ACCEPTED MANUSCRIPT

Research skills in upper secondary education and in first year of university

Louise Maddens^{ab}, Fien Depaepe^{ab}, Rianne Janssen^c, Annelies Raes^{ab}, Jan Elen^a

^aFaculty of Psychology and Educational Sciences, Centre for Instructional Psychology and Technology, KU Leuven Campus Kulak Kortrijk, Belgium

^bITEC, imec research group at KU Leuven, imec, Leuven, Belgium

^cFaculty of Psychology and Educational Sciences, Centre for Educational Effectiveness and Evaluation, KU Leuven, Belgium

This is a PDF file of an unedited manuscript that has been accepted for publication. The manuscript has undergone copyediting, typesetting and review of the resulting proof before it was published in its final form. The Version of Record of this manuscript has been published and is available in Educational Studies (January 28 2020)

<https://www.tandfonline.com/doi/full/10.1080/03055698.2020.1715204>

Research skills in upper secondary education and in first year of university

Louise Maddens^{ab}, Fien Depaepe^{ab}, Rianne Janssen^c, Annelies Raes^{ab}, Jan Elen^a

^aFaculty of Psychology and Educational Sciences, Centre for Instructional Psychology and Technology, KU Leuven Campus Kulak Kortrijk, Belgium

^bITEC, imec research group at KU Leuven, imec, Leuven, Belgium

^cFaculty of Psychology and Educational Sciences, Centre for Educational Effectiveness and Evaluation, KU Leuven, Belgium

Abstract: The acquisition of research skills is an important goal of secondary education. Still, insights on what learner characteristics stimulate and hamper research skills are lacking. Therefore, this study investigates the relationship between students' research skills (assessed by means of the Leuven Research Skills Test (LRST) and students' gender, educational track, grade, autonomous and controlled motivation. The LRST was administered to 405 11th and 12th grade students, and to 156 first-year university students in Flanders. Multiple regression analyses reveal a significant relationship of grade and track with students' research skills in upper secondary education. Also in higher education, a significant relationship was found between students' former educational track and their research skills. In both groups, no significant differences were found in research skills based on gender and on students' controlled and autonomous motivation. This study contributes to future research by stressing the importance of curricular characteristics for fostering student's research skills.

Keywords

Research skills, secondary education, learner characteristics, gender, controlled motivation, autonomous motivation, tracks, grade

Introduction

In the 21st century, education aims at empowering students to acquire a broad set of skills with regard to (a) the understanding of the acquisition of scientific knowledge in several disciplines (and the role of specific concepts and methods in this acquisition), and (b) the examination of the validity and the relevance of science-related claims, methods and concepts. These scientific concepts and methods can then be used to identify and to solve (science, work and daily life-related) problems (Fischer et al., 2014; Opitz, Heene & Fischer, 2017). In this article, these skills are referred to as research skills. It is argued by the National Research Council (2010) that the engagement of students in these scientific processes not only builds science proficiency, but also may help to develop other important 21st century skills (such as complex communication skills and nonroutine problem-solving skills). The international attention for research skills proficiency is expressed in highly valued outcomes in curriculum documents for upper secondary education (Engelmann, Neuhaus, & Fischer, 2016; Fischer et al., 2014; Osborne & Dillon, 2008). Accordingly, three specific objectives related to research are formulated for students in the 11th and 12th grade in the Flemish educational guidelines for the general track¹ (EACEA, 2018). These curriculum objectives state that students have to be able (a) to orientate themselves towards a research problem by collecting, organizing and editing information in a well-considered way; (b) to prepare, carry out and evaluate research, and (c) to report the research results and conclusions and confront them with alternative points of view (Departement Onderwijs en Vorming, 2017).

¹ In Flanders, four different types of education are offered from the second stage of secondary education onwards (EACEA, 2018) (general secondary education, technical secondary education, secondary education in the arts and vocational secondary education).

The attention for the acquisition of research skills comes along with an emerging research agenda focusing on the assessment (Opitz, Heene, & Fischer, 2017) and support (Engelmann et al., 2016; Fischer et al., 2014; Linn, Davis, & Bell, 2004) of students' research skills proficiency. Still, no conclusive answers can be formulated with respect to the question what learner characteristics affect adolescents' research skills proficiency. Therefore, and as a baseline for future research, this study aims to assess students' research skills proficiency (based on an administration of the Leuven Research Skills Test or LRST) in relation to students' learner characteristics.

In the following sections, first, the term research skills is conceptualized. Second, several learner characteristics associated with research skills proficiency are addressed. Third, the method and the results are presented. To conclude, some important findings and pitfalls of this study are discussed.

Theoretical framework

Conceptualization of research skills

The international shared conviction that the acquisition of research skills is an important goal of (upper) secondary education results in a variety of terms and definitions referring to these skills across and within disciplines (Kestens, Elen, & Verburgh, 2016), such as: scientific reasoning skills (Engelmann et al., 2016; Fischer et al., 2014; Opitz, Heene, & Fischer, 2017; Zimmerman, 2000), scientific literacy (Norris, Phillips, & Burns, 2014) or research methods skills (Earley, 2014). In the present study, the term research skills is used, as it adequately reflects the target concept as a broad set of skills (not merely referring to reasoning, literacy or research methods skills). In this manuscript, this term will be used solely when referring to a comprehensive set of skills (Fischer et al., 2014; Opitz et al., 2017). This study draws upon the theoretical framework of Fischer et al. (Fischer et al., 2014; Opitz et al., 2017), deconstructing research skills into eight scientific activities in which students have to be able to engage, namely: (a) problem identification, (b) questioning, (c) hypothesis generation, (d) construction and redesign of artefacts, (e) evidence generation, (f) evidence evaluation, (g) drawing conclusions and (h) communicating and scrutinizing. Table 1 offers a short description of each of these skills.

Table 1
Conceptualization of research skills (Fischer et al., 2014; Opitz et al., 2017)

Skill name	Skill description
Problem identification	By analyzing the situation (observing a shortcoming concerning the available explanation of a specific problem), a problem representation is built
Questioning	During questioning, an applicable research question is formulated or identified based on the representation developed during problem identification
Hypothesis generation	The process of hypothesis generation involves (a) formulating possible hypotheses according to scientific standards, and (b) a good understanding of the relationship between variables
Construction and redesign of artefacts	This skill includes the construction, testing and revision of some kind of artefact (for example interview guidelines or a questionnaire). This artefact is tested in an authentic environment (Kolodner, 2007)
Evidence generation	During evidence generation, various approaches such as observation, experimenting, interviewing or surveying lead to the production of evidence
Evidence evaluation	The process of evidence evaluation involves the evaluation of pieces of evidence against the available data, or against the methods used to collect the evidence

Drawing conclusions	This skill involves the critical analysis and interpretation of data in relation to the method used and to the rules and criteria of the research discipline
Communicating and scrutinizing	By communicating and scrutinizing, findings and methods are shared, presented or discussed within a team or a broader community by means of objective, scientific language

Learner characteristics associated with research skills proficiency

As depicted in Figure 1, this study focuses on the link between learner characteristics and research skills. More particularly, it is investigated whether (and to what extent) students' scores on the LRST (see Method section) are influenced by specific learner characteristics. Based on Drachler and Kirschner (2012), a distinction is made between *personal* learner characteristics, *academic* learner characteristics, *cognitive* learner characteristics and *affective* learner characteristics. As a personal learner characteristic, students' gender is included as an independent variable. Academic characteristics relate to education or learning context (Drachler & Kirschner, 2012). More specifically, students' track and grade are included as academic independent variables. In addition, students' autonomous and controlled motivation towards research are the two affective independent variables. The scores on the LRST are included in the model as a cognitive dependent variable.

In what follows, a short overview is presented of the existing literature on the relationship between these variables. However, it should be noted that the literature investigating these learner characteristics in relationship with students' research skills proficiency (as defined in this paper) is scarce. Nonetheless, many articles and policy documents focus on the relationship of learner characteristics with *science* proficiency. Although we argue that science proficiency (or scientific reasoning skills) and research skills proficiency do not entirely have the same meaning; literature on science proficiency is discussed from the perspective of research skills as conceptualized above in this paper. The Programme for International Student Assessment (PISA) for scientific literacy, for example, assesses how well students (a) explain scientific issues, (b) explain phenomena scientifically and (c) use scientific evidence (OECD, 2006). Although some aspects of the PISA-assessments do not relate to our operationalization of research skills proficiency (such as the focus on conceptual knowledge in natural sciences) some important similarities are found (such as the focus on the interpretation of findings, the identification of issues, drawing conclusions, and so on). Obviously, the analogies between the PISA-tests (and other comparable tests) on the one hand and the LRST on the other hand must be interpreted with caution.

Figure 1. Conceptual framework (based on and adapted from Drachsler & Kirschner, 2012).

Research skills proficiency by gender. In 1988, Eylon and Linn (1988; see also Linn & Hyde, 1989) argued that research findings regarding gender differences in science vary with science topic, which makes generalizations hard. Overall, these authors indicate that, if the level of courses being taken is controlled for, studies of gender differences in science learning report few differences between males and females (Eylon & Linn, 1988). The pattern of few and inconclusive findings regarding gender differences in science proficiency can still be observed today (Raes, 2015). In 2006, PISA (OECD, 2007, p. 61), concluded that “overall, gender performance in science is remarkably even, with only a few OECD countries showing significant gender differences”. A more recent administration does indicate differences between boys’ and girls’ research skills proficiency: according to PISA (OECD, 2018), there are more male top performers than female top performers in science (except for Finland). Still, these gender differences in science performance seem to be small (OECD, 2018). A similar pattern in PISA-scores is found in Belgium: boys seem to outperform girls within the group of top performers (with a significant difference of 3,9%), and girls seem to outperform boys in the group of low achievers (with a non-significant difference of 1,5%) (OECD, 2018). On average, for Belgian students, PISA (OECD, 2018) indicates a mean score for boys of 508, and a mean score of 496 for girls, indicating a rather small proficiency difference between boys and girls. As such, the results regarding gender differences in scientific reasoning skills are inconsistent and thus, the question should be raised whether it still makes sense to include gender as a possible predictor of scientific reasoning or research skills proficiency. In an attempt to take a stand in the ongoing debate, and in order to decide whether the focus on gender differences is still worthwhile to consider in the specific case of research skills proficiency in upper secondary education and in first year of university, gender was included as an independent variable.

Research skills proficiency by academic track. Already in 1981, Eder argued that one should consider the influence of the environment in which learning occurs on students’ academic achievements (Eder, 1981). One of the characteristics of this environment is whether students are

assigned to different tracks (this practice is mainly referred to as tracking) (Eder, 1981, Raes, 2015). In Flanders, from the age of 15, students are tracked in different trajectories¹. As such, one could argue that students in technical or vocational tracks, are possibly less confronted with science or research related courses during their trajectory, compared to students in general secondary education. In line with this assumption, OECD points out that “students enrolled in vocational programmes receive, on average, around 80 minutes less per week of regular science instruction than their peers in academic tracks.” (OECD, 2016, p. 233). In Belgium, there is a difference of 29.6% in science learning time at school between students enrolled in a vocational track, and students enrolled in an academic track. In the vocational tracks, 68.4% of the students indicate that they take at least one science lesson per week at school, while, in the academic tracks, 97.9% indicates to do so (OECD, 2016). This can be seen as a logical consequence of the characteristics of the Flemish curricula, as reflected in written educational standards. While three specific educational objectives relating to science education are formulated for the general tracks, these objectives are absent in the vocational and technical tracks, and in the curriculum documents for secondary education in the arts. The relationship between academic tracks and students’ research skills proficiency for the specific context of Flemish education is further investigated in this paper.

Research skills proficiency by grade. Based on the PISA-report (OECD, 2016), a growth in research skills proficiency is expected after one year of education. This has been seen in the case of Germany, where a longitudinal study assessed a cohort (9th-graders) in 2003, and in 2004 (as 10th-graders). The authors showed that, in one year, students gained about 21 score points in a science test (Prenzel et al., 2007). Based on this finding and other similar studies, the PISA report equates 30 score points with about one year of schooling (OECD, 2016). Consequently, for the specific case of this study, one would expect students in 12th grade to outperform students in 11th grade (after one year of schooling). Of course, the assumption here is that the groups receive the same amount of instruction in research skills in both grades (which is not yet thoroughly investigated). In addition, it should be noted that an aging effect can interfere with the effect of grade. Unfortunately, no studies were found regarding grade effects controlling for aging effects and instruction time.

Research skills proficiency by autonomous and controlled motivation. In line with self-determination theory, previous research distinguishes between autonomous and controlled study motivation (SDT) (Ryan & Deci, 2000; Vansteenkiste, Lens, & Deci, 2006; Vansteenkiste, Sierens, Soenens, Luyckx, & Lens, 2009). By autonomous motivation we understand the intrinsic or identified regulation towards a task (be it because the task is considered interesting, or because the task is considered personally relevant). Controlled motivation refers to the external or introjected regulation towards the task (as a consequence of internal or external pressure) (Vansteenkiste et al., 2009). As such, each motivational type consists of two subcomponents. Autonomous motivation consists of intrinsic motivation and well-internalized extrinsic motivation (or identified regulation), and controlled motivation consists of external regulation and introjected regulation. Intrinsic motivation refers to excitement, enjoyment and interest inherent to learning, while internalized extrinsic motivation refers to perceived personal relevance of learning. In addition, external regulation relates to external pressured regulation, while introjected regulation relates to internal pressured motivation (Vansteenkiste et al., 2009, see Table 2 for a sample item related to each construct). The existence of these two factors (autonomous and controlled motivation) has been repeatedly confirmed in previous research by means of exploratory and confirmatory factor analyses (see for example: De Naeghel, Van Keer, Vansteenkiste, & Rosseel, 2012). Vansteenkiste et al. (2009) identified four motivational profiles in high school and college students by means of a cluster analysis: a good quality motivation group (high autonomous, low controlled); a poor

quality motivation group (low autonomous, high controlled), a low quantity motivation group (low autonomous, low controlled), and a high quantity motivation group (high autonomous, high controlled), and suggested that, if the motivation is controlled rather than autonomous in nature (and thus of poorer quality), higher levels of motivation do not necessarily lead to more desirable outcomes (Ryan & Deci, 2000; Vansteenkiste et al., 2006, 2009). The authors conclude that the good quality motivation group displayed the most optimal learning pattern. In addition, it is argued that controlled motivation predicts lower achievement (Soenens & Vansteenkiste, 2005), while autonomous motivation is associated with higher achievement (Soenens & Vansteenkiste, 2005). In this paper, the relatedness of autonomous motivation and controlled motivation with students' research skills proficiency is studied.

Table 2

Autonomous and controlled motivation (based on Vansteenkiste et al., 2009)

Type of motivation	Example of an item
Autonomous motivation	
Intrinsic motivation	I am studying ... because I am highly interested in doing this.
Identified regulation	I am studying ... because it is personally important to me.
Controlled motivation	
External regulation	I am studying ... because I am supposed to do so.
Introjected regulation	I am studying ... because I want others to think I am smart.

The current study. Relying on the aforementioned theoretical framework and (research) findings, the specific research question to be answered in this study is: how does research skills proficiency vary as a function of gender, grade level, academic track, autonomous motivation and controlled motivation? In order to cross-validate our study, this research question is investigated with two different cohorts, namely students in upper secondary education, and students in the first year of university.

Method

Participants and procedure. Two different cohorts participated in the study. First, the LRST was administered to 405 students in 11th (41%) and 12th (59%) grade (in February 2018). Nineteen respondents were excluded from the analyses because they did not complete the whole test. The cohort consisted of students in three types of secondary education: general secondary education (behavioral sciences track: 55% and non-behavioral sciences track: 32%), technical secondary education and secondary education in the arts (13%). The sample consisted of 68% girls.

The second cohort consisted of 156 first-year university students, enrolled in an Educational Sciences bachelor's program. Three students were excluded from the analyses because they were enrolled in a non-comparable study program, and one student was excluded because he or she did not complete the entire test. Students' mean age was 18,54 years ($SD = 0.67$). The sample consisted of 85% girls, which is in line with the expectations of prevalence of sexes in the specific study program. In their 12th grade of secondary education, 43% of the cohort of university students was enrolled in general secondary education, behavioral sciences track, 42% of the cohort was enrolled in general secondary education, non-behavioral sciences track, 9% followed technical secondary education and 6% was enrolled in secondary education in the arts.

All students provided informed consent for their participation. The total administration (on computer) lasted maximum 100 minutes. On average, students in upper secondary education spent 85 minutes on the test, and students in university spent 62 minutes on the test. A supervisor was present during data collection.

Instrument. Students completed the 37-item computer-based Leuven Research Skills Test (LRST), (Maddens, Depaepe, Janssen, Raes, & Elen, 2019), which consists of a mixture of multiple-choice and open end questions related to research practices in a social sciences context. In the LRST, research skills are operationalized as a set of 37 items relating to the eight scientific activities formulated by Fischer et al. (2014) (see: Theoretical Framework). The test indicates good internal consistency ($\alpha = .81$). In addition, students filled in a motivation questionnaire (The Academic Self-Regulation Scale, Vansteenkiste et al., 2009; adapted version of Academic Self-Regulation Scale, Ryan & Connell, 1989). The 16-item scale contains eight items (each answered by means of a likert-type scale) per subscale (see: theoretical framework), and also indicates good internal consistency (with α -levels from .83 for controlled motivation to .91 for autonomous motivation).

Analysis. In order to investigate the relationship of students' grade, track, gender, controlled motivation and autonomous motivation (as independent variables) with research skills proficiency (as a dependent variable), a multiple regression analysis was conducted. First, two dummy variables were created for students' grade and students' gender. Since students in higher education follow the same grade (while students in upper secondary education follow either the 11th or the 12th grade), grade was not included in the regression model for the second cohort. Second, students' track was coded as two dummy variables (one for general education, behavioral sciences and one for general education, non-behavioral sciences). Students in the technical tracks or education in the arts served as the reference category. The variable 'grade' is coded as 0 for 11th grade and 1 for 12th grade; the variable 'general track behavioral sciences' is coded as 1 if the student follows a behavioral sciences track, and as 0 if this is not the case. The same goes for the coding of the variable 'general track non-behavioral sciences'. The variable 'gender' is coded as 0 for males and as 1 for females. Students' autonomous and controlled motivation were included as two continuous variables. The sum scores were divided by the number of items, resulting in two means (ranging from 1 to 5) for the two subscales (autonomous and controlled motivation), with a higher score indicating stronger motivation (De Naeghel, Van Keer, Vansteenkiste, & Rosseel, 2012). The scores on the LRST were included as a continuous variable. First, the means on the subscales were calculated by dividing the total score on a subscale by the number of items in a subscale. Second, the total test scores were calculated by adding the subscale scores and dividing them by the number of subscales. As such, the means vary between 0 and 1, and can be interpreted as percentages.

Results

Descriptive characteristics with respect to research skills proficiency and motivation

An overview of the mean scores for the LRST and the autonomous and controlled motivation scores for both cohorts is presented in Table 3. The mean on the total test is .566 (or 56,6%) for upper secondary education, and .686 (or 68,6%) in first year of university. An independent samples t-test reveals a significant difference between LRST scores of students in upper secondary education, and students in first year of university ($t(543) = -7.881, p = < .05$). No significant differences were found in autonomous ($t(542) = -1.671, p = > .05$) and controlled ($t(542) = .082, p = > .05$) motivation scores between the two cohorts.

Table 3

Mean scores and standard deviations for the LRST-scores and autonomous and controlled motivation

	Upper secondary education		First year of university	
	Mean	SD	Mean	SD
1.Overall LRST score	.566	.165	.686	.144
2.Problem identification	.623	.276	.701	.253
3.Questioning	.650	.368	.711	.375
4.Construction of artefacts	.625	.373	.768	.305
5.Hypotheses generation	.436	.357	.727	.337
6.Evidence generation	.814	.205	.927	.122
7.Evidence evaluation	.313	.346	.461	.328
8.Drawing conclusions	.520	.288	.558	.255
9.Communicating and scrutinizing	.543	.292	.634	.248
10.Autonomous motivation	3.355	.733	3.466	.583
11.Controlled motivation	2.286	.727	2.280	.678

Upper secondary education

Model parameters. A multiple linear regression analysis was performed to predict participants' LRST scores based on their grade, educational track, gender, autonomous motivation and controlled motivation. A significant regression equation was found ($F(6, 385) = 5,327$, $p < .001$) with an R^2 of .077.

As can be seen in Table 4, results from the multiple linear regression analysis indicate that grade, and both variables related to track had significant positive regression weights. As such, students' score on the LRST increases by 5,5% as grade increases by one unit (from 11th grade to 12th grade). In addition, both general tracks (behavioral sciences and non-behavioral sciences) contribute positively to the LRST-score, compared to the reference category (technical secondary education and education in the arts). As such, students in general secondary education, behavioral sciences track (compared to students in technical tracks or in secondary education in the arts) have higher scores (9%) on the LRST. The same holds for students in general secondary education, non-behavioral sciences tracks (11%). An additional independent samples t -test revealed no significant differences between LRST scores of students in general education, behavioral sciences track ($M = .57$), and students in general education, non-behavioral sciences track ($M = .60$) ($t(340) = -1,820$, $p = > .05$). Finally, nor gender, nor autonomous or controlled motivation seem to be significant predictors of students' research skills proficiency ($p > .05$).

Table 4

Linear model of predictors of research skills proficiency with beta values, standard errors, standardized beta values and significance values

Variable	B	$B (SE)$	β	p
Constant	.459	.056		.000
Autonomous motivation	.003	.011	.015	.772
Controlled motivation	-.007	.011	-.029	.564
Grade	.055	.016	.164	.001
Track: non behavioral sciences	.114	.027	.322	.000
Track: behavioral sciences	.085	.025	.258	.001

First year of university

Model parameters. A multiple linear regression was calculated to predict participants' LRST scores based on their educational track, gender, autonomous motivation and controlled motivation. As the respondents are situated in the same year, the independent variable grade was not included in the model for higher education. A significant regression equation was found ($F(5, 146) = 4,778, p < .001$) with an R^2 of .141.

As indicated in Table 5, results from the multiple linear regression analysis show that both variables related to track had significant positive regression weights. In line with the results in upper secondary education, both general tracks (behavioral sciences and non-behavioral sciences) contribute positively to the LRST-score, compared to the reference category (technical secondary education and education in the arts). As such, students in general secondary education, behavioral sciences track (compared to students in technical tracks or in secondary education in the arts) have higher scores on the LRST (9%). The same holds for students in general secondary education, non-behavioral sciences tracks (15%). An additional independent samples t -test revealed significant differences between LRST scores of students in general education, behavioral sciences track ($M = .67$), and students in general education, non-behavioral sciences track ($M = .73$) ($t(130) = -2,793, p < .05$). However, the differences in LRST scores between students in general education, non-behavioral sciences track and students in general education, behavioral sciences track are smaller (.065) than the difference between students in general education, non-behavioral sciences track and students in technical secondary education (.144). In addition, students in general education, behavioral sciences track outperform students in technical secondary education (.078). Finally, in accordance with the results in upper secondary education, nor gender, nor autonomous or controlled motivation seem to be significant predictors of students' research skills proficiency ($p > .05$). An independent samples t -test revealed significant differences between the LRST-scores of students in the general ($M = .70$) and the vocational ($M = .59$) tracks ($t(150) = -3,287, p < .05$).

Table 5

Linear model of predictors of research skills proficiency with beta values, standard errors, standardized beta values and significance

Variable	<i>B</i>	<i>B</i> (SE)	β	<i>P</i>
Constant	.481	.092		.000
Autonomous motivation	.029	.020	.118	.136
Controlled motivation	.018	.017	.083	.293
Track: non behavioral sciences	.152	.036	.522	.000
Track: behavioral sciences	.087	.036	.302	.016
Gender	-.047	.033	-.113	.156

Discussion and conclusion

The acquisition of research skills is an important goal of upper secondary education (Departement Onderwijs en Vorming, 2017; Engelmann et al., 2016; Fischer et al., 2014). Still, research on what learner characteristics affect research skills proficiency is scarce. Therefore, results from a multiple regression analysis on a test administration with 405 students in upper secondary education were explored. Next, the study was replicated in a second cohort (with 156 students enrolled in a first year course in university) in order to examine the generalizability of the results. Personal, academic and affective learner characteristics were included as independent

variables, and a cognitive learner characteristic was included as a dependent variable (Drachsler & Kirschner, 2012) in the model. The independent variables included in the model were gender, track, grade (only for upper secondary education), autonomous motivation and controlled motivation. The dependent variable was students' research skills proficiency (operationalized as the score on the LRST). The LRST (Leuven Research Skills Test, Maddens et al., 2019) and the Academic Self-Regulation Scale (Vansteenkiste et al., 2009) served as reliable measures for research skills proficiency and students' motivation towards research skills (education).

The overall mean on the LRST is 57% for students in upper secondary education, and 69% for students in higher education. This difference in LRST-scores between the two cohorts is significant. Results from the multiple regression analysis confirm that, in secondary education, educational grade and educational track are significantly related to students' research skills proficiency. However, no significant effects of autonomous motivation, controlled motivation or gender were found on students' LRST scores. In the second cohort, in higher education, educational track is found to be significantly related to students' research skills proficiency. Moreover, in accordance with the results in upper secondary education, no significant effects were found regarding autonomous and controlled motivation and gender in relation to LRST scores.

A first considerable finding of this study is that, nor in upper secondary education, nor in higher education, controlled and autonomous motivation were found to be (respectively negative and positive) significant predictors of students' research skills proficiency. However, based on previous research (Soenens & Vansteenkiste, 2005) one would hypothesize that controlled motivation (external and introjected regulation) is predictive for a broad variety of undesirable outcomes, including lower achievement. Autonomous motivation (intrinsic motivation and identified regulation), in contrast, is expected to be associated with positive outcomes, such as higher achievement. A possible explanation for the result of the current study on controlled motivation is that only a small percentage of students seemed to be highly or moderate highly motivated in a controlled way. Therefore, a lot of controlled motivation means are centered below neutral scores, and consequently a low variation exists in controlled motivation scores. As such, the results concerning controlled motivation must be interpreted cautiously. A possible explanation for the non-significant relationship of autonomous motivation with LRST-scores could be that students have difficulties to estimate the extent to which they are motivated towards research because they do not have a good understanding concerning research education at the time of the assessment. Since the high variability in schools (and thus curricula) in upper secondary education, there is no evidence that explicit instruction on research skills is covered (and thus it cannot be guaranteed that all these students know what research education is). As such, these results on motivation in upper secondary education can be flawed. However, the respondents in higher education are subscribed to a methodological course called 'methods and techniques for behavioral sciences research' and a statistics course, in which they learn about research (methods) in an explicit way, and thus, one would expect that these students can interpret the meaning of research skills education in an accurate way. As such, an alternative possible explanation could be that students' motivation towards research skills does not impact students LRST-scores in a significant way. These findings call for further investigation.

As a second important finding of this study, the variable grade was found to be a significant predictor for research skills proficiency in upper secondary education. As such, students in 12th grade of upper secondary education outperformed students in 11th grade of secondary education. However, as indicated in the Theoretical framework section, the assumption here is that the groups receive the same amount of instruction in research skills in both grades (which is not yet thoroughly investigated). In addition, it should be noted that an aging effect can interfere with the effect of

grade. As such, these findings call for further research on the characteristics of the implemented curriculum regarding research skills education.

A third important finding of this study is that, nor in upper secondary education, nor in higher education, gender was found to be a significant predictor of students' research skills proficiency. In addition, a significant number of previously published studies on the relationship of gender with research skills proficiency, science performance or scientific reasoning skills report inconsistent findings (Eylon & Linn, 1988; Linn & Hyde, 1989; Raes, 2015). Therefore, a reasonable consequence of this observation of a tradition of inconsistent findings regarding gender differences would be to, at least for the specific target group of upper secondary school students in a behavioral sciences context, question the inclusion of gender as a possible exploratory factor in future studies. However, when applying this finding to another sample, one should evaluate this statement with caution, since demographic characteristics of the Flemish school population (such as the predominance of girls in the behavioral sciences track) may not be generalizable to other groups.

Fourth, the impact of track on research skills proficiency has proved to be significant in both upper secondary education and in university. More particularly, students in general secondary education, or coming from general secondary education (behavioral sciences and non-behavioral sciences) outperform students in technical secondary education and secondary education in the arts. One possible explanation for this mean difference between the general and the non-general tracks is that students in technical education possess insufficient background information and domain specific knowledge (e.g., Kruglanski & Gigerenzer, 2011). As Eylon and Linn (1988) indicate, this could be either due to differences in personal cognitive abilities, or this could reflect what education makes or fails to make available to specific groups of students. The latter explanation might indicate the importance of curricular and instructional characteristics in relation to research skills proficiency. Therefore, an important area for future research is to focus on curricular characteristics or instructional guidelines contributing to research skills proficiency. Important questions to be answered by future research are thus what to instruct, and how to design this instruction. This promising research area can be realized by means of interdisciplinary research such as (curriculum) document analyses and intervention research on the role of curricular characteristics and instructional design principles (Fischer et al., 2014). Only when we have insights in the quality of instruction and in important instructional design principles succeeding in fostering research skills proficiency, we can further assess the possible role of motivation in research skills acquisition processes.

Apart from some important issues addressed above (such as a low variation in controlled motivation scores and the black box regarding educational practices), some other important limitations of this study must be addressed. First, the results from this study can only partly be compared to results from other studies or assessments mentioned (such as OECD, 2006), because the conceptualization of the term research skills differs from the conceptualization of the term scientific literacy. Still, it is argued that the two terms overlap (such as the focus on the interpretation of findings, the identification of issues, drawing conclusions...). Therefore, these studies were interpreted as important indicators for expected findings regarding research skills proficiency. In accordance with the previous comment, a second problem is that these studies differ in the way they operationalize motivation and analyze motivation results. For example, the studies on motivation differ in the way they clustered motivation results. In this study, and in line with De Naeghel and colleagues (2012), two continuous variables were computed for the two corresponding subscales (autonomous and controlled motivation), with a higher score indicating stronger motivation. Third, as mentioned earlier, results depend highly on contextual factors such as national

specific curricula. As such, an important limitation of this study is that its results rely partly on the characteristics of the Flemish curriculum (such as early tracking). Still, by offering a transparent description of this educational landscape, the important findings (such as the role of tracks and grade research skills proficiency) of this study can be transferred to or replicated in similar educational contexts. As a fourth limitation of this study, it should be mentioned that the models only explain 8 (in upper secondary education) to 14 (in higher education) percent of the variance in LRST scores. As such, an important challenge for future research remains to examine what other characteristics are affecting students' research skills proficiency.

To conclude, this study contributes to future research on research skills (education) by stressing the (possible) importance of curricular characteristics and instructional design principles for fostering student's research skills proficiency as a primary goal rather than focusing on trivial learner characteristics such as gender differences. In addition, the role of motivation on students' LRST scores should be further investigated in the future. As such, the remaining challenges are (1) to define what to instruct, (2) to decide how to design this instruction, (3) to test the possible effects of such instruction, and (4) to refine these instructional design principles for research skills education.

Acknowledgments

This study was carried out within imec's Smart Education research programme, with support from the Flemish government.

Declaration of interest

The authors declare that they have no conflict of interest.

References

- Departement onderwijs en vorming. (2017). *Secundair onderwijs - Derde graad ASO - Humane wetenschappen - Specifieke eindtermen*. Retrieved June 6, 2018, from <http://eindtermen.vlaanderen.be/secundair-onderwijs/specifieke-eindtermen-aso/-/humane-wetenschappen/specifieke-eindtermen.htm>
- De Naeghel, J., Van Keer, H., Vansteenkiste, M., & Rosseel, Y. (2012). The relation between elementary students' recreational and academic reading motivation, reading frequency, engagement, and comprehension: A self-determination theory perspective. *Journal of Educational Psychology*, 104(4), 1006–1021. <https://doi.org/10.1037/a0027800>
- Drachsler, H., & Kirschner, P. A. (2012). Learner characteristics. In N. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 1743-1745). Springer, Boston. MA.
- EACEA. (2018). *Belgium - Flemish Community Overview*. Retrieved June 6, 2018, from https://eacea.ec.europa.eu/national-policies/eurydice/content/belgium-flemish-community_en
- Earley, M. A. (2014). A synthesis of the literature on research methods education. *Teaching in Higher Education*, 19(3), 242–253. <https://doi.org/10.1080/13562517.2013.860105>
- Eder, D. (1981). Ability grouping as a self-fulfilling prophecy: A micro-analysis of teacher-student interaction. *Sociology of education*, 151-162.
- Engelmann, K., Neuhaus, B. J., & Fischer, F. (2016). Fostering scientific reasoning in education – meta-analytic evidence from intervention studies. *Educational Research and Evaluation*, 22(5–6), 333–349. <https://doi.org/10.1080/13803611.2016.1240089>
- Eylon, B.-S., & Linn, M. C. (1988). Learning and Instruction: An Examination of Four Research

- Perspectives in Science Education. *Review of Educational Research*, 58(3), 251. <https://doi.org/10.2307/1170256>
- Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., ... Eberle, J. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, 4, 28–45. <https://doi.org/http://dx.doi.org/10.14786/flr.v2i2.96>
- Kestens, J., Elen, J., & Verburch, A. (2016). *Kritisch denken en onderzoekscompetenties: de nexus*. (Unpublished master's thesis). KU Leuven, Leuven.
- Kolodner, J. L. (2007). The Roles of Scripts in Promoting Collaborative Discourse in Learning by Design. *Scripting Computer-Supported Collaborative Learning*, 6, 237–262. <https://doi.org/10.1007/978-0-387-36949-5>
- Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, 118(1), 97–109.
- Linn, M. C., & Hyde, J. S. (1989). Gender, Mathematics, and Science. *Educational Researcher*, 18. Retrieved from <http://journals.sagepub.com/doi/pdf/10.3102/0013189X018008017>
- Linn, M., Davis, E., & Bell, P. (2004). *Internet Environments for Science Education BT - Internet environments for science education*. New Jersey: Lawrence Erlbaum Associates. Retrieved from <https://books.google.be/books?hl=nl&lr=&id=wERBomDOPqMC&oi=fnd&pg=PA29&dq=linn+eylon+davis+&ots=cYTwsOU4N&sig=1xrOtLmWuepqBLf9cG4dFuKyPts#v=onepage&q=linn+eylon+davis&f=false>
- Maddens, L., Depaepe, F., Janssen, R., Raes, A., & Elen, J. (2019). Evaluating the Leuven Research Skills Test for 11th and 12th Grade. *Journal of Psychoeducational Assessment*, 0734282918825040.
- National Research Council. (2010). *Exploring the Intersection of Science Education and 21st Century Skills: A Workshop Summary*. Margaret Hilton, Rapporteur, Board on Science Education, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academics Press.
- Norris, S. P., Phillips, L. M., & Burns, D. P. (2014). Conceptions of Scientific Literacy: Identifying and Evaluating Their Programmatic Elements. In M. Matthews (Ed.), *International Handbook of Research in History, Philosophy and Science Teaching* (pp. 1317–1344). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-7654-8_40
- OECD. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris: OECD.
- OECD. (2007). *PISA 2006: Science Competencies for Tomorrow's World. Volume 1, Analysis*. Paris: OECD.
- OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education, PISA*. OECD Publishing, Paris, <https://doi.org/10.1787/9789264266490-en>.
- OECD. (2018). *PISA 2015 Results in Focus*. Paris: OECD. Retrieved from <https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf>
- Opitz, A., Heene, M., & Fischer, F. (2017). Measuring scientific reasoning – a review of test instruments. *Educational Research and Evaluation*, 23(3–4), 78–101. <https://doi.org/10.1080/13803611.2017.1338586>
- Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections* (Vol. 13). London: The Nuffield Foundation.
- Prenzel, M. (ed.) (2007). *Studies on the educational quality of schools. The final report on the DFG Priority Programme*. Münster: Waxmann.

- Raes, A. (2015). *The impact of web-based collaborative inquiry for science learning in secondary education* [Doctoral Dissertation]. Universiteit Gent, Gent.
- Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57(5), 749.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25, 54–67. <https://doi.org/10.1006/ceps.1999.1020>
- Soenens, B., & Vansteenkiste, M. (2005). Antecedents and Outcomes of Self-Determination in 3 Life Domains: The Role of Parents' and Teachers' Autonomy Support. *Journal of Youth and Adolescence*, 34(6), 589–604. <https://doi.org/10.1007/s10964-005-8948-y>
- Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic Versus Extrinsic Goal Contents in Self-Determination Theory: Another Look at the Quality of Academic Motivation. *Educational Psychologist*, 41(1), 19–31. https://doi.org/10.1207/s15326985ep4101_4
- Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology*, 101(3), 671–688. <https://doi.org/10.1037/a0015083>
- Zimmerman, C. (2000). The Development of Scientific Reasoning Skills. *Developmental Review*, 20(1), 99–149. <https://doi.org/10.1006/drev.1999.0497>

Accepted manuscript