers in the field
PhD student has done this at least once

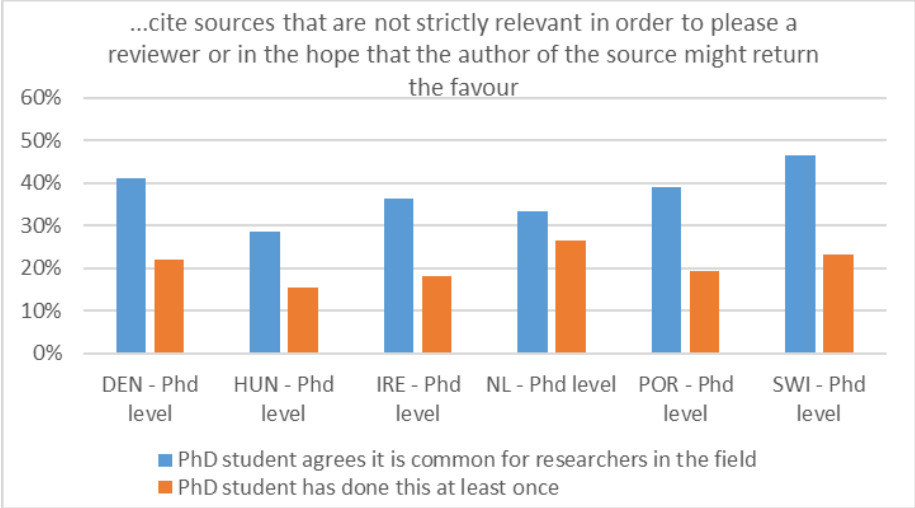
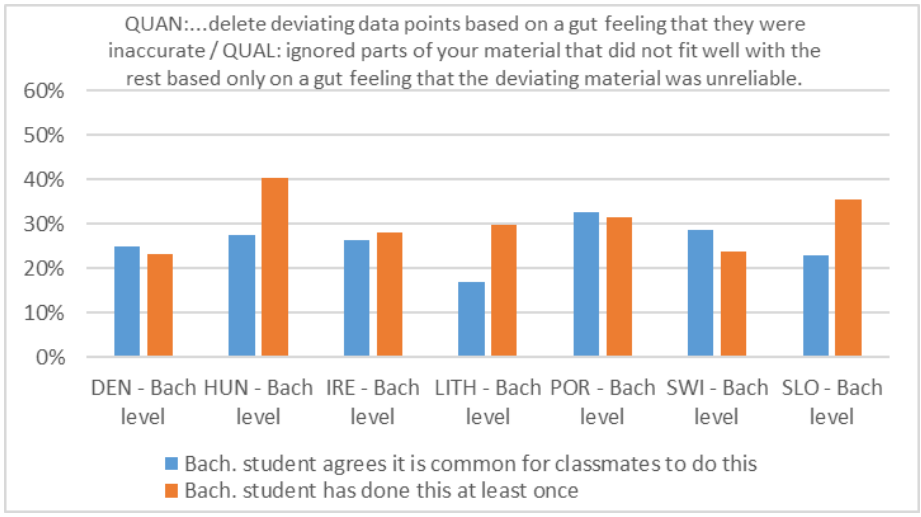


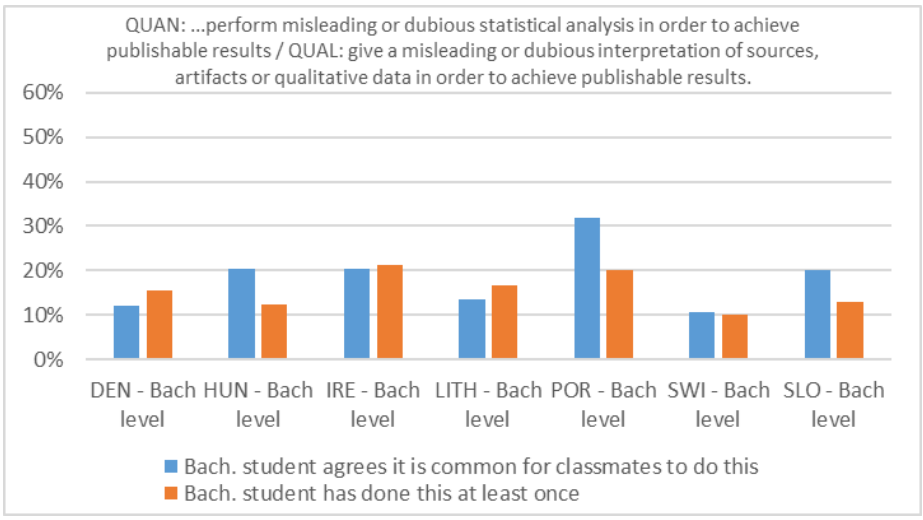
Figure 11. Perceptions of peer behaviour and whether the student has behaved similarly during the PhD study period (Shares within groups are reported), PhD students, stratified by country.

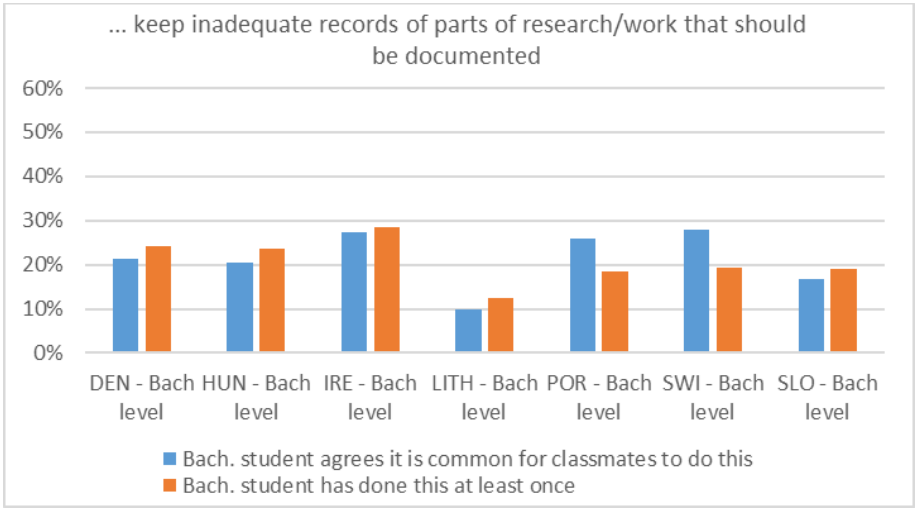
6.2. Perceived peer practice and own practice among bachelor students

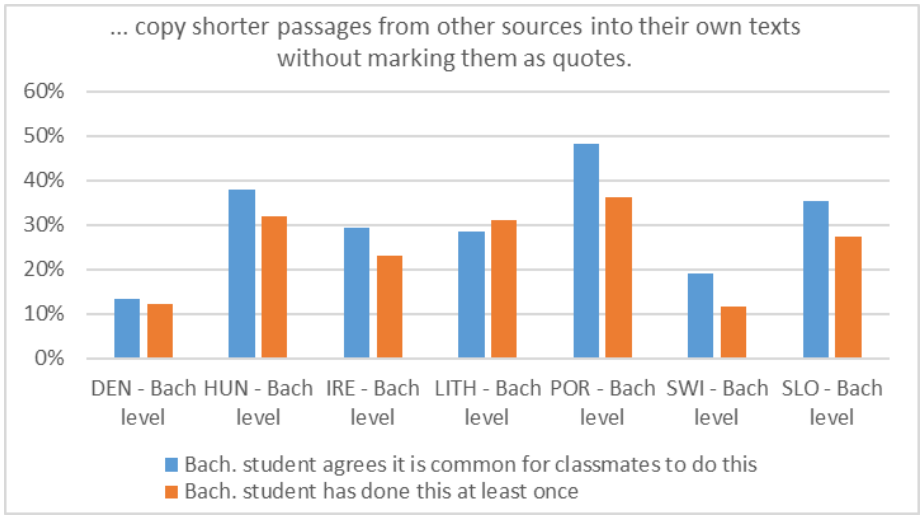
Bachelor students perceived "...receiving help from other students or family members on assignments you were supposed to complete on your own" and "...adding students as co-authors of group assignments, even though they did not contribute" to be the most common questionable behaviours among their peers. Furthermore, in Hungary and Portugal, bachelor students also believed it to be very common to "...copy shorter passages from other sources into their own texts without marking them as quotes."

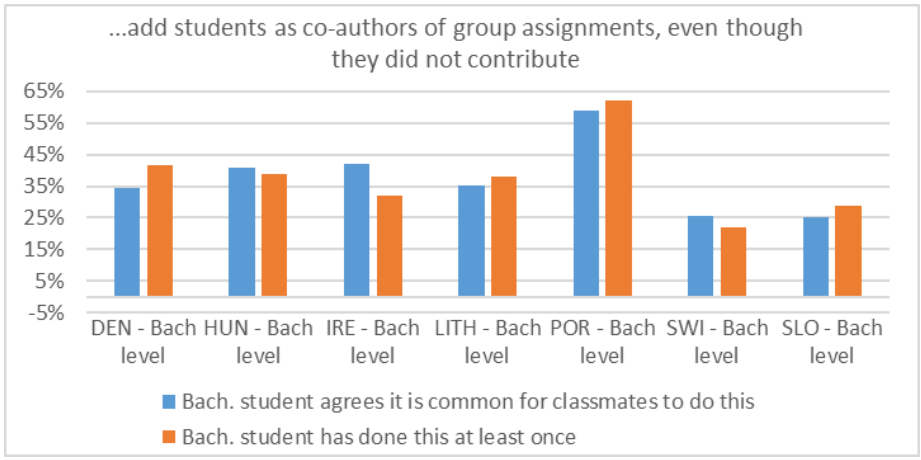
A similar share of bachelor students who reported that their peers behave this way also reported that they do it themselves. Specifically, across countries, 22% (Switzerland) to 62% (Portugal) had "...added students as co-authors of group assignments, even though they did not contribute", and across countries, 40% (Denmark) to 62% (Portugal) had "...received help from other students or family members on assignments you were supposed to complete on your own".


QUAN:...delete deviating data points based on a gut feeling that they were inaccurate / QUAL: ignored parts of your material that did not fit well with the rest based only on a gut feeling that the deviating material was unreliable.
60%
50%
40%
30%
20%
10%
0%
DEN - Bach level
HUN - Bach level
IRE - Bach level
LITH - Bach level
POR - Bach level
SWI - Bach level
SLO - Bach level
Bach. student agrees it is common for classmates to do this
Bach. student has done this at least once


QUAN: ...perform misleading or dubious statistical analysis in order to achieve publishable results / QUAL: give a misleading or dubious interpretation of sources, artifacts or qualitative data in order to achieve publishable results.
60%
50%
40%
30%
20%
10%
0%
DEN - Bach level
HUN - Bach level
IRE - Bach level
LITH - Bach level
POR - Bach level
SWI - Bach level
SLO - Bach level
Bach. student agrees it is common for classmates to do this
Bach. student has done this at least once


... keep inadequate records of parts of research/work that should be documented
60%
50%
40%
30%
20%
10%
0%
DEN - Bach level
HUN - Bach level
IRE - Bach level
LITH - Bach level
POR - Bach level
SWI - Bach level
SLO - Bach level
Bach. student agrees it is common for classmates to do this
Bach. student has done this at least once





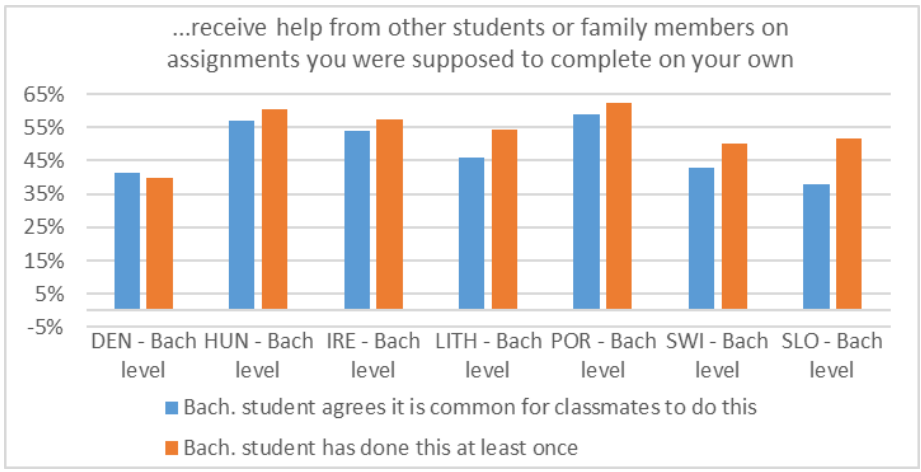
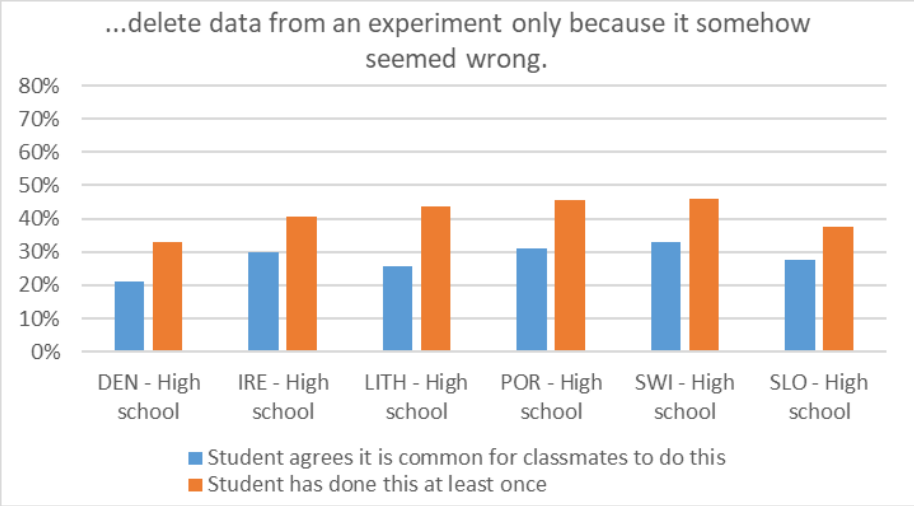


Figure 12. Perceptions of peer behaviour and whether the student has behaved similarly during their bachelor studies (Shares within groups are reported), bachelor students, stratified by country.

6.3. Perceived peer practice and own practice among high school students

In general, high school students seem to believe that questionable practice is quite common, irrespective of the topic (deleting data, copying shorter passages from other sources into their own texts, adding students who did not contribute as co-authors, and receiving help from others). The high school students perceived "...receiving help from other students or family members on assignments you were supposed to complete on your own" to be the most common questionable practice. This was also the most frequently reported own conduct, as 68% (Lithuania) to 77% (Slovenia) of the students across countries admitted to having performed this practice.



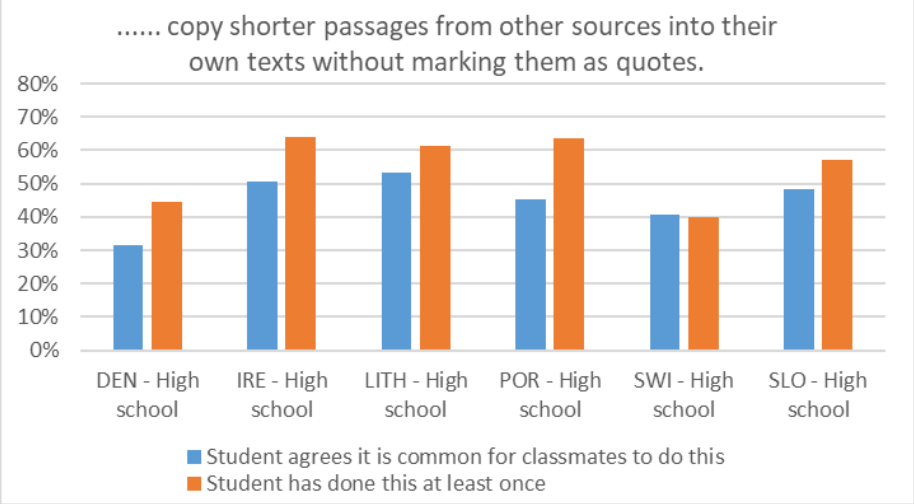


Figure 13. Perceptions of peer behaviour and whether the student has behaved similarly during high school (Shares within groups are reported), high school students, stratified by country.

6.4. Comparison across study levels

An overall view of the three figures (**Figures 11**, **12**, and **13**) shows a striking difference across study levels in the students' perception of others' behaviour vis-a-vis their own behaviour. Among PhD students, there is a particularly large difference between what students think others do and what they report to do themselves. It was relatively more common for students to believe that other researchers in the field participate in questionable practice (see the blue bars) compared to what they reported to do themselves (red bars). This difference between the perception of peer behaviour and own behaviour was lower among bachelor students, with no difference found for most countries and for many of the behaviours. Similarly, there was no difference between the perception of peer behaviour and own behaviour among high school students.

This could reflect a difference in how academic and research integrity norms are seen as part of one's own self-image. Whereas high school upper secondary and bachelor level students may see these norms as external, the PhD students may see them as part of their own identity as future scientists, and therefore on a personal level, distance themselves more from any violations.

6.5. Differences in perceived peer practice and own practice across gender identities

In this section we investigate whether the respondents' gender identity influenced the perception of peer practice and the tendency to engage in questionable practices. The questions laid out in the previous section (section 6.4) were used. At each study level and for each indicator variable, we conducted multivariable regression analyses to identify possible gender identity differences, while controlling for possible confounding from a number of factors. At PhD and bachelor level, the following predictor variables were used: age, country, type of data used, study direction, and gender identity. At high school level, the following predictor variables were used: age, country, and gender identity. Gender identity emerged as a significant predictor (at the 0.05 significance level) in relation to the following subset of questions:

- Citing sources that are not strictly relevant (for PhD students, perception of others and own behaviour)
- Adding other students as co-authors of group assignments, even though they did not contribute (for bachelor and high school students, own behaviour)
- Ignoring deviating material (for high school students, own behaviour)
- Receiving help from friends or family (for high school students, perception of others and own behaviour)

An overview of the gender distribution for the answers to these questions is reported in detail in Figure 14a-g below. Although most of the differences are small, it is striking that the male-identifying high school students were more than ten percentage points more likely than female-identifying students to add students who did not contribute to group assignments (but the opposite effect can be seen in the case of bachelor students). It is also worth noting that the male-identifying PhD students were more likely to believe that their peers cite sources that are not strictly relevant. It is also worth noting that gender identity is not associated with questions relating to power relations for PhD students (e.g. granting people in power undeserved co-authorship).

For all cases, the answers are reported using the four different gender categories used in the questionnaire (see section 3.5). It should be noted that the results for the sub-group who answered "none of the above" are uncertain as very few students identified with this category (see Table 3).

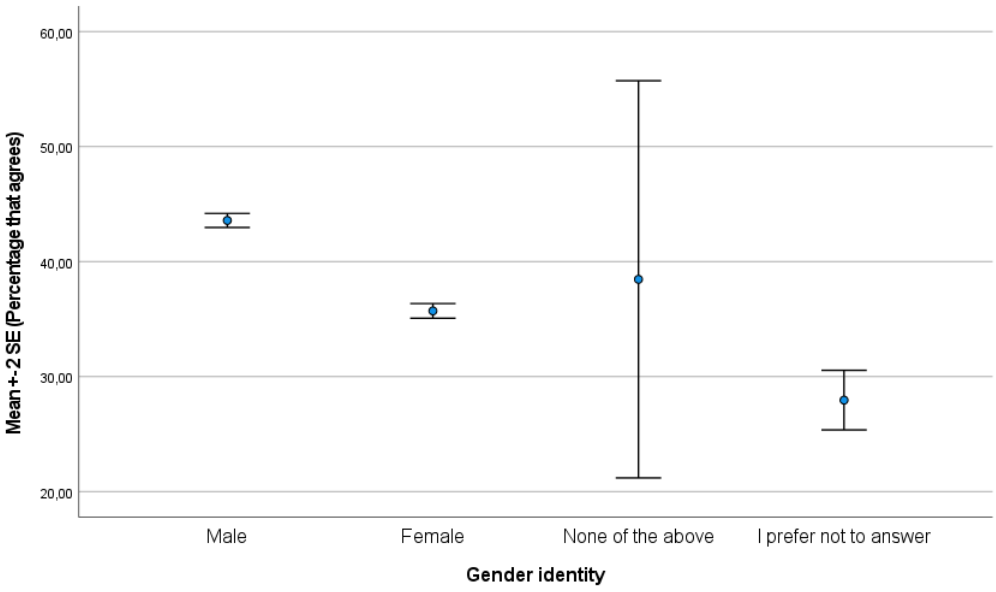


Figure 14a. Percentage of PhD students (reported as mean and SE of predicted probability) who agreed that it is common for researchers in the field to cite sources that are not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour – across countries and gender identity.

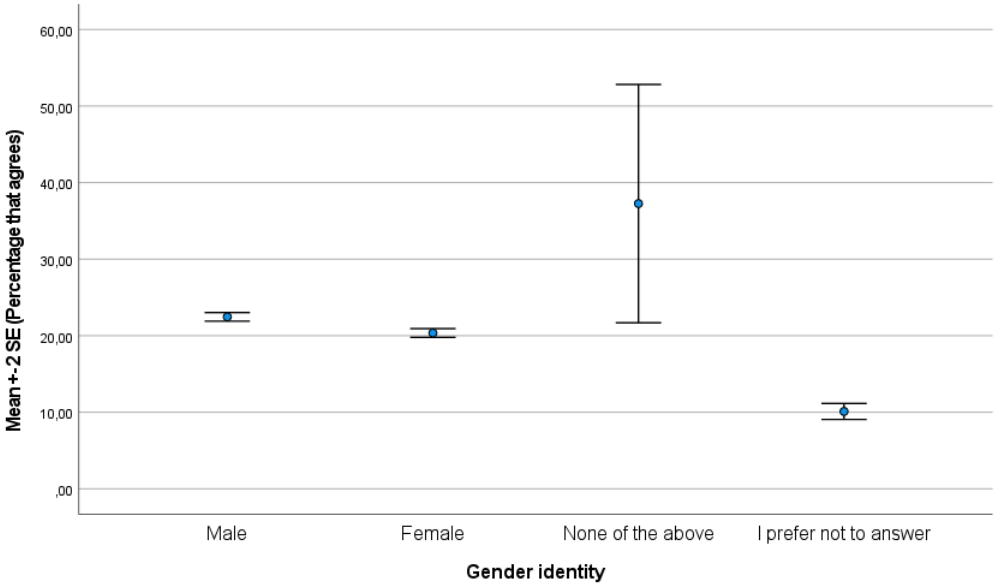


Figure 14b. Percentage of PhD students (reported as mean and SE of predicted probability) who at least once cited sources that were not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour – across countries and gender identity.

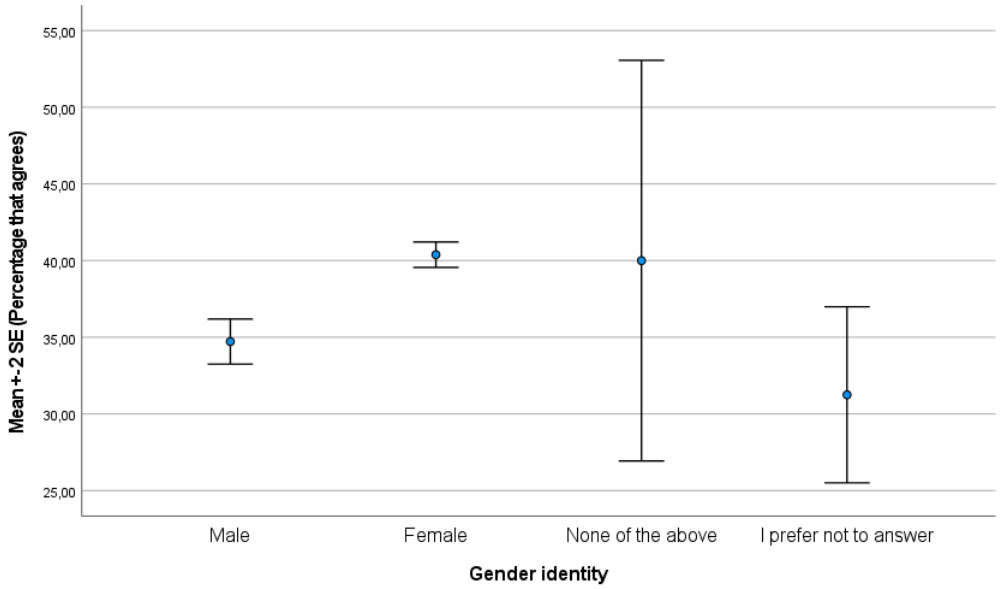


Figure 14c. Percentage of bachelor students (reported as mean and SE of predicted probability) who at least once added other students as co-authors of group assignments, even though they did not contribute – across countries and gender identity.

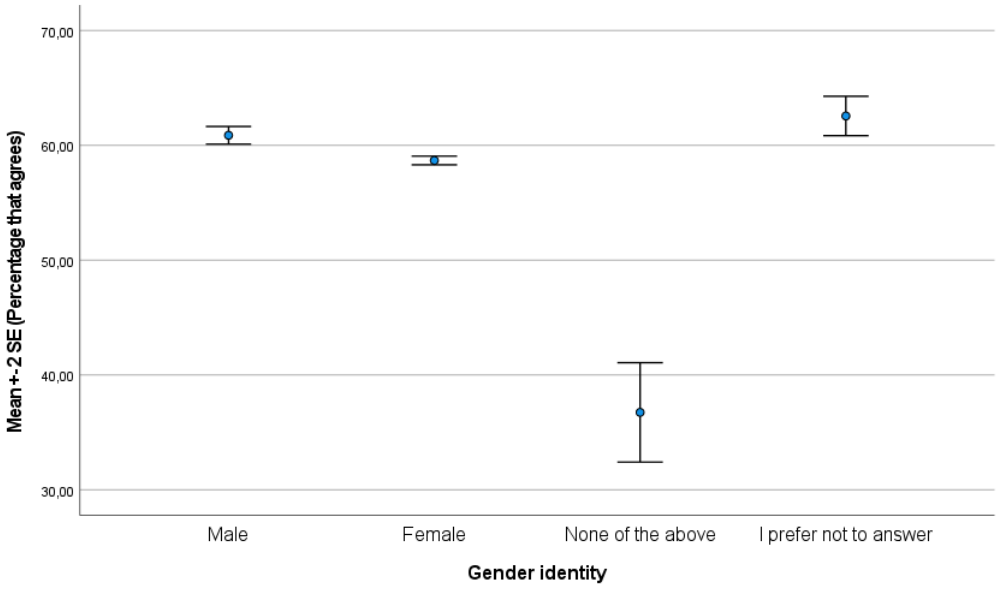


Figure 14d. Percentage of high school students (reported as mean and SE of predicted probability) who at least once ignored or deleted deviating data based on a gut feeling – across countries and gender identity.

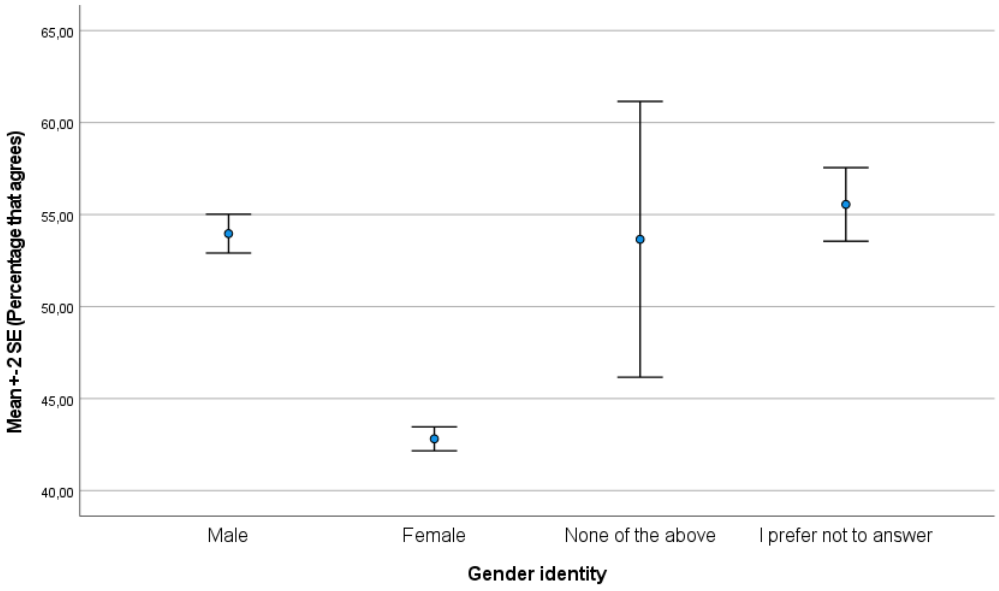


Figure 14e. Percentage of high school students (reported as mean and SE of predicted probability) who at least once added another student as a co-author of a group assignment even though the student did not contribute – across countries and gender identity.

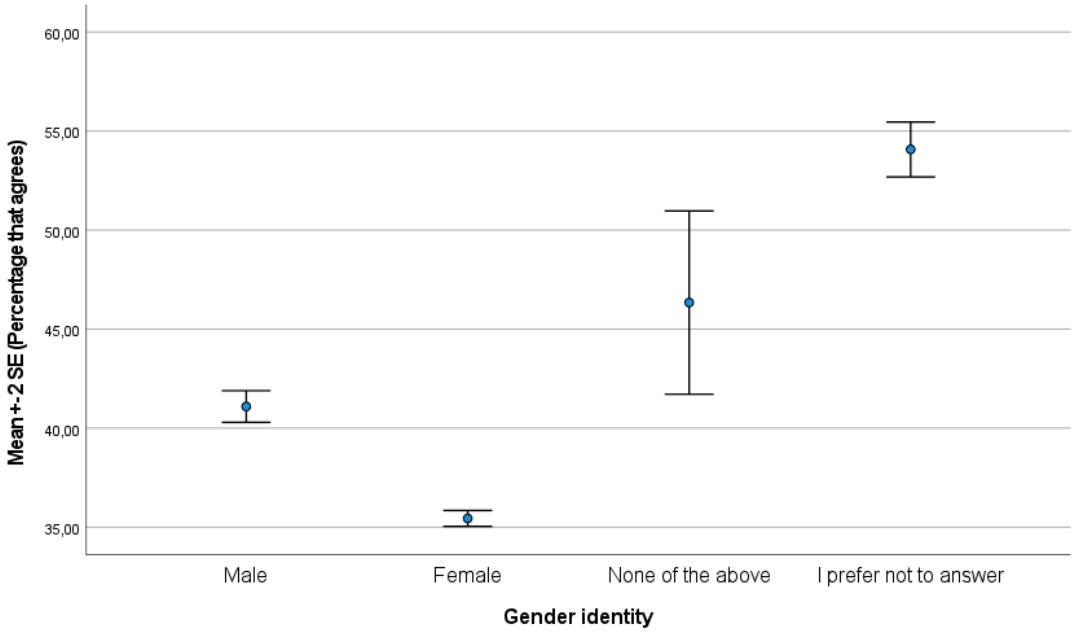


Figure 14f. Percentage of high school students (reported as mean and SE of predicted probability) who agreed that it is common for classmates to receive help from other students or family members on assignments they were supposed to complete on their own – across countries and gender identity.

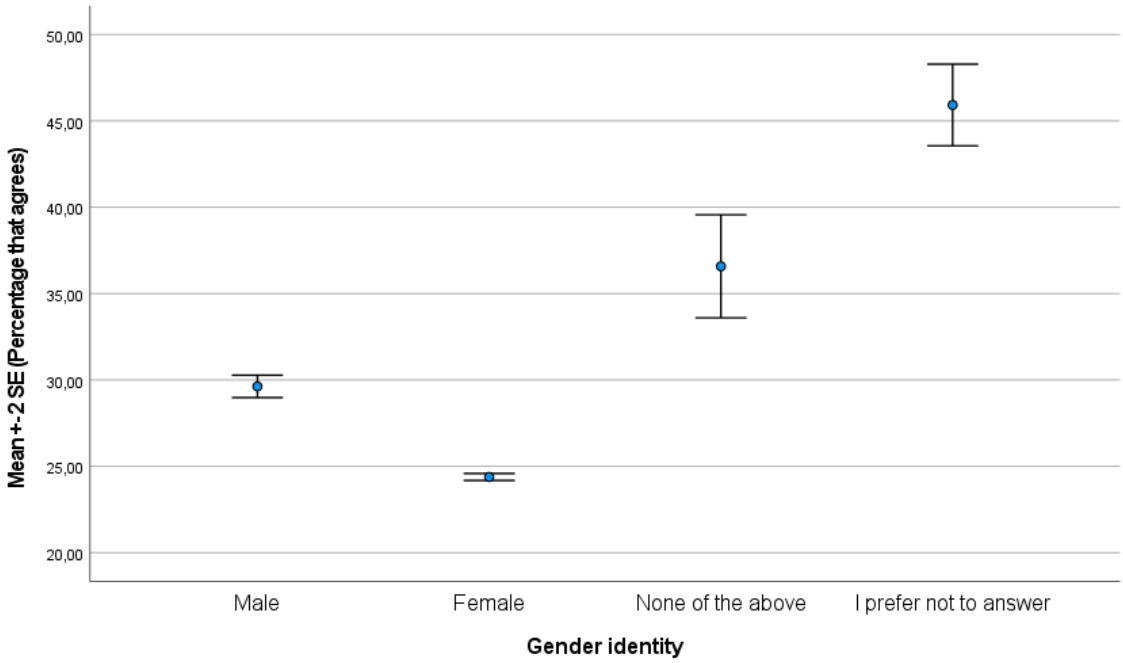


Figure 14g. Percentage of high school students (reported as mean and SE of predicted probability) who at least once received help from other students or family members on assignments they were supposed to complete on their own – across countries and gender identity.

7.Learning about integrity: courses and lectures taken, and stated needs

7.1. Courses and lectures taken

We asked the students about the kind of training they had received concerning rules and/or ethically correct behaviour in relation to our three core themes – see **Figure 15a**. There are major differences across countries. About two thirds of bachelor students in Hungary (64%), Portugal (72%) and Slovenia (64%) had not received any type of training in the form of a course, lectures, or e-learning sessions. This level is similar to Swiss high school students and in stark contrast to Danish and Swiss bachelor students, where only 28% and 18% stated that they had not received any form of training, respectively. A large proportion of the PhD students in Portugal (54%) and Switzerland (46%) had not taken any courses, while a very high proportion of the PhD students in Denmark (83%), Ireland (85%), and the Netherlands (75%) had received some training. Surprisingly, more bachelor students than PhD students in Switzerland had received training (and many of the bachelor students had even had dedicated courses), which is probably because a very high fraction of PhD students in the participating Swiss universities are internationals. Finally, many of the PhD students in Denmark had taken an entire course.

In Figure 15b, PhD students who answered "No" to having received training are shown by country and field of study. Again, there are clear differences between countries. It is also clear that there are striking differences across the different fields of study within the individual countries – natural science/engineering in Portugal in particular stands out here – yet there are no clear patterns such as a particular field of study being generally overrepresented.

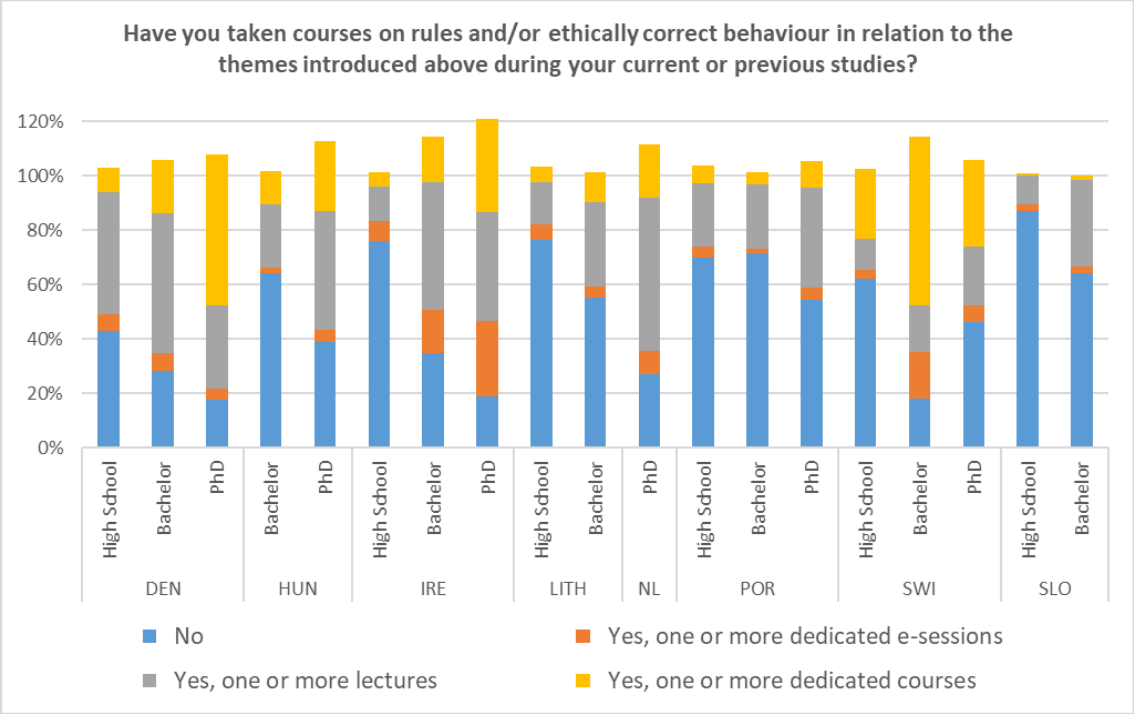


Figure 15a. Types of courses taken on cheating and misconduct (Stacked bars; responses can sum to more than 100% because this was a multiple response question), stratified by country and study level.

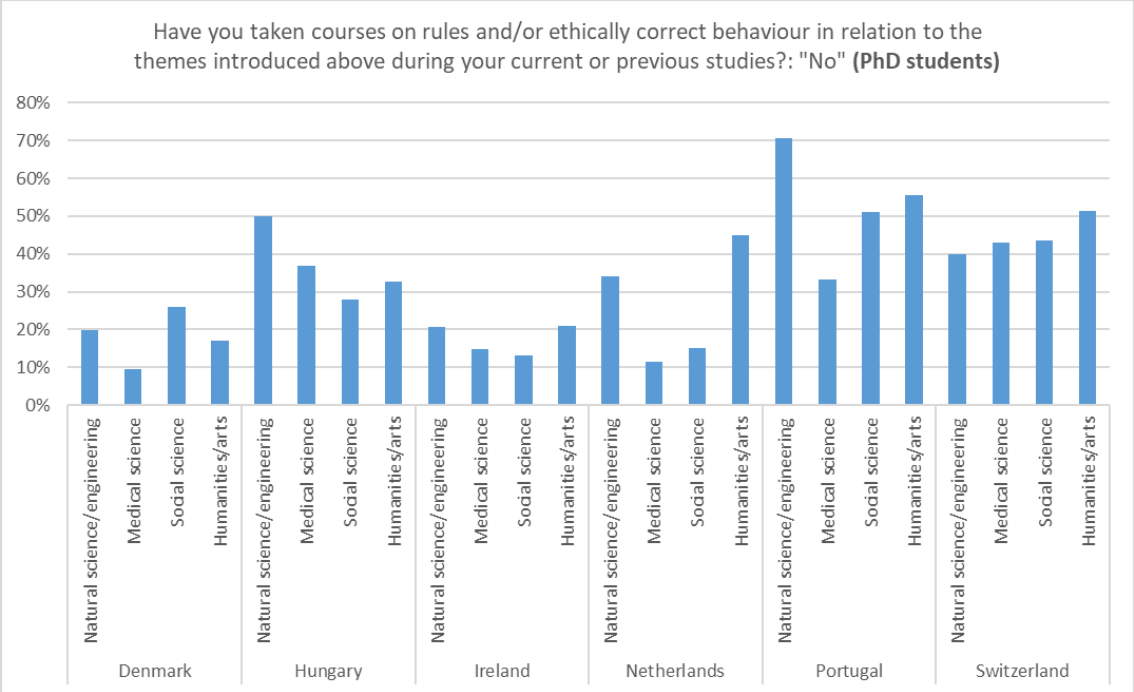


Figure 15b. PhD students who answered "No" to having received any kind of training, stratified by country and field of study.

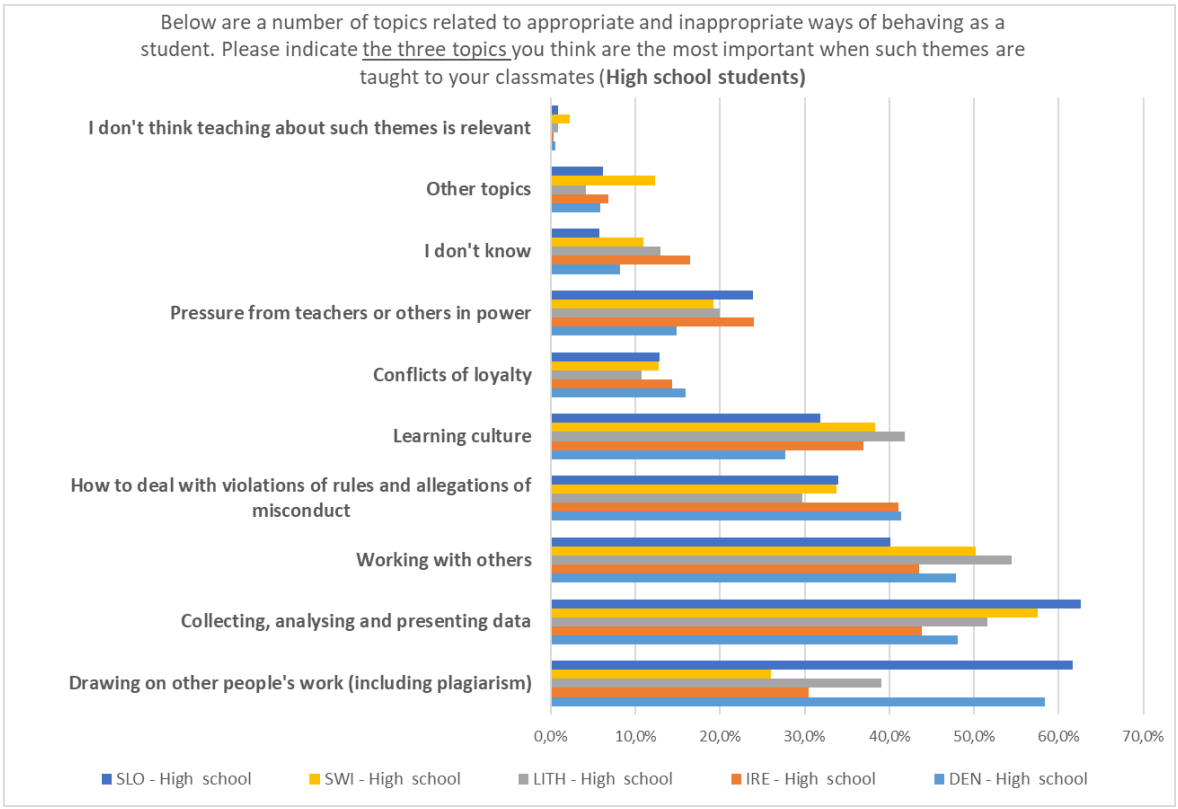
7.2. Stated needs

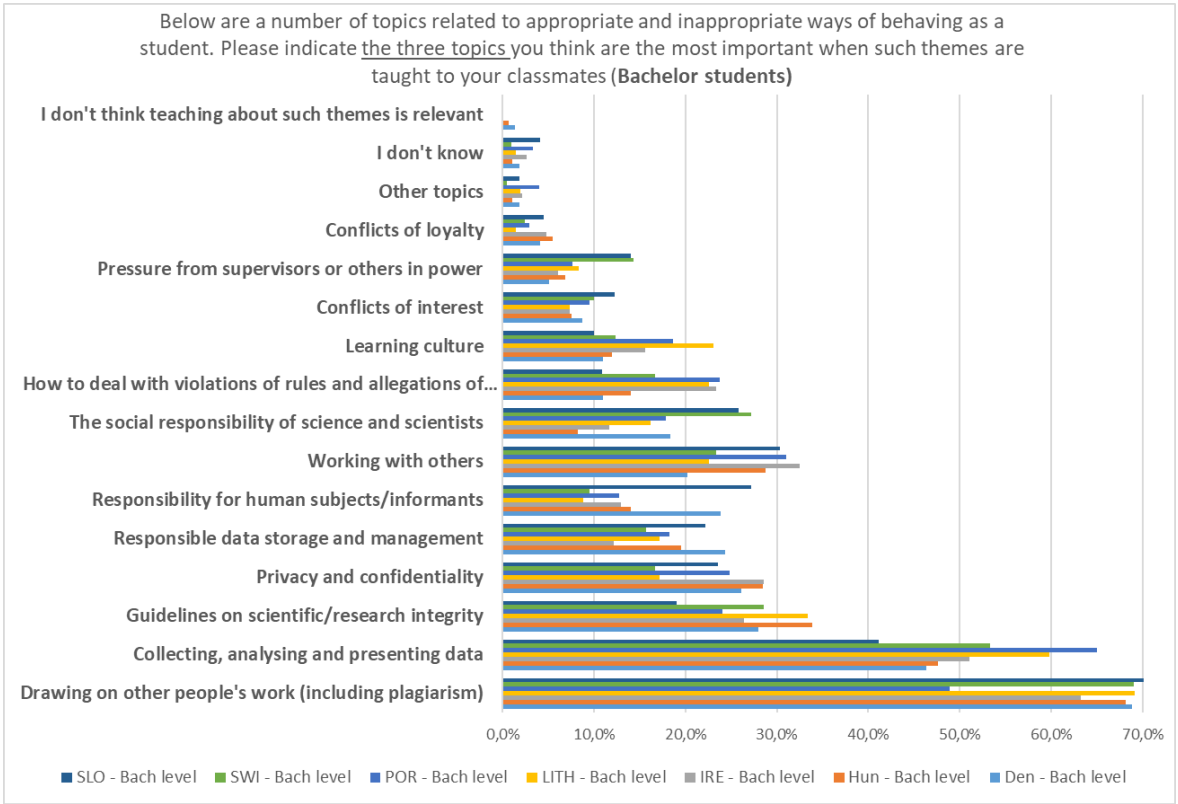
For each study level, students were presented with a list of possible topics within academic and research integrity and asked to pick the three topics they thought were "most important when such themes are taught to your [peers]". The results are laid out in **Figure 16**. The idea behind making students tick a maximum of three topics among many more possibly relevant topics was to ensure they would prioritise, and in effect pinpoint what they thought was most interesting and/or most needed. One downside is that if there are topics that students perceive to overlap slightly (e.g. "working with others" and "conflicts of loyalty"), the limited number of choices might lead them to choose only one these, even when they think both are important, in order to save the other options that they perceive to be clearly distinct.

Among *high school students*, working with others and data collection and analysis were the generally most important topics, while there were considerable differences across countries in the perception of how important it is to learn about plagiarism (ranging from 26% in Switzerland to 62% in Slovenia).

Among *bachelor students*, data collection and analysis and plagiarism were the most important topics. The fact that "working with others" was less prioritised can be put into perspective with the findings in Chapter 6 – that many bachelor students believed that their peers add people as co-authors even though they have not written anything, and that many of them have also done this themselves.

Among *PhD students*, data collection and analysis and plagiarism were also the clearly most important topics, although there was also an expressed need to learn about integrity guidelines. It is also worth noting that even though many PhD students thought it common for researchers in the field to allow people in power to become co-authors (even though they have not contributed), and they had also allowed this themselves (cf. results in Chapter 6), we can see that PhD students to a lesser extent prioritised learning about "working with others" and "pressure from superiors or others in power".


Below are a number of topics related to appropriate and inappropriate ways of behaving as a student. Please indicate the three topics you think are the most important when such themes are taught to your classmates (High school students)
I don't think teaching about such themes is relevant
Other topics
I don't know
Pressure from teachers or others in power
Conflicts of loyalty
Learning culture
How to deal with violations of rules and allegations of misconduct
Working with others
Collecting, analysing and presenting data
Drawing on other people's work (including plagiarism)
0,0% 10,0% 20,0% 30,0% 40,0% 50,0% 60,0% 70,0%
SLO - High school
SWI - High school
LITH - High school
IRE - High school
DEN - High school


Below are a number of topics related to appropriate and inappropriate ways of behaving as a student. Please indicate the three topics you think are the most important when such themes are taught to your classmates (Bachelor students)
I don't think teaching about such themes is relevant
I don't know
Other topics
Conflicts of loyalty
Pressure from supervisors or others in power
Conflicts of interest
Learning culture
How to deal with violations of rules and allegations of...
The social responsibility of science and scientists
Working with others
Responsibility for human subjects/informants
Responsible data storage and management
Privacy and confidentiality
Guidelines on scientific/research integrity
Collecting, analysing and presenting data
Drawing on other people's work (including plagiarism)
0,0% 10,0% 20,0% 30,0% 40,0% 50,0% 60,0% 70,0%
SLO - Bach level
SWI - Bach level
POR - Bach level
LITH - Bach level
IRE - Bach level
Hun - Bach level
Den - Bach level

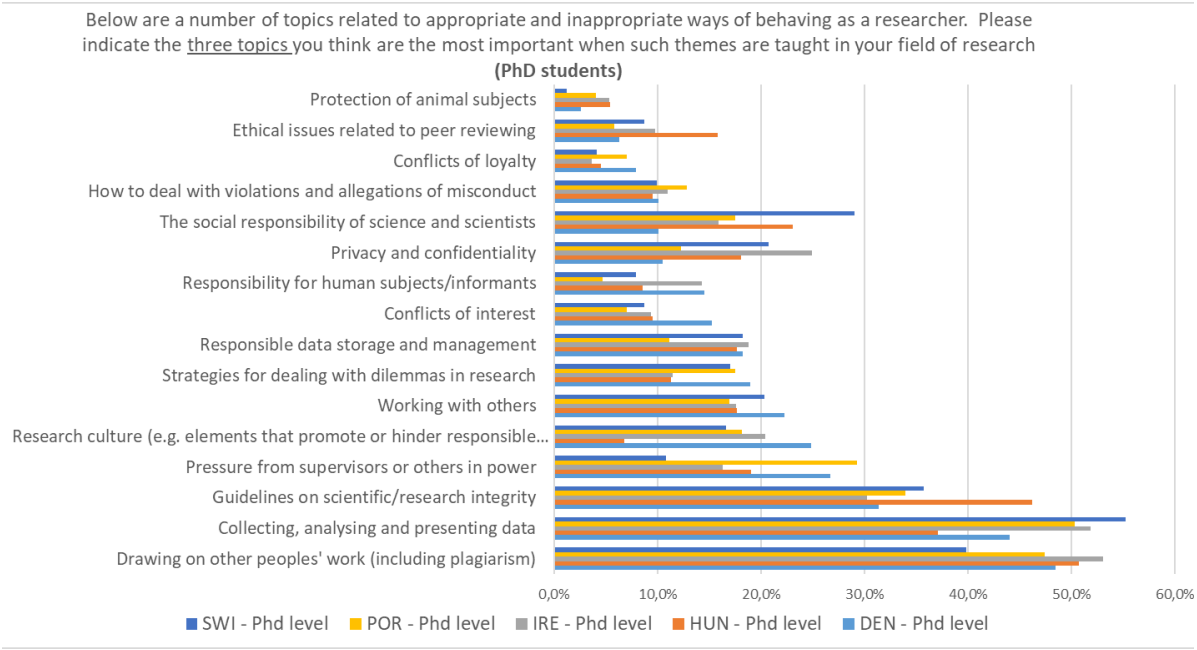


Figure 16. Topics that students would like to learn about (Reported as share of students for each topic) – three graphs for each study level, stratified by country when possible.

8. Does RI training reduce the rate of questionable practice among PhD students?

Approximately 45% of the students had received some form of dedicated RI training at the time of answering the survey (Table 6). Of these, most had participated in less than 1 working day and only about 6% had more than a working week.

Table 6. Amount of time spent on courses about rules and/or ethically correct behaviour in relation to the themes introduced

	Frequency	Percent	Cumulative
Do not recall	435	7.9	7.9
No time spent	2744	49.8	54.0
Less than 1 working day	735	13.3	68.5
1 full working day	456	8.3	77.5
1-3 full working days	586	10.6	89.0
4-5 full working days	202	3.7	93.0
1-2 working weeks	155	2.8	96.1
More than 2 working weeks	199	3.6	100.0

To investigate the effect that teaching has on the likelihood that students would engage in the questionable practices discussed in Chapter 5, we constructed an index based on the six questions in Figure 11 (section 6.2 above) about what sort of behaviour PhD students believed to be common among peers in their field of study. This index measured *perceived questionable practice among peers*. The index has acceptable internal validity (avg. Cronbach's alpha 0.78, ranging between 0.75 for natural science/engineering and 0.81 for humanities/arts). We constructed the index by summing responses to the six questions in Figure 11. The index ranges from 0 to 1, where 0 indicates a very low, 1 indicates a very high, and 0.5 indicates a medium level of propensity to think that peers are involved in the six examples of misconduct (Mean 1.76 (SD 1.59); n=1387).

Using this, we constructed a correlation matrix (Table 7) displaying the level of correlation between *perceived questionable practice among peers*, *own questionable practice* (all six questions), *self-reported knowledge* (from Chapter 3), *courses taken* (three variants) and amount of *time spent on these courses* for

PhD students. The matrix shows the correlation between these twelve variables. The matrix is a Spearman's correlation to account for the fact that several of the variables are not continuous or normally distributed. To read the matrix, a perfect positive correlation between two variables will result in the correlation coefficient 1, no correlation will result in the coefficient 0 and a perfect negative correlation in the coefficient -1.

The variable "own misconduct" reported in the table is based only on answers ranging from "No" to "Yes, many times" (whereas "Not Applicable", "Don't know", and "Prefer not to answer" are not included).

Table 7. Spearman's correlations between own questionable practice, type and amount of RI training, and perceived questionable practice among peers (n=1217-1595) – for PHD students

Have you ever…	Perceived questionable practice among peers	One or more dedicated courses	One or more lectures	One or more e-learning sessions	Time spent on courses
deleted deviating data points based on a gut feeling that they were inaccurate?	0.193**	-0.067*	-0.063*	0.014	-0.026
performed a misleading or dubious statistical analysis in order to achieve results that a reviewer would accept?	0.183**	0.029	0.015	0.015	-0.026
kept inadequate records of parts of your work that should be documented?	0.287**	0.006	0.040	0.006	0.001
copied shorter passages from other sources into your own text without marking them as quotes?	0.147**	0.024	-0.036	0.041	0.059*
allowed research group leaders, supervisors or others in power to become co-authors of papers, even though they did not make a significant contribution to them?	0.356**	-0.009	-0.000	0.013	0.022
cited sources that were not strictly relevant in order to please a reviewer or in the hope that the author of the source might return the favour?	0.321**	-0.003	0.022	-0.009	0.018

In general, the table shows a positive correlation between the students' own propensity to conduct dubious acts and the perceived questionable practice of others (the magnitude of the associations ranges from small to medium), while there is no discernible correlation between students' own behaviour and the type and amount of time spent on research integrity courses.

9.References

ALLEA (2017). *The European Code of Conduct for Research Integrity - revised edition*. Available from: https://www.allea.org/publications/joint-publications/european-code-conduct-research-integrity/

D2.1 (2019). *Preliminary report on combined results of questionnaire and qualitative interviews*. INTEGRITY. URL= http://h2020integrity.eu/about-us/documentation/

De Vries, R., Anderson, M.S, Martinson, B.C (2006). Normal Misbehavior: Scientists Talk about the Ethics of Research. *Journal of Empirical Research on Human Research Ethics*, *1*(1), 43–50. https://doi.org/10.1525/jer.2006.1.1.43

Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PloS One*, *4*(5), e5738–e5738. https://doi.org/10.1371/journal.pone.0005738

Goddiksen, M. et al (2020). Good friend or good student: An interview study of perceived conflicts between personal and academic integrity in three European countries. Accountability in Research. Doi: https://doi.org/10.1080/08989621.2020.1826319

ICMJE (2017). Defining the Role of Authors and Contributors. *icmje.org* Available from: http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html [visited June, 2020].

Jensen, L.A, Arnett, J. J, Feldman, S.S, Cauffman, E. (2002). It's Wrong, But Everybody Does It: Academic Dishonesty among High School and College Students. *Contemporary Educational Psychology*, 27(2), 209–228. https://doi.org/10.1006/ceps.2001.1088

Johansen, M. & Christiansen, F. (2020). Handling Anomalous Data in the Lab: Students' Perspectives on Deleting and Discarding. *Science and Engineering Ethics 26*, 1107-1128

Knobloch-Westerwick, S., Glynn, C. J. Huge, M. (2013). The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest. *Science Communication, 35*(5), 603–625. https://doi.org/10.1177/1075547012472684

Maliniak, D. Powers, R., Walter, B. F. (2013). The Gender Citation Gap in International Relations. *International Organization, 67*(4), 889–922. https://doi.org/10.1017/S0020818313000209

Martinson, B. et al (2005). Scientists behaving badly. *Nature 435*, 737-738

McCabe, D. (1999). Academic dishonesty among high school students. *Adolescence*, *34*(136), 681–687.

McCabe, D., Treviño, L.K., Butterfield, K. D.. (2001). Cheating in Academic Institutions: A Decade of Research. *Ethics & Behavior, 11*(3), 219–232. https://doi.org/10.1207/S15327019EB1103_2

Monroe, K., Ozyurt, S., Wrigley, T. Alexander, A. (2008). Gender Equality in Academia: Bad News from the Trenches, and Some Possible Solutions. *Perspectives on politics*, *6*(2), 215–233. https://doi.org/10.1017/S1537592708080572

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). *Science faculty's subtle gender biases favor male students*. Yale University.

Reuben, E., Sapienza, P., Zingales, L. (2014). How stereotypes impair women's careers in science. Proceedings of the National Academy of Sciences. *PNAS, 111*(12), 4403–4408. https://doi.org/10.1073/pnas.1314788111

Roig, M. (1997). Can undergraduate students determine whether text has been plagiarized? *The Psychological Record 47*, 113-22