Write Right -

A Practical Action Research into Teaching Information Literacy with the 4C/ID Model

Schoon Schrift -

Een Practical Action Research naar Onderwijs in Informatievaardigheden met het 4C/ID model

Chantal Mülders

Master Onderwijswetenschappen

Open Universiteit

Datum: 6 maart 2018

Studentnummer: 851838542

Begeleiding: Prof. dr. F. L. J. M. Brand - Gruwel

Examinator:

Table of contents

Summary   **Write Right**   **Chantal Mülders**

BACKGROUND - Prior research (e.g., Van Merriënboer, Jelsma & Paas, 1992; Melo & Miranda, 2014) discovered positive transfer effects for 4C/ID (4 Components Instructional Design) designs. Wopereis, Frerejean and Brand-Gruwel (2015) and Wopereis, Frerejean and Brand-Gruwel (2016) reported positive perceptions in teacher and students after a course designed with 4C/ID. This model is intended to plan instruction in complex skills, so it is suitable for Information Problem Solving (IPS). IPS or information literacy belongs to the 21st century skill set that students should acquire to be able to operate in a fast-changing, technological environment (Partnership for 21st Century Skills, 2009).

AIM - In order to analyze how information literacy could be taught using 4C/ID principles, the main question was as follows: What are the effects on teacher perceptions, student perceptions, and student performance in higher education in a course incorporating information literacy designed according to the 4C/ID model?

DESIGN - A Practical Action Research was executed with teacher perceptions, student perceptions and student performance as variables. Linguistic ability in English was considered a covariate. The control group consisted of 72 students, and the experimental group of 115 students, divided into a Dutch (48) and English (67) stream.

MEASURES - Data on the perceptions of teachers and students were gathered via field notes, observation, interviews and a questionnaire. Teachers were interviewed three times. Students completed the Teaching and Learning Quality (TALQ) questionnaire (Frick, Chadha, Watson and Zlatkovska, 2010). This measures Merril's (2002) First principles, time, progress, satisfaction, and quality.

RESULTS - Results indicated that the redesigned program did not influence the number and type of errors students made compared to the control group. Students who completed the program made as many mistakes as the control group and on the same aspects. Students who did not take part did have higher error rates. Both teachers and students perceived the course positively although student perception correlated with participation. English ability did not predict improved performance, but participation did.

CONCLUSION - The current research shows that teachers and students rate a 4C/ID design positively. Student performance is closely linked to participation, but it is unclear which aspect of the renewed approach causes that. This research does demonstrate that information literacy differs from language ability and underscores the need for explicit instruction in this area. Further experimental research is required, ideally at different levels and incorporating other aspects of the IPS skill.

Key words: 4C/ID model, information literacy, higher education, perceptions

Nederlandse samenvatting   **Schoon Schrift**   **Chantal Mülders**

ACHTERGOND - Onderzoek (o.a. Van Merriënboer, Jelsma & Paas, 1992; Melo & Miranda, 2014) vond positieve transfereffecten voor 4C/ID (4 Components Instructional Design) ontwerpen. Wopereis, Frerejean en Brand-Gruwel (2015) en Wopereis, Frerejean en Brand-Gruwel (2016) meldden positieve percepties in docenten en studenten na een cursus gebaseerd op 4C/ID. Dit model richt zich op instructie in complexe vaardigheden en is daarmee geschikt voor Information Problem Solving (IPS) of informatie-vaardigheden. Deze behoren tot de 21^e-eeuwse vaardigheden die studenten moeten verwerven om in een veranderende, technologische omgeving te kunnen werken (Partnership for 21st Century Skills, 2009).

DOEL - Om te analyseren hoe informatievaardigheden volgens 4C/ID principes kunnen worden getraind, is de hoofdvraag: Wat zijn de effecten op docentpercepties, studentpercepties en studentprestaties in het hoger onderwijs in een cursus waarin informatievaardigheden zijn opgenomen, ontworpen volgens het 4C/ID model?

DESIGN - Een Practical Action Research is uitgevoerd met docentpercepties, studentpercepties en studentprestaties als variabelen. Taalvaardigheid in Engels is meegenomen als covariaat. De controlegroep bestond uit 72 studenten. De experimentele groep bestond uit 115 studenten verdeeld over een Nederlandse (48) en Engelse (67) stroom.

MEETINSTRUMENTEN - Data over de percepties van docenten en studenten werd verzameld via veldnotities, observatie, interviews en een vragenlijst. Docenten werden drie keer geïnterviewd. Studenten vulden de Teaching and Learning Quality (TALQ) vragenlijst in (Frick, Chadha, Watson and Zlatkovska, 2010). Deze meet First Principles van Merril (2002), tijd, voortgang, tevredenheid en kwaliteit.

RESULTATEN - Resultaten tonen dat het herontwerp geen invloed had op de hoeveelheid en het type fouten dat de studenten maakten in vergelijking met de controlegroep. Studenten die deelnamen maakten evenveel fouten als de controlegroep en op dezelfde aspecten. Studenten die niet deelnamen maakten wel meer fouten. Zowel docenten- als studentenpercepties waren positief, al waren studenten positiever als ze meer deel hadden genomen. Niveau Engels was geen voorspeller voor prestaties, maar deelname wel.

CONCLUSIE - Het huidige onderzoek toont aan dat docenten en studenten positief zijn over een 4C/ID ontwerp. Studentprestaties hangen nauw samen met deelname, maar het is onduidelijk welk aspect van de vernieuwde aanpak hieraan ten grondslag ligt. Informatievaardigheden verschillen wel van taalvaardigheid en dit onderzoek benadrukt de noodzaak van expliciete instructie. Verder experimenteel onderzoek is nodig, idealiter op verschillende niveaus en over andere aspecten van de IPS vaardigheid.

Keywords: 4C/ID model, informatievaardigheden, hoger onderwijs, percepties

Write Right -

A Practical Action Research into Teaching Information Literacy with the 4C/ID Model

Chapter 1: Introduction

 "Over the last few decades biologists have reached the firm conclusion that … man … is also an algorithm" (Harari, 2017, p. 98). An algorithm is a mathematical formula, a set of steps, a flowchart, that allows for processes to continue along certain paths as the intelligent actor determines based on the external circumstances. Harari (2017) argued that humans themselves are algorithms too: "You can use numbers and mathematical symbols to write … the series of steps a brain takes when it is alarmed by the approach of a lion." (p. 131). Harari did not question how the brain knows to follow those exact steps, how the person *learned* to recognize the lion and analyze the surroundings. If humans are algorithms constantly working to decide the next move, how can we teach others what steps the algorithm should consist of and what external circumstances to consider?

The 4C/ID Model

The Four Component Instructional Design (4C/ID) model assumes that complex learning requires the formation of something akin to a complicated algorithm in students' minds. A basic algorithm suffices for simple tasks, but for sophisticated problems, the student needs to slowly build to full complexity, allowing more branches to be attached progressively. The addition of such knowledge branches requires the Goldilocks principle, getting it just right: not too little to bore the learner, but not too much as to overwhelm him.

   Cognitive load and schema construction. The 4C/ID model recognizes these considerations as it stems from Cognitive Load Theory. Sweller, van Merriënboer and Paas (1998) posited that working memory capacity is limited to two or three items, so they deduced that instruction should be designed to minimize the burden. They suggested that long-term memory has no limits and is organized in patterns, or schemata, so they define learning as the building, expanding, and automating of such schemata. Because even a highly complex schema is a single entity, it occupies a single slot in working memory (Van Merriënboer & Sweller, 2005). Sweller et al. (1998) explained that the required load to learn a new piece of information can be a result either of the material's complexity (intrinsic) or how it is presented and how students are to process it (extrinsic). Van Merriënboer and Sweller (2005) stated that cognitive overload can occur if all cognitive loads together are too high or if modalities are mixed. In other words, a poor design hampers progress with too high a load, but a superior design contributes to the student's schema construction.

Designing with 4C/ID. A 4C/ID design can be achieved after careful preparation and employing specific practice types because fragmentation will occur when information is presented in small pieces (Dolmans, Wolfhagen & Van Merriënboer, 2013). Van Merriënboer, Jelsma and Paas (1992) described that in the analysis phase, the designer identifies the recurrent and non-recurrent component skills and prior knowledge. In the design phase the tasks and presentation of information are formed, including variability of practice and demonstration. Sweller et al. (1998) suggested using goal-free problems, worked examples, and completion problems. Van Merriënboer, Kirschner and Kester (2003) further elaborated on Sweller et al.'s ideas with the right amount of scaffolding, authentic tasks at appropriate complexity, employing simplifying conditions, and the completion strategy. Van Merriënboer and Kirschner's (2013) book presented the comprehensive 4C/ID model to design instruction for complex learning. This model focuses on a whole-task approach with the designer creating task classes of increasing complexity. Each task class offers appropriate supporting information available to students at all times and just-in-time information to explain specific procedures. The learning tasks are organized from high to low scaffolding, and part-task practice can automate specific skills.

Effectiveness of the 4C/ID model. According to Van Merriënboer et al. (1992), the true test for the efficacy of any 4C/ID design is whether transfer of the practiced skills to new situations occurs. In their three experiments the experimental group outperformed the control group on the transfer test. Similar positive, experimental results were found, for example, in 14-year-olds building electric circuits (Melo & Miranda, 2014), an integrated information literacy task at a teacher's college (Brand-Gruwel, Wopereis & Poortman-Cremers, 2005), increasing technical expertise in Ghanaian secondary schools (Sarfo & Elen, 2007), and prospective teachers constructing a grade book in Excel (Lim & Reiser, 2006).

Apart from these experimental studies focusing on transfer, other later studies examined different success factors. Holtslander, Racine, Furniss, Burles and Turner (2012) and Wopereis, Frerejean and Brand-Gruwel (2015) found that their 4C/ID designs achieved high student ratings in the evaluation. In a follow-up study, Wopereis, Frerejean and Brand-Gruwel (2016) found that the course's instructors perceived the whole-task approach, structure, process worksheets, and feedback as strengths. Kicken, Brand-Gruwel, Van Merriënboer and Slot (2009) discovered that the success of a 4C/ID design depends on the acceptability of such an approach to both students and teachers.

Information Literacy

The 4C/ID model as described above is a useful tool for designing education for complex skills, so it is applicable to 21st century skills, more specifically information literacy. The term 21st century skills encompasses core subjects and themes with the necessary life, career, learning, innovation, information,

media, and technology skills (Partnership for 21st Century Skills, 2009). Although the 21st century skill set overall does not diverge much from the ones required in earlier decades, information literacy is strikingly different from the past.

The need for information literacy. Information literacy requires more consideration than before because digitalization and the internet have had a major impact. The quantity of available information has increased tremendously, as well as the speed at which it can be updated and found. Redecker et al. (2011) described that, as result of this, "the shelf-life of knowledge is decreasing; the amount of information is increasing" (p. 26). The Partnership for 21st Century Skills (2009) and Association of College and Research Libraries (ACRL, 2015) stressed this as well: Anyone can access more information than ever before, using new technological tools, and work together with others on a large scale. Current students need to balance the amount and flow of information, so they can apply this behavior to new problems. As Wilson (1999) defined it, information behavior is the actions taken to determine what information is needed, to find it, and use it. This behavior can solve an information problem, when new information must be found to answer a question (Brand-Gruwel, Kammerer, Van Meeuwen & Van Gog, 2017). Students, however, tend to find it exceedingly difficult to tackle such problems (e.g., Brand-Gruwel, Wopereis and Walraven, 2009; Brand-Gruwel et al., 2017).

Models for information literacy. Because of the complexity of information literacy, various models exist that explain the concept, and three of those models will be discussed here. First, the general model by the Partnership for 21st Century Skills (2009) included information literacy as one of its constituent parts. They separated it into two subsets: first is access and evaluation of information, and second is the use and management of information. The first subset includes (a) the ability to efficiently and effectively access information and to then (b) evaluate this information critically. The second subset covers (a) the accurate and creative use of information for the problem, (b) managing the information flow from many different sources, and (c) understanding the ethical and legal issues in accessing and using information. The second more specific model for only information literacy by ACRL (2015) defined six frames. These are brought together by "reflective discovery" (p. 3), understanding the value and production of information, knowing how information is used to create new knowledge, and ethical behavior. The third model of Information Problem Solving (IPS) by Brand-Gruwel et al.(2009) placed information literacy in a sequence of five steps. Students define the problem, search for the information, scan, process, and then organize and present the information. In each step they actively regulate the process.

Comparing the three models. These three models mostly overlap in their treatment of information literacy, but a few discrepancies can be found. The ACRL (2015) description perfectly aligns with the

Partnership for 21ˢᵗ Century Skills' (2009) model: the first two with the first subset and the last two with the second subset. Although the three intermediate steps of IPS (Brand-Gruwel et al., 2009) clearly intersect with the other two models, the first and last step are less obviously related. Problem definition is explicated by ACRL as "the practice to formulate questions for research based on information gaps or on reexamination of existing, possibly conflicting, information" (p. 7). It could be included in the Partnership's second subset, part (a). The last step of actual information presentation, the ACRL describes as synthesizing ideas, drawing conclusions, and citing others. The Partnership for 21ˢᵗ Century Skills did not categorize this as information literacy, but communication and collaboration: "Articulate thoughts and ideas effectively using oral, written and nonverbal communication skills in a variety of forms and contexts" (p. 4). However, in an academic context, information literacy and communication interact as convention demands a clear representation of what source was used where, as ACRL aptly recognizes. Therefore, the IPS model is the more useful one in illustrating the sequence of steps while ACRL offers in-depth insight into a student's skill set at each point.

Because information literacy is increasing in importance, but students struggle with its complexity, it is necessary to carefully design education on this issue. Therefore, the 4C/ID model, as described, is a coherent, structured, and evidence-based approach for designing a course that includes this skill.

This Research: Teaching Information Literacy with 4C/ID

In order to achieve practical insight into *how* information literacy could be taught in higher education, this research examines the English semester 4 course on writing that was redesigned with the 4C/ID model. The steps that the students take align with the IPS model by Brand-Gruwel et al. (2009). The added value of this study lies in the actual application of the 4C/ID model as a tool in higher education in general and to teach information literacy in particular. The 4C/ID model assumes that the reduction of fragmentation and compartmentalization will increase the transfer of the desired skill by creating appropriate cognitive schemata (Van Merriënboer & Kirschner, 2013). If this assumption holds true, students should produce fewer or different errors on a transfer test of information literacy. Moreover, the experiences of teachers and students during the course can highlight strengths and weaknesses in the design.

The main question this thesis aims to answer is as follows: What are the effects on student performance, student perceptions, and teacher perceptions in higher education in a course incorporating information literacy designed according to the 4C/ID model? The first variable is the performance of students in the new approach compared to the old method (mainly the amount and type of errors), the second variable covers the perceptions of the students who are taking this new course, and the third variable investigates the perceptions of teachers on the course. For the student performance variable,

students in the new course can be compared to those who have taken the old course and resitters who did not take the new course.

These three variables are expected to influence each other differently as illustrated in Figure 1.
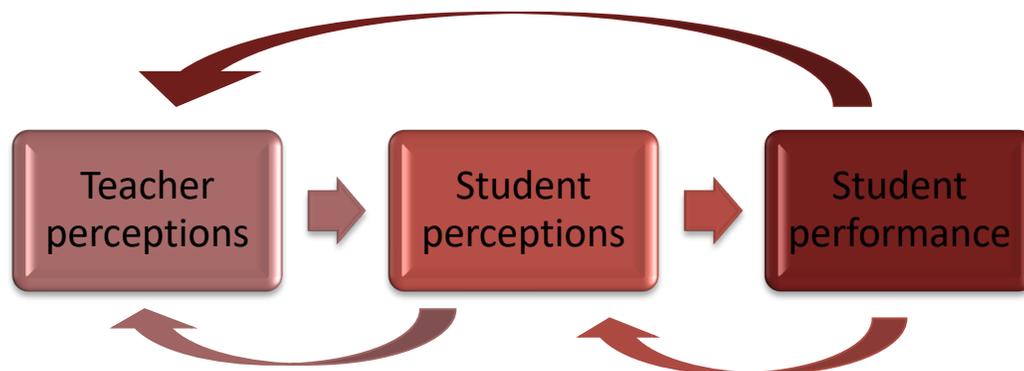
Figure 1. Relationship between the three variables in this research

The main direction of influence is expected to be that teacher and student perceptions affect student performance. Jossberger, Brand-Gruwel, van de Wiel and Boshuizen (2015) found that "[t]eacher beliefs and values influence the implementation of change" (p. 290). Wopereis et al. (2016) investigated their course on IPS modeled on 4C/ID and reported that teachers perceived it to be of high quality. They found few differences between teacher perceptions and student perceptions of the course. Kicken et al. (2009) reported on self-directed learning via the 4C/ID model and found that acceptability of a design to both teachers and students is crucial to its success. Therefore, positive teacher and student perceptions will probably result in improved student performance while negative perceptions will likely adversely affect performance.

However, student and teacher perceptions are unlikely to affect student performance independently. Kicken et al. (2009) stated that the frequency of use in their case study was mainly due to student characteristics. Jossberger et al. (2015) described how "[t]he interplay between students and teachers forms a reciprocal loop that can influence the overall atmosphere either positively or negatively" (p. 309). Although both teachers and students need to be willing to work with the new design, the student needs to complete the tasks. Therefore, teacher perception can influence student perception, which in turn affects student performance. However, as Jossberger et al. highlighted, perceptions of students and teachers interact, so a loop needs to exist between these two groups: Student perceptions may alter the way a teacher views the new design. They might perceive the 4C/ID course as having high quality (similar to

Wopereis et al., 2016), but view it as ineffective if student performance is low. Finally, the way a student performs will probably also influence the perceptions of both students and teachers on the course's effectiveness. Two studies, one by Braga, Paccagnella and Pellizzari (2014) and another by Carrell and West (2010), found that students rated their teachers lower in courses that were more effective in the long term. These students achieved lower scores on the initial course, so they rated that course lower as well, despite this course teaching them more in the long run.

In this research the following three hypotheses will be tested:

1. H_0: Student performance in the new program does not differ from the previous program, also when accounting for English ability, participation in the design, or the background factors gender, nationality, ethnicity, and previous diploma.

   H_1: Student performance differs based on gender, nationality, ethnicity, or previous diploma.

   H_2: Students with higher participation in the design perform better.

   H_3: Students with higher English ability perform better.

   H_4: Student performance in the new program is better than previous students.

2. H_0: Student perceptions on the new course are neutral, irrespective of performance.

   H_1: Student perceptions on the new course are positive.

   H_2: Student perceptions on the new course are more positive if they performed better.

3. H_0: Teacher perceptions on the new course are neutral.

   H_1: Teacher perceptions on the new course are positive.

Methods

Design

In order to draw conclusion about cause and effect for a 4C/ID design, it is necessary to adopt a true or quasi-experimental design. This is the most rigorous of approaches, but executing that design for this research carries ethical repercussions. Pass rates from previous cohorts hovered at 50%, so it would be unethical to exclude certain students from an adjusted approach that potentially benefits them in this course, their further studies, and future employment.

In addition, a purely quantitative design excludes potentially valid and informative qualitative data. The opinions and ideas of teachers and students impact the execution of any curriculum. Furthermore, classrooms are complex environments in which the instructional design is one of a multitude of factors, such as teacher interaction and student engagement to mention but a few. This redesign is a multi-faceted adjustment that includes many more factors than can be statistically controled. Apart from adjusted course material, the design incorporates more individual feedback to students, more writing practice, and a

different role for the teacher. Therefore, mixing research methods allows for greater detail in studying the redesigned course and can include some of those factors that cannot be captured quantitatively.

Because the goal is to study the implementation of a new course in detail, a practical action research design was chosen. This allows for a mixed methods approach, and the researcher can systematically identify problem areas in the redesigned course and improve it further, to enhance student learning in this specific situation. The researcher can further benefit from the views of the co-teachers, creating a team inquiry and reflective spirit. Data collection can include all three aspects of action research design (Mills, 2011 in Creswell, 2014): experiencing via observations and field notes, enquiring via interviews and questionnaires, and examining by studying the students' work.

Participants

Teachers as well as students participated in this research. After discussing the three teachers, information about the students is presented. All teachers and students are part of the International Business and Languages (IBL) program at Rotterdam University of Applied Sciences.

Table 1

Background Information on the Course's Teachers

Teacher	Gender	Nationality	Native language	At IBL since	Years of teaching experience		
					General	Writing	In higher education
J	F	Canadian	English	Sep. 2016	8	8	1[a]
N	M	American	English	Sep. 2016	6	6	1[a]
Researcher	F	Dutch	Dutch	Feb. 2012	10	9	10

[a]Teachers J and N had several years of experience in the final years of secondary education, an age group that partially overlaps with the IBL students. Both had taught writing with sources at those levels.

Teachers. Relevant background information about the three teacher participants of this research is displayed in Table 1. Information about teachers J and N was gathered from the formal mid-way interview. Management allocated the teachers to the classes according to availability and time.

Student control group. The control group for this research consisted of two subsets. Subset A included 37 students from previous years whose reports could be found in the archives. Those reports constituted first attempts, but not necessarily a pass. Subset B consisted of 35 students who took part in a resit in April 2017. All these students had taken the course before but had not passed the report writing

part. They created their own groups of three or four to complete the report, so they had possibly taken the original course in different years. Although the course was modified slightly every year, no significant changes were made until the redesign in this academic year. No students of either subset retook the course. An Independent Samples T-Test confirmed that the two subsets did not differ significantly on the errors they made in the reports, their English score at high school, their English scores at IBL, or the number of resits for English. Therefore, the two subsets were collapsed into a single control group.

The control group consisted of 43 females and 29 males, and all but five students had the Dutch nationality. Forty-six students followed the Dutch stream and 26 the English stream. Most students had white ethnicity (37), followed by non-western (31) and white non-Dutch (4). Of all control group participants, 48 had a Dutch high school diploma, 14 had a vocational college degree, five had completed a 21+ entrance test, and five had a foreign diploma. These background factors were fairly evenly divided across the two subsets.

Table 2

Background Information on Experimental Group

Class	Stream	N	Gender (male : female	Dutch nationality	Non-white ethnicity[a]	First year completed[b]	Teacher
201N	Dutch	19	5 : 14	19	3	16	Researcher
202N	Dutch	14	4 : 10	14	6	9	Researcher
203N	Dutch	15	3 : 12	15	7	8	Researcher
211E	English	19	5 : 14	17	12	5	J
213E	English	21	4 : 17	21	4	14	J
214E	English	27	13 : 14	14	3	26	N

[a]Students with at least one non-white parent. [b]Number of students in this class who had completed all first-year courses as of May 1, 2017.

Student experimental group. The experimental group consisted of 115 students divided in six classes (see Table 2). Students had been randomly placed into classes at the beginning of the academic year by the institute's administration. The main distinction into Dutch and English stream reflects the student's personal decision made at the start of the program. Access to Honors class 214E was based on other selection criteria as well, such as academic performance and motivation.

Materials

Intervention. The intervention consisted of six booklets designed to follow 4C/ID principles in teaching information literacy and writing skills. This redesign was planned as part of the Open University (Heerlen, the Netherlands) Master's course in Educational Design with the practical conversion completed afterwards. See Appendix A for a shortened blueprint of the redesigned course.

Students received five booklets, copied double-sided, one staple at the top, and four holes punched on the side. These were formative practice booklets, offering students relevant exercises organized and designed according to 4C/ID principles (also see the blueprint in Appendix A). Those booklets contained two real-life cases; the first case consisting of exercises (up to eleven) and the second case being the conventional writing task of one or two paragraphs. The exercises were designed to aid students with specific issues in their development as writers employing secondary sources, such as writing topic sentences (booklet 2), evaluating sources for trustworthiness (booklet 3), and executing a Google search (booklet 5). The final conventional task was discussed during individual feedback sessions with the teacher. Students could write their answers in the booklets, except when using materials from Bailey (2015). Students received the first booklet in the first class, and consecutive booklets were distributed when that student had completed the final conventional task of the previous booklet satisfactorily, according to the assessment scheme. Each booklet included such an assessment scheme for the teacher to evaluate the task. This ensured transparency of quality standards, mainly to students, but also to teachers.

The sixth booklet contained all the information for the summative test of writing a research report. The format was identical to previous years: it could not be altered as the course description, written in May 2016, had already established this testing method. Although the sixth booklet therefore did not follow the 4C/ID model as closely as the formative booklets, having an identical summative test method allowed for comparison with the control group. Booklet 6 was also available digitally.

Students completed two preparatory assignments and then their research report, which involved both group and individual work. They could choose from three cases, resembling their projects of year 1 and 2, but could also add their own topic. One case was identical to previous years, and the other two presented a similar setup as before. Each individual answered two subquestions, maximum one page each and using at least six sources. It is those individual answers that were analyzed for this research

Next to the six hard copy booklets, students had access to additional materials digitally. This contained the PowerPoints of the classes, answer keys for the booklets (published after the scheduled feedback sessions), a template for the report, and the possibility to upload their work.

Instruments. Performance and perceptions were measured using interviews, field notes, student reports, and a questionnaire. This section covers what measure was used for which variable.

The first variable, student performance, was measured quantitatively via error analysis of the student's final reports. Three categories were used to group the errors. The first category was labeled APA Application, which involved the mechanics of applying APA style in a text, e.g., correct usage of brackets, using a page or paragraph number for a quotation, or employing a lower case letter after the year for publications with the same author-year combination. The second category was termed Source Integration, which included the correct placement of a citation, e.g., providing clarity about the origin of each piece of information and foregoing repetitive use of the (author, year) format. Source Integration requires more in-depth analysis of the writing than APA Application: the latter is merely applying the technical rules of APA style whereas the former obliges a student to critically examine whether to place a citation and where. The style mechanics consist of constant rules that can be researched straightforwardly, but the rules of placement have more flexibility and depend on the actual text the student wrote. The last category, Trust, entailed issues of trustworthiness, in particular avoiding plagiarism and employing reliable sources. Appendix B displays a detailed overview of the errors, including one or two examples and their corrections. The number of errors for the categories APA Application and Source Integration can be high as each repeat mistake was counted. Trust will logically include fewer errors as most instances of plagiarism are unintentional, and the number of sources used tends to be the required minimum of three per subquestion.

The second variable, student perceptions, was measured both qualitatively and quantitatively. Qualitative data collection consisted of interviews, field notes, and observation. During class or feedback sessions students were interviewed informally, which was recorded in the field notes. Students were asked leading questions, such as "What did you like about booklet X? What not?" and "Which parts were easy and which were difficult?". The researcher also served as a privileged, active observer for the three Dutch stream classes. Quantitative and further qualitative data regarding student perceptions was gathered via a paper survey (see Appendix C) in the last class of the course. The first section surveyed student participation in the practice booklets, time effort for the course, and reasoning for certain actions, such as not attending a feedback session. The second part was based on Frick, Chadha, Watson and Zlatkovska's (2010) Teaching and Learning Quality (TALQ) questionnaire. They had designed the TALQ to align with Merrill's (2002) First Principles, whose focus on authenticity of problems and tasks is fundamental to 4C/ID designs (Frick et al., 2010). The survey was validated across subjects and disciplines, including business. Five scales were used to measure Merril's First Principles and four more scales for time, progress, satisfaction and quality. The Cronbach α's for these scales were Authentic Problems 0.690, Activation 0.812, Demonstration 0.830, Application 0.758, Integration 0.780, Total First Principles 0.881,

Academic Learning Time 0.763, Learning Progress 0.935, Satisfaction 0.926, and Global Quality 0.915. The survey included 35 statement on a 5-point Likert scale (*completely disagree - completely agree*). Each scale constituted the aggregate of three to five statements in the questionnaire; reverse-coded questions were converted to regular coding prior to analysis. Statements were not presented per scale, but jumbled. The original language level was deemed too high in some statements, so they were simplified slightly to a B2 level to ensure understanding. Furthermore, all statements were modified to include the phrase "in the written part of the course" to avoid students rating other components. The survey was not tested before distribution to the students.

The third variable, teacher perceptions, was measured via three interviews. The first interview was an unstructured group discussion during which the teachers voiced their opinions on the design and the practical execution. The two other interviews were one on one and semi-structured. This approach allowed for the collection of the required information without rigid procedures, in line with the collegial relationship that was already established between the teachers and the researcher. Questions centered on their opinion of the program, student feedback, and suggestions for improvement (see Appendix D for the interview questions). The one-on-one interviews were recorded digitally and transcribed while the group discussion was entered into the field notes.

Procedure

Data collection started in February 2017 and was completed in July 2017. All data was gathered within the IBL program of the Rotterdam University of Applied Sciences, the Kralingse Zoom location. Classes and feedback sessions took place in different rooms of this building as per the schedule with occasional additional feedback at their teacher's desk.

The course was taught following an ON/OFF schedule over fourteen teaching weeks as suggested by teacher J during the informal interview prior to the course on February 1st. During an ON week students attended class as a group; during an OFF week, students could schedule five-minute individual sessions with their teacher using Calendly.com. Mostly, an ON week would follow an OFF week, with exceptions after the mid-term exam week and two consecutive weeks of ten-minute final feedback sessions for the research report. Two classes had a switch in weeks 9 and 10 due to their required attendance at the in-house trade fair for their project. During the mid-term exam week in April, the first formal interview with the teachers was conducted. In the 14th week students submitted their research report; the last ON class was used for oral assessment and the questionnaire. Teachers graded the work, after which all reports were gathered by the researcher for error analysis. The final formal interview took place in July, after the course had finished.

The above describes the approach for experimental group; however, the control group did not take the renewed classes. These students had followed English semester 4 classes in previous years, which had been modified only slightly. Grading had been identical, but classes had met every week for three hours. The materials had been offered in a fragmented way with few possibilities for students to practice their writing and receive feedback.

Data Analysis

The gathered data was analyzed both qualitatively and quantitatively. The informal interviews with students and co-teachers were noted as the researcher's field notes while the one-on-one interviews with the co-teachers were recorded digitally and transcribed. Background information about the students was gathered from the institute's student information program Osiris. Students received the end-of-term questionnaire hard copy in their last class. That data was transferred to an Excel file for further analysis with open questions copied for text analysis. Error analysis noted per student how often each person made a specific error, which included only the two individual subquestions.

SPSS was used to conduct statistical analysis on the quantitative data, such as ANOVAs and correlations. Results were deemed significant at $p < .05$ because the sample size never exceeds 200.

Results

In this research, three experienced teachers taught six classes of second-year students a course that incorporated information literacy, redesigned following 4C/ID principles. The writing of 187 students was collected for error analysis, 115 in the experimental group and 72 in the control group. The experimental group was further divided into a Dutch and English stream of 48 and 67 students respectively. Independent Samples T-Tests confirm that the Dutch classes ($n = 36$, $M = 6.56$, $SD = .74$) differed significantly from the English classes ($n = 40$, $M = 7.33$, $SD = .69$) in their high school English scores ($t(72.07) = -4.68$, $p < .001$). This was also found for the Dutch ($n = 48$, $M = 6.81$, $SD = .40$) and English ($n = 67$, $M = 7.62$, $SD = .59$) classes for their scores on their three previous IBL English courses ($t(112.64) = -8.76$, $p < .001$) and the Dutch ($n = 48$, $M = 1.48$, $SD = 1.68$) and English ($n = 67$, $M = .64$, $SD = .16$) classes for number of English resits ($t(87.17) = 2.86$, $p = .005$). No difference was found between the groups on the academic performance measure of completing the first year, so their general academic ability is similar. Because a difference in English ability can have an impact on their performance in an English writing course, the experimental group was divided in two accordingly. The sample size of each group still suffices for statistical analysis ($N > 30$). The results are analyzed per hypothesis: effect of the program on student performance, student perception, and teacher perception.

Hypothesis 1: Student Performance

The first hypothesis of this research tests whether student performance within the experimental group differed and compares this to the performance of the control group. The background factors gender, nationality, ethnicity, and entry diploma are examined first for all students; it was hypothesized these would not display an effect. Next, level of participation in the new program is considered for the experimental group; hypothesis 1 predicts that higher levels of participation lead to an increased performance. Third, the effect of English ability of the experimental group on student performance is investigated; it is hypothesized that higher English ability would coincide with improved performance. Finally, the error patterns and types of the experimental and control group are compared; hypothesis 1 posits that the new course should have had a positive effect, creating a change in the pattern, number, and types of errors.

In order to test hypothesis 1, three measures were employed. To begin, the background factors were extracted from the Osiris system for all students. Second, the questionnaire was analyzed for the amount of participation in the program as reported by the students themselves. The last step was to examine all students' final reports for errors in the three aspects of information literacy included in this research: APA Application, Source Integration, and Trust.

Background factors of the control and experimental group. Analyses were performed controling for the background factors gender, nationality, ethnicity, and entry diploma. The information on the background factors was extracted from the Osiris system, and the error rates stem from the mistakes on APA Application, Source Integration, and Trust students made in their final reports. These four background factors were assumed to not influence the error rates.

Males ($n = 63$) and females ($n = 124$) did not differ significantly in their error rates. Similarly, no gender difference was found only within the experimental group with 34 males and 81 females. Therefore, gender appears not to be an influencing factor in this design.

Non-Dutch students ($n = 20$) and Dutch students ($n = 167$) did not differ significantly on their error rates. Within the experimental group non-Dutch students ($n = 15$) differed significantly from Dutch students ($n = 100$) on APA Application ($t(24.63) = -2.23$, $p = .035$), but not on the other three error rates. Considering the small sample size of the non-Dutch students, no conclusions can be drawn for this lone significant result. Therefore, nationality is not a highly significant factor in this design.

A one-way between groups ANOVA was conducted comparing ethnicity (white Dutch $n = 102$, white non-Dutch $n = 19$, and non-western Dutch $n = 66$) to the error rates. No significant differences were found for APA Application, Source Integration, or Total errors, but a significant result appeared for Trust [$F(2, 184) = 3.37$, $p = .036$]. Post hoc comparisons using Tukey HSD test indicated that for Trust non-

western Dutch students (M = .85, SD = .17) differed significantly from non-Dutch white students (M = 3.05, SD = 1.59). This seems to indicate that white foreign students were performing worse; however, this group consisted of only 19 students. A one-way between groups ANOVA comparing ethnicity for only the experimental group (white Dutch n = 65, white non-Dutch n = 15, and non-western Dutch n = 35), yielded no significant results.

A one-way between groups ANOVA was conducted comparing the diplomas with which students had entered the program to their error rates. Students could have a Dutch high school diploma (n = 113), a 21+ test (n = 13), a vocational college degree (n = 37), or a foreign diploma (n = 24). Results revealed no significant differences for APA Application and Trust, but those for Source Integration [$F(3, 183) = 4.92$, $p = .003$] and Total errors [$F(3, 183) = 3.84$, $p = .011$] were significant. Post hoc comparisons using Tukey HSD revealed significant differences for Source Integration for students with a foreign diploma compared to students with either a Dutch diploma or vocational college degree. For Total errors only the difference between foreign diploma and the Dutch diploma was significant. A one-way between groups ANOVA comparing diplomas for only the experimental group (Dutch high school n = 65, 21+ test n = 8, vocational college n = 23, and foreign diploma n = 19) showed similar results: Source Integration [$F(3, 111) = 5.14$, $p = .002$] and Total errors [$F(3, 111) = 4.59$, $p = .005$] differed significantly. Post hoc comparisons using Tukey HSD revealed significant differences for Source Integration for students with a foreign diploma compared to students with a Dutch diploma or the 21+ test. On Total errors, again only the difference between foreign diploma and the Dutch diploma reached significance. These results indicate some effect for diploma, but the sample size for students with a foreign diploma is small.

The background factors are unlikely to influence error rates, so this part of hypothesis 1 cannot be rejected. Gender and nationality displayed no effects on the error rates. Ethnicity demonstrated one effect for Trust, but due to the small sample size of white foreign students, no conclusions can be drawn. The previous education seems to affect some aspects of information literacy, with foreign students performing worse. However, due to the small sample size of foreign students, no definitive conclusions can be drawn.

Participation in the new program for the experimental group. The questionnaire asked students in detail how much they participated in the formative program (see Appendix C, question 1 and 2). The answers to these questions were combined with the errors students had made in their reports on the three information literacy categories: APA Application, Source Integration, and Trust. It was assumed that increased participation would coincide with fewer errors in information literacy.

Out of 102 completed questionnaires, 19 students did not supply their student number, so they could not be matched to their reports. An additional six students had completed the questionnaire, but not their

reports. Therefore, all information was available for 77 students: 43 Dutch and 34 English.

Answers to question 2 (Exactly what parts of the booklets did you do?) were coded on a five-point scale per booklet to reflect the student's effort as Table 3 illustrates. Scores were consequently tallied per student to reach a total completion score between 0-25.

Table 3

Coding for Booklet Completion

	Booklet completion		
Code	Final paragraph	Exercises	Feedback session
0	-	-	-
1	v	-	-
2	-	v	-
3	v	-	v
4	v	v	-
5	v	v	v
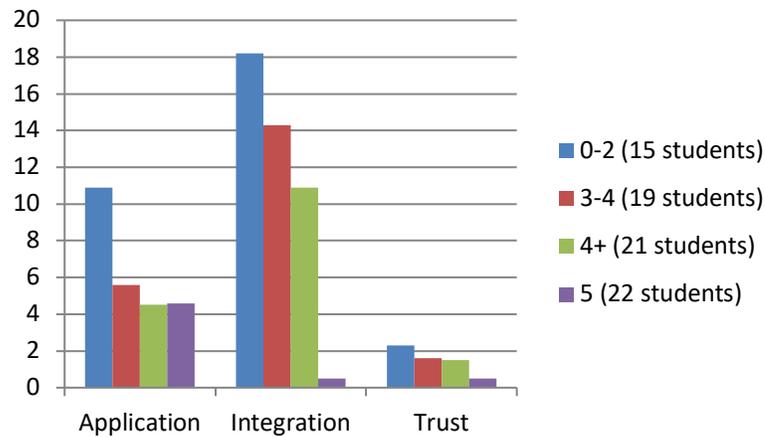
Figure 2. Error patterns based on number of booklets completed

Significant yet moderate negative correlations were found between booklet completion and three measures: APA Application ($r = -.37, p = .001$), Source Integration ($r = -.42, p < .001$), and Total errors ($r = -.47, p < .001$). The correlation with Trust was also negative, but small and not significant ($r = -.15, p = .202$), probably because of a floor effect. However, the negative direction indicates that those who spent more time on the booklets tended to make fewer mistakes in their sources. When grouping students

according to their booklet completion, the number of errors displays a downward trend, especially on Source Integration, but not on APA Application (see Figure 2).

The connection described above holds for full completion rates, so correlations were also drawn per booklet. These bivariate correlations ($n = 70$) showed significant effects for booklet 1 and APA Application $r = $ -.30, $p = .011$), booklet 2 and APA Application ($r = $ -.31, $p = .009$), booklet 3 and APA Application ($r = $ -.29, $p = .017$), booklet 4 and APA Application ($r = $ -.26, $p = .028$), booklet 5 and Source integration ($r = -.43$, $p < .001$), and booklet 5 and Total errors ($r = $ -.39, $p = .001$). All correlations were negative, so higher completion of that specific booklet is linked to fewer errors of the given type. Students who do not start the program or drop out along the way tend to make more mistakes, particularly in APA Application. Booklet 5 is especially linked to Source Integration. However, most students who stopped making booklets did not return to the program. Only seven students were found to return after skipping a booklet; otherwise, once dropped out of the program, they did not restart. In addition, completion rates declined progressively. Although booklet 1 was completed fully by 80 students out of the 91 who answered this question, the number steadily dropped to 30 by booklet 5.

In conclusion, the significant correlations indicate that the more students participated in the program, the fewer mistakes they made. This means that this part of hypothesis 1 can be rejected. The significant correlations per booklet reflect what is already found in the overall completion rates. They do not indicate that students who completed a specific booklet acquired certain identifiable knowledge or skills from it.

Effect of English ability for the experimental group. The experimental group consisted of two streams, Dutch and English, that differed in English ability. Their previous English scores, both in high school and at IBL, were extracted from the Osiris system. This was combined with the number of errors they made in their research reports on APA Application, Source Integration, and Trust. It was presumed that students with better English would outperform those with worse English because the program was fully taught in English and required them to write in English.

Bivariate correlations with errors and English ability found significant, positive relationships with APA Application ($r = .24$, $p = .010$), Source Integration ($r = .26$, $p = .005$), and Total errors ($r = .33$, $p < .001$). The correlation with Trust was weakest and slightly above significance ($r = .18$, $p = .056$), probably due to a floor effect on this aspect. These results indicate an effect contrary to expectation: higher English ability does not coincide with fewer errors but with more.

This rather unforeseen result can be explained, however, by re-examining the participation measure. The Dutch and English stream differed considerably in their participation: The English stream averaged a completion rate of 13.47 ($SD = 9.53$), but the Dutch stream displayed an average of 21.47 ($SD = 4.79$).

An Independent Samples T-Test confirmed this is a significant difference ($t(46.04) = 4.47$, $p < .001$). When drawing partial correlations with errors and English ability using stream as the control variable, none of the correlations reached significance. Therefore, the positive correlations with errors and English ability are actually another reflection of participation in the program.

In conclusion, students with higher English ability do not perform similarly to those with lower ability; in fact, the higher ability correlates with weaker performance. However, this effect can be linked to the participation in the program. Particularly in Honors class 214E students had decided the booklets were beneath them and completed none or very little. The Dutch stream, in contrast, probably saw more relevance in the program because they do not study in English. They apparently acquired more mechanical and integrative skills along the way despite their poorer English skills. Therefore, this part of the hypothesis cannot be rejected.

Error patterns for the control and experimental groups. The most relevant test for the effectiveness of the new program stems from a comparison on how the groups performed on the summative test. All reports were analyzed for error rates on the three categories APA Application, Source Integration, and Trust. Error patterns on these three aspects of information literacy were expected to be different between the control and experimental groups. In previous years, students seemed to struggle most with Source Integration, second APA Application, and last Trust. Therefore, the control group was expected to follow that pattern, but the experimental group's pattern should reflect an adjustment as a result of the redesign that emphasized those aspects with the 4C/ID approach. Within the experimental group, the English stream was expected to perform better than the Dutch stream due to the differences in English ability.
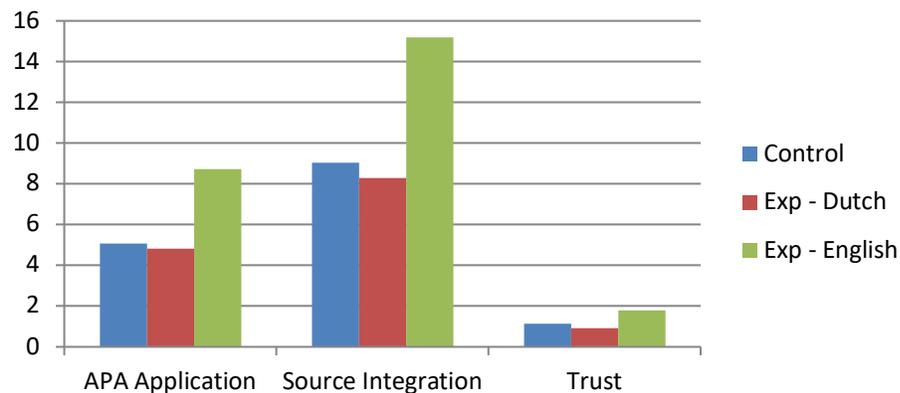
Figure 3. Error patterns for the control and experimental groups based on average number of errors per student.

Results revealed an unchanged error pattern across the three groups (see Figure 3): Source Integration had the most errors in each group, then APA Application, and finally Trust. The English experimental group displayed the most errors while the Dutch group scored on par with the controls.

A one-way between groups ANOVA was conducted to compare the errors in the three groups. There were significant effects for APA Application [$F(2, 184) = 8.25, p < .001$], Source Integration [$F(2, 184) = 10.86, p < .001$] and Total errors [$F(2, 184) = 16.05, p < .001$]. Post hoc comparisons using Tukey HSD test indicated that the mean scores for the English group differed significantly from the Dutch group and the control on those three error categories. On average, the Dutch group made fewer errors than the control group, but the difference was not significant. This suggests that the Dutch group and the control perform similarly on the error categories, but that the English group scores worse on APA Application, Source Integration, and Total errors. Considering the experimental group was expected to perform best, and the English stream in particular, these results indicate the opposite.

Because the Honors class 214E had such deviating participation rates, it was considered that the diverse error rates for the English stream perhaps only resulted from that class. Two one-way between groups ANOVA analyses were conducted to compare the three error categories per class within the two streams. As expected, the three Dutch classes displayed no significant results on the error rates. The English stream showed a significant difference only on APA Application [$F(2, 64) = 6.29, p = .003$]. The post hoc comparisons using the Tukey HSD showed that class 211E differed significantly from both 213E and 214E. The other error categories did not have significant differences, and the results showed that also 213E differed. Therefore, it remains warranted to consider the English stream as a single group.

Error types for the control and experimental groups. The section above groups the errors together in the three categories APA Application, Source Integration, and Trust; however, this possibly fails to reveal more detailed patterns in the error types. APA Application consists of nineteen mistakes, Source Integration includes eight and Trust two. All reports were analyzed for error rates of each specific type to see if more specific issues of information literacy differed between the control and experimental groups. It was expected that the experimental group would not display certain error types anymore or show them less frequently than the control group because the redesign had taught them how to eliminate those.

First, it was examined whether certain errors never occurred in certain groups. Table 4 highlights that only infrequent errors were absent from certain groups in no particular pattern. Therefore, the new design has not taught students to eliminate certain error types.

Table 4

Errors Present (Y) or Absent (N) per Group

Error (number of errors in all groups)	Control	Exp-Dutch	Exp-English
"and" in brackets (6)	Y	N	Y
& outside brackets (2)	Y	N	N
Comma missing author, year (8)	Y	N	N
Et al. at first mention (4)	Y	Y	Y
Repeating (author, year) (34)	Y	N	Y
No space between author and year (8)	Y	Y	N
Publisher or title used as author when author is known (22)	N	Y	Y
Invalid source specified (Word) (3)	Y	Y	N
Year not in brackets (3)	N	Y	N
Year before author name (6)	N	N	Y
Space before comma (20)	N	N	Y
No space before or after citation (13)	N	N	Y
Providing wrong source (3)	N	Y	Y

Next, it will be investigated whether certain errors occurred more frequently in certain groups than in others. Errors made only 10 times or fewer were removed, which left 25 errors. Percentages were calculated per error type: What percentage of the all errors of that type were made by each different group?

Table 5

Descriptive Statistics for All Groups with Expected Means and Confidence Intervals (CI)

Group	Range		M		SD	95% CI
	Minimum	Maximum	Found	Expected		
Control	.00	1.00	.34	.39	.22	[-.04, .82]
Exp-Dutch	.00	.48	.16	.26	.11	[.03, .48]
Exp-English	.08	1.00	.51	.36	.24	[-.01, .84]

Assuming that each group was equally likely to make a certain error and given the varying sizes of the groups, it could be determined to what extent these error distribution patterns conformed to expectations. Assuming a 95% confidence interval meant that any scores ± 1.96 standard deviations of the expected mean were significantly different. Table 5 offers the descriptive statistics and Table 6 the percentages per error for all groups.

Table 6

Error Percentages for All Groups[a]

Error (number of errors in all groups)	Group		
	Control	Exp-Dutch	Exp-English
Misplaced period (301)	.35	.12	.52
Separate brackets for authors (12)	.42	.08	.50
Initial or first name used (41)	.37	.15	.49
Title or publication used (68)	.72	.04	.23
Author also in brackets (113)	.32	.48	.20
nd or ND for no date (75)	.40	.20	.40
Para no for quotation missing (138)	.24	.20	.57
Mismatch with bibliography (151)	.29	.37	.34
Misusing a/b after year (137)	.07	.12	.81
Repeating (author, year) (34)	.21	**.00**	.79
Publisher or title used as author when author is known (22)	.00	.09	**.91**
Only first author mentioned (17)	.65	.12	.24
Space before comma (20)	.00	**.00**	**1.00**
No space before or after citation (13)	.00	**.00**	**1.00**
APA Application total (1189)	.31	.20	.50
Misplaced citation (68)	.38	.18	.44
Misplaced year (89)	.20	.29	.51
Source missing (1797)	.29	.21	.51
Year missing (127)	.64	.11	.25
Author missing (12)	**.92**	**.00**	.08
Source Integration total (2101)	.31	.21	.49

20

Source not trustworthy (136)	.21	.13	.66
Sentence plagiarized (110)	.50	.23	.27
Trust total (246)	.34	.17	.49
All errors (3536)	.31	.20	.49

[a]Bold indicates percentage falls outside the 95% CI

Table 6 demonstrates that few scores vary enough from the mean to be considered aberrations. The control group only forgot the author more. The Dutch experimental group scored significantly lower than expected on repeating (author, year), space before comma, no space before or after citation, and author missing because these errors did not occur at all in this group. The English experimental group displayed three issues much more than expected: publisher or title used as author when author is known, space before comma, and no space before or after citation. However, for all of these aberrations, the number of errors remains quite low, 12-34, so one mistake more or less appreciably impacts the error percentage.

It can be concluded that for the more common error types, the group errors conform to a normal distribution. Because no specific error separates one or more groups from any other, the redesign has not caused specific errors to be eliminated or made much more or less.

Conclusions. Based on the above, hypothesis 1 cannot be rejected. The background factors did not reveal much impact, and participation was indeed linked to performance. However, English ability had an opposite effect on performance, and the experimental group did not outperform the control group.

Hypothesis 2: Student Perceptions

Hypothesis 2 tests whether students perceived the redesigned course positively or not, also taking their performance into account. This was tested by analyzing the TALQ section of the questionnaire, mixed with informal student discussions as recorded in the field notes. The TALQ results are discussed in general, per stream, per class, and adjusted for participation. Finally, the answers to the open question in the questionnaire and the field notes are discussed before conclusions are reached.

General TALQ results. The overall results from the TALQ section can indicate whether the experimental group as a whole was positive about the course or not. The results of the 102 students who completed this TALQ questionnaire showed fairly high scores for all aspects, with Authentic Problems and Integration scoring lowest and Application and Global Quality the highest scores (see Table 7). No scale scored below 3. Because 3 represents the neutral score, this would be the hypothesized value for all scales. Because this neutral 3 was not included in the 95% CI except for Authentic Problems (see Table E1), those scales differed significantly from the neutral score. Therefore, the students of the experimental

Table 7

TALQ Questionnaire Average Results Overall and per Class

		Classes							
		Dutch				English			
Scale (*SD*)[a]	All	All	201N	202N	203N	All	211E	213E	214E
N	102	49	20	15	14	53	19	11	23
First Principles									
Authentic Problems	3.15 (.78)	**3.41** (.08)	**3.40** (.12)	**3.62** (.14)	3.19 (.16)	2.91 (.12)	**3.42** (.15)	3.06 (.20)	**2.41** (.18)
Activation	**3.49** (.61)	**3.70** (.07)	**3.64** (.10)	**3.81** (.07)	**3.66** (.17)	**3.31** (.09)	**3.83** (.10)	3.31 (.14)	2.87 (.13)
Demonstration	**3.67** (.67)	**3.73** (.07)	**3.62** (.13)	**3.70** (.09)	**3.94** (.12)	**3.61** (.09)	**3.90** (.11)	**3.38** (.15)	**3.47** (.15)
Application	**3.94** (.67)	**4.12** (.07)	**4.10** (.10)	**4.11** (.10)	**4.14** (.16)	**3.78** (.11)	**4.25** (.10)	**3.85** (.14)	3.37 (.19)
Integration	**3.26** (.75)	**3.49** (.08)	**3.38** (.12)	**3.51** (.13)	**3.64** (.15)	3.04 (.12)	**3.76** (.12)	3.00 (.17)	**2.47** (.16)
Total First Principles	**3.49** (.55)	**3.67** (.06)	**3.61** (.09)	**3.71** (.08)	**3.73** (.12)	**3.33** (.09)	**3.82** (.08)	**3.31** (.12)	2.92 (.13)
Other scales									
Academic Learning Time	**3.60** (.59)	**3.67** (.07)	**3.73** (.10)	**3.70** (.14)	**3.57** (.12)	**3.54** (.09)	**3.88** (.12)	**3.59** (.20)	3.23 (.14)
Learning Progress	**3.63** (1.08)	**4.26** (.08)	**4.23** (.15)	**4.23** (.12)	**4.32** (.17)	3.05 (.15)	**3.81** (.20)	3.32 (.27)	**2.29** (.19)
Satisfaction	**3.57** (.94)	**3.98** (.08)	**3.94** (.13)	**3.86** (.15)	**4.16** (.16)	3.19 (.14)	**4.20** (.13)	2.95 (.23)	**2.47** (.16)
Global Quality	**3.80** (0.70)	**3.99** (.08)	**3.93** (.13)	**3.98** (.14)	**4.07** (.14)	**3.63** (.11)	**3.97** (.14)	3.39 (.21)	**3.46** (.19)

[a]Figures in bold indicate 95% CI the score deviates from the neutral 3

group rate the course higher than neutral.

TALQ results per stream. The experimental group can be divided in two streams, Dutch and English, and there is no reason to assume that stream impacts the TALQ results. An Independent Samples T-Test was performed comparing the Dutch and English streams on the questionnaire scales. Significant differences were found for Authentic Problems ($t(90.48) = 3.47$, $p = .001$), Activation ($t(92.98) = 3.44$, $p = .001$), Application ($t(84.78) = 2.62$, $p = .010$), Integration ($t(87.21) = 3.20$, $p = .002$), Total 5 Principles ($t(88.03) = 3.38$, $p = .001$), Learning Progress ($t(80.59) = 6.91$, $p < .001$), Satisfaction ($t(82.45) = 4.75$, $p < .001$), and Global Quality ($t(94.43) = 2.68$, $p = .009$). As Table 7 shows, the Dutch stream students rated all scales consistently higher than the English stream, except for Satisfaction. Considering the further developed language skills of the English stream, this is to be expected on Learning Progress, where the Dutch stream rated 1.21 higher. Except for the English stream's Authentic Problems, students rated all aspects above 3. The 95% CI for all scales (see Table E2) showed that 3 was not included for any of the scales in the Dutch stream while the English stream rated four scales not significantly different from neutral: Authentic Problems, Integration, Learning Progress, and Satisfaction. It can be concluded that the Dutch stream students rate all parts of the course above neutral, but that the English stream is only neutral on four of the aspects.

TALQ results per class. Because the streams differed in their perceptions of the course, it is relevant to examine this in more detail by analyzing the TALQ results per class (also see Table 7). The experimental group consisted of six classes, three Dutch and three English, who evaluated the course quite diversely. All Dutch stream classes rated the course significantly higher on all scales (also see Table E3). As Table E4 illustrates, this also held for class 211E, but 213E and 214E displayed another pattern. Class 213E evaluated Authentic Problems, Activation, Integration, Learning Progress, Satisfaction, and Global Quality not significantly different from neutral; they rated only Demonstration, Application, Total First Principles, and Academic Learning Time higher. Class 214E rated Demonstration and Global Quality significantly higher, but did not veer from neutral on Activation, Application, Total First Principles, and Academic Learning Time. Unlike any of the other classes, though, class 214E rated Authentic Problems, Integration, Learning Progress, and Satisfaction significantly lower than neutral. In conclusion, the Dutch stream and class 211E perceive the course positively, class 213E somewhat positively, but class 214E mostly neutrally or negatively.

TALQ results and performance. It is likely that student performance influences how they rate the course. The results for hypothesis 1 showed that participation was the core measure for performance. Also, this measure (booklet completion) yields a single score per student rather than four (the three error categories plus total), so the analyses can be simplified by studying participation instead of the error rates. As stated before, the Dutch stream classes completed much more of the formative booklets with an average of 21.00 ($SD = .71$) than the English stream with 12.3 ($SD =$

1.28). An Independent Samples T-Test confirmed this is a significant difference ($t(80.54) = 5.93$, $p <$ .001). It was expected that students who performed (so participated) less would rate the course lower.

First, the TALQ results are examined for those who scored "0" on booklet completion. This is a very small group of eleven students, but their ratings are considerably lower than even those who completed one booklet. Seeing as they did not actually participate, they logically experienced little Learning Progress ($M = 2.02$, $SD = .26$) or felt they did not work on Authentic Problems ($M = 2.12$, $SD = .26$). This small group displays the first indication that the participation (performance) and course evaluation are linked as expected.

Correlations between booklet completion and all scales indicate moderate positive relations with Authentic Problems ($r = .53$, $p < .001$), Activation ($r = .52$, $p < .001$), Application ($r = .52$, $p < .001$), Integration ($r = .49$, $p < .001$), Total First Principles ($r = .52$, $p < .001$), Academic Learning Time ($r = .38$, $p < .001$), Learning Progress ($r = .61$, $p < .001$), and Satisfaction ($r = .47$, $p < .001$). These correlations indicate that in general, high ratings on the course coincide with high booklet completion, and therefore that students who performed better rated the course higher.

A step-wise multiple regression was conducted to predict booklet completion from all scales. The model ($F(4, 97) = 23.00$, $p < .001$, $R^2 = .487$) consisted of Learning Progress, Application, Demonstration, and Time. Learning Progress and Time were directly linked to booklet completion: Spending more time on the booklets means completing more of them and experiencing more progress. On its own, Learning Progress reached an R^2 of .367 ($F(1, 100) = 58.00$, $p < .001$).

Reasons for participation. The questionnaire offered further insight into the reasons for completing the booklets or not by answering question 3 of the questionnaire (see Appendix C). Students provided five main reasons for completing all or some of the booklets: because they are helpful (6), because it is practice for the final (13), because it increases knowledge (12), to obtain a high score (5), and because they were mandatory (9). Some students erroneously assumed the booklets were compulsory. One class representative approached the researcher about this in week 10 as students thought policies unfairly varied across teachers, a view that was swiftly corrected. The reasons for not completing any or some of the booklets proved to be time (24), lack of relevance (8), and a combination of time and relevance (8). During the classes, students complained about the length of the booklets because this required them to spend much time working on the problems. The booklets were intended to cover the lost in-class time, but various students noted they needed more time to complete the work. Numerous students report that the required time became too high a burden, and that led them to discontinue their participation.

Conclusions. Hypothesis 2 aimed to discover whether students perceived the course positively or not. It can be rejected as most students responded positively to the redesigned course. Although some may have dropped out due to time constraints or not finding the program relevant, those who

participated gave higher ratings than neutral. In fact, as expected, the better students performed (i.e. the more they participated), the more positive they were. Dutch stream students were particularly pleased with their learning progress. The Honors class was negative about certain aspects of the course, but they also had the lowest participation rates. It seems logical that students cannot judge a course if they have not actually taken part.

Hypothesis 3: Teacher Perceptions

The third hypothesis of this research tests whether teacher perceptions of the course were positive or not. This was assessed via two semi-structured, fairly informal interviews with each teacher individually. The interview questions (see Appendix D) focused on their perceptions of the course setup, feedback sessions, and effect on students. After discussing their comments in detail, conclusions will be drawn about hypothesis 3. To clarify when a comment was made, -1 is added to those made during the mid-way interview and -2 to those in the final interview.

Course setup. Both teachers agreed the booklets constituted a sound approach to offer useful practice and feedback to students. Teachers J-1 liked the material in the booklets. Teacher N-1 mentioned "the progression is good because it starts off with easier stuff and then builds up their knowledge". Effective aspects were the one-on-one sessions (J-2), the ON/OFF schedule (N-2), and the booklet 6 kick-off session (N-2). Teacher J-1 reported her students appreciated the one-on-one time.

All teachers employed the same booklets, but they added their own input when needed. Instructor J-1 allowed students to skip exercises as long as it was not detrimental. Teacher N-1 commented that some parts were easy for his Honors class students, but regarding their refusal to take part, he said: "I kind of found it not really valid, their criticism, because a lot of it came at the onset of the course before we'd even really looked at this stuff." He explained how the Honors class students had very poor attendance in all of their courses and that the late hour in the day affected the students adversely. He also posed that students could have been skipping English classes due to being busy in other courses. Teacher N-1 further commented some students were confused about the instructions, "although I think the instructions were pretty clear overall" (N-1).

Regarding possible drawbacks of the new design, the teachers mentioned as the main issues length and lack of class time for writing. Both instructors highlighted the potentially adverse effect of length: "too long was a major, was a big comment" (J-1). The solution is either shortening the booklets (N-2) or reducing the number of booklets to three (J-2). Teacher N-2 liked the idea of students having more autonomy in deciding which exercises to complete based on their previous knowledge and experience. Teacher J-1 commented that the navigation and reader-friendliness of the booklets could use some more work. Additionally, both teachers would like to see some writing covered in class. Teacher J-2 wanted to work in class on bibliography, structuring, and referencing practice. Teacher N mentioned

this already during the mid-way interview and observed in the final interview that students could have worked more in groups and benefited from each other's knowledge.

When asked to rate their experience with the course, the teachers were positive. Teacher N-2 awarded his experience a 7 and J-2 an 8 out of 10. Teacher N-2 was "very positive about the way the course is designed", but also commented that "most of them [the students] were very passive about it". Teacher J-2 gave higher marks due to the one-on-one feedback.

Feedback sessions. Both instructors commented that the on/off schedule worked for them. Teacher J-1 said that five minutes is a very brief time with the students, so "I feel a bit panicked during the feedback sessions with time. I don't feel like I have enough time with each student." Teacher N-1 liked the setup, remarking that "it forces them to be independent." One student commented to the researcher that he would have preferred the more traditional in-class time rather than the on/off schedule, but other students seemed pleased about the reduced contact time. Teacher J-1 worked with Calendly to organize the feedback sessions, but she did not find it very user-friendly. Teacher N-1 decided against using Calendly. His class took place in the late afternoon, and he had so few students attending the feedback sessions that it did not feel useful to him.

During these feedback sessions, teacher J-2 reported giving more corrective feedback, but teacher N-2 saw more need for cognitive feedback. Teacher J-2 more often commented on the technical aspects of APA style, timing of references, when to use quotations, and layout: "Where they break up a paragraph on another line, instead of [creating a visually coherent piece]. Or they haven't; it's all part of the same paragraph. It drives me nuts." Teacher N-2's Honors class performed poorly in the parts that required deeper reflection. He spoke about students misjudging the trustworthiness of sources, not supporting statements with a source, or making generalizations and assumptions, particularly in case of the cultural topic. He remarked how the information about generalizations, covered in class with a view to the oral debate, did not carry over into the writing. Both teachers also explained on what they complimented the students: sentence flow (J-2), including references (J-2), register (N-2), and style (N-2).

Effect on students. The teachers assessed their class's performance in a varied way. Teacher N-2 spoke very negatively about his class: "[I]t was just piss-poor, to put it mildly". Their reports were acceptable, but they exhibited a poor attitude during the course. He could not say whether the Honors class simply did not need such a supported setup or whether they did all the work at the end. Teacher J-2 was more positive, rating her two groups at 7.5. She based this on the fact that only one group did very badly on the report and because "there were probably two good papers in each class, two wow papers".

Considering the effect on performance, teacher J-1 reported improvements: "Yes, yes, I am [seeing progress]. With some students. … I think that that's just the level of effort that the individual student

puts forward" (J-1). Teacher N-1, due to the low level of involvement from his students, felt he could not comment on this. He reported: "I know from the students themselves that they said, you know, this is helpful." He declined to comment on the final report because he could not compare the performance (N-2). Teacher J-1 saw options for students to take more ownership of their learning.

   Conclusions. In general, teachers perceived the course positively, so hypothesis 3 can be rejected. They appreciated the course design with the booklets. Because the Honors class of teacher N exhibited poor participation, he felt he could not remark on the effect, but teacher J saw positive outcomes. She felt the feedback sessions were effective in helping the students advance in writing. Their points for improvement consisted of reviewing the booklets' length and returning some writing instruction to the ON classes. Because the researcher also designed of the program, the teachers possibly felt reluctant to be more critical when explicitly asked to grade the program. However, their criticism and points for improvement revolved around certain organizational aspects, less the quality of tasks.

<div align="center">

Chapter 4: Conclusions and discussion

</div>

This thesis aimed to investigate the effects on teachers and students in higher education in a course on information literacy designed according to the 4C/ID model. The examined variables were student performance, student perceptions, and teacher perceptions. Each of these variables was linked to its own hypothesis (also see page 6).

The Hypotheses

Hypothesis 1 tested whether performance in the new program differed from the previous program, taking English ability, participation, and background factors into account. This hypothesis cannot be rejected fully. Student performance in the redesigned program was not improved compared to the previous program, as evidenced by the identical error patterns. Previous research (e.g., Van Merriënboer et al., 1992; Melo & Miranda, 2014; Brand-Gruwel et al., 2005; Sarfo & Elen, 2007; Lim & Reiser, 2006) suggested that a 4C/ID design improves performance, but this could not be determined here. Students in the new course performed similarly or even worse than those in previous years or resitters, and error patterns remained unchanged. It was interesting to find that English ability ran contrary to expectations: The Dutch stream with the lowest English ability performed best. However, the reason for this can be found in the participation in the design: as expected, completing more of the program resulted in fewer errors. Because the English stream contained more students who chose not to complete the formative practice booklets, they did not acquire the necessary knowledge and skills. This aligns with findings from Kicken et al. (2009) where poorer students used a voluntary self-directed learning system more frequently. The linguistically poorer Dutch students in this research also partook more in the optional sections of the program than those with higher abilities. The background factors gender, ethnicity, and nationality appear not to affect the errors

made, but some effect was found for the entry diploma. Foreign students performed worse than students with a Dutch diploma on some errors, but this was again linked to participation.

Hypothesis 2 tested whether students perceived the new course positively or not, also considering their performance. This hypothesis can be rejected. Some students were confused about the formative nature of the booklets, and they felt the booklets were too long, but overall, students awarded the course with high ratings. Similar positive students evaluations of 4C/ID designs had been found by Holtslander et al. (2012) and Wopereis et al. (2015). The English stream was less positive than the Dutch stream; the English Honors class 214E even scored various factors negatively. However, student perception was connected with performance as measured by the level of participation, and the Honors class had the lowest rates on that measure of all classes.

Hypothesis 3 tested whether teachers perceived the redesign positively or not. This hypothesis can be rejected as well. Teachers evaluated the course well despite high dropout rates and poor student performance. They appreciated the materials, the structure, and the feedback sessions. These are the same elements that teachers valued in Wopereis et al.'s (2016) study. That study had also found few differences between teacher and student perceptions, but in the current research student perceptions varied more. Teacher comments concerned the length of the booklets and the lack of in-class time for writing. It should be remembered that they might not have felt comfortable expressing more critical comments because the researcher was also the program designer.

Interactions Between the Variables

Regarding the main interactions between the variables, one conclusion can be drawn. The main direction of influence was assumed to be teacher perception influencing student perception and student perception influencing student performance. The current research design did not offer insight into the connection between teacher and student perception, so this interaction cannot be determined. However, student perception and student performance were linked: Students who performed better (i.e. participated more) also perceived the course more positively. This finding concurs with those of Braga, Paccagnella and Pellizzari (2014) and Carrell and West (2010) whose students also rated their teachers higher if they received higher scores.

Furthermore, the variables were assumed to display further interactions with student performance influencing student perception, student performance influencing teacher perception, and student perception influencing teacher perception. As mentioned before, the interaction between student performance and perception runs via the fundamental factor of participation. The reciprocal interaction between student and teacher perceptions, as Jossberger et al. (2015) found, was not clearly identified in this research. Teachers assessed the course positively, irrespective of what the student thought of it. Additionally, they even remained positive about the redesign when students displayed poor performance. Although it is possible they were evaluating the students' English more than their

application of information literacy concepts, their comments in the interviews highlight the teachers' awareness of the students' flaws in this area.

Conclusions

To answer the main question, the current research offers insight into the practical execution of a 4C/ID design in higher education on the subject of information literacy, expanding upon the work of Wopereis et al. (2015) and Wopereis et al. (2016). The positive teacher and student perceptions serve as an endorsement of the 4C/ID model to create designs that both teachers and students appreciate. Performance of students was linked to their participation, but not to English ability or background factors. It cannot be determined that the new program was an improvement on the old one when examining the error patterns, but keeping the students on board is clearly crucial.

Additionally, this research highlights the fact that linguistic skills and information literacy are two different constructs, and that affects teaching in higher education. The cognitive load required for researching and writing would be higher if one had a lower ability in that language. All students had already passed the minimum threshold needed to complete the report writing task in a foreign language, but still the Dutch and English groups differed in their English levels. Because those students who participated more in the program practiced the whole task more and received more individual feedback, their skills in information literacy increased as evidenced by their lower error rates. The unexpected direction of the result (higher English ability correlated with more errors) underlines the relevance of teaching information literacy to students of all language levels. Earlier research had also found the need for explicit instruction to achieve information literacy (e.g., ACRL, 2000; Larkin & Pines, 2005; Walton & Archer, 2004). Because the students in the English stream opted out of the voluntary parts of the program, special attention should be devoted to convincing such students of the necessity to learn more about information literacy. A student with a high level of language may still need to work on applying APA style, integrating sources into a text, selecting trustworthy sources, and avoiding plagiarism. This is supported by Wopereis, Brand-Gruwel and Vermetten (2008) who compared a course with and without embedded instruction on IPS for students of Psychology. Students who had followed the embedded course regulated their search process more actively and judged the information found more often than those whose course had not included this component. Their online course had been taught in the students' native language, so linguistic ability was not an issue, but the instruction impacted student behavior in this particular aspect. In order to achieve the desired higher level of participation, the redesigned course needs to be adjusted based on the student and teacher comments, mainly the length.

Limitations

The current research faces some clear limitations. Findings suggest that amount of completion shapes student performance, and this supports the idea that the 4C/ID model is effective. However, this effect

might have also been observed in an alternative design. Although the students who participated demonstrated more of the knowledge and skills intended to be acquired than those who avoided the booklets, this could very well be due to the increased levels of practice and feedback they received, rather than the whole-task approach which according to Van Merriënboer and Kirschner (2013) sets the 4C/ID model apart from other methods. No specific element of the design could be identified as more effective than another. Students dropped out at some point, usually to not return anymore until the final report. Therefore, inferences can be based only on how much of the full program was completed, and not the influence of a specific booklet.

Further limitations to this research stem from time constraints. Long-term transfer effects outside of the immediate course, such as the internship or third-year project, could not be included. Therefore, a logical follow-up study tracks this student subset longitudinally. Also, a slightly modified version of the course will run in the spring semester of the 2017-2018 academic year, and this allows for additional future comparison.

Apart from those methodological and time-related issues, the research is limited to one course (English semester 4) for students in a specific program (IBL) at one Dutch institute of higher education (Rotterdam University of Applied Sciences) at bachelor's level. This restricts generalizability. More insight into the validity of the 4C/ID model could be gleaned from extending the design across several IBL-courses related to information literacy, to other programs in- or outside the business field, to other institutes, or to programs at a different level than bachelor. Additionally, the research focuses on certain aspects of information literacy, described in the error types of APA Application, Source Integration, and Trust. Other aspects of information literacy, such as defining a research question, key-word selection, and organizing the found information, are not or partially covered in the current design. Errors in these areas are therefore not taken into account here. Not only could students benefit from including said aspects, but also the research scope could be extended to include more subskills. The APA Application aspect is highly mechanical and therefore the likeliest candidate to replace by a subskill that requires more in-depth analysis on the student's part.

References

Association of College and Research Libraries. (2010). *Information literacy competency standards for higher education.* Retrieved from:
http://www.ala.org/acrl/sites/ala.org.acrl/files/content/standards/standards.pdf

Association of College and Research Libraries. (2015). *Framework for information literacy for higher education.* Retrieved from:
http://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/infolit/Framework_ILHE.pdf

Bailey, S. (2015). *Academic writing for international students of business.* New York, NY: Routledge.

Braga, M., Paccagnella, M., & Pellizarri, M. (2014). Evaluating students' evaluations of professors. *Economics of Education Review, 41*, 71-88. doi:10.1016/j.econedurev.2014.04.002

Brand-Gruwel, S., Wopereis, I., & Poortman-Cremers, S. (2005). Informatieproblemen oplossen geïntegreerd in het curriculum: Een pilot binnen een Lerarenopleiding. [Information problem solving integrated in the curriculum: A pilot in a teacher's college.] *Tijdschrift voor Hoger Onderwijs, 2,* 88-104.

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education, 53*, 1201-1217. doi:10.1016/j.compedu.2009.06.004

Brand-Gruwel, S., Kammerer, Y., Van Meeuwen, L., & Van Gog, T. (2017). Source evaluation of domain experts and novices during Web search. *Journal of Computer Assisted Learning, 33,* 234-251. doi:10.1111/jcal.12162

Carrell, S. E., & West, J. E. (2010). Does professor quality matter? Evidence from random assignment of students of professors. *Journal of Political Economy, 118*(3), 409-432.

Creswell, J. W. (2014). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (4th ed.). Harlow, UK: Pearson.

Dolmans, D. H. J. M., Wolfhagen, I. H. A. P., & Van Merriënboer, J. J. G. (2013). Twelve tips for implementing whole-task curricula: How to make it work. *Medical Teacher, 35,* 801-805.

Frick, T. W., Chadha, R., Watson, C., & Zlatkovska, E. (2010). New measures for course evaluation in higher education and their relationship with student learning. Paper presented at the annual conference of the American Educational Research Association, Denver, CO.

Harari, Y. N. (2017). *Homo deus: A brief history of tomorrow.* London: Harvill Secker.

Holtslander, L. F., Racine, L., Furniss, S., Burles, M., & Turner, H. (2012). Developing and piloting an online graduate nursing course focused on experiential learning of qualitative research methods. *Journal of Nursing Education, 51*(6), 345-348.

Jossberger, H., Brand-Gruwel, S., Van de Wiel, M. W. J., & Boshuizen, P. A. (2015). Teachers' perceptions of teaching workplace simulations in vocational education. *Vocations and Learning, 8*, 287-318.

Kicken, W., Brand-Gruwel, S., Van Merriënboer, J. J. G., & Slot, W. (2009). Design and evaluation of a development portfolio: How to improve students' self-directed learning skills? *Instructional Science, 37,* 453-473. doi:10.1007/s11251-008-9058-5

Larkin, J. E., & Pines, H. A. (2005). Developing information literacy and research skills in introductory psychology: A case study. *The Journal of Academic Librarianship, 31*(1), 40-45.

Lim, J., & Reiser, R. A. (2006). The effects of part-task and whole-task approaches on acquisition and transfer of a complex cognitive skill. Paper presented at the Association for Educational Communications and Technology (AECT), Annual Convention, October 10-14, Dallas, TX.

Melo, M., & Miranda, G. L. (2014). Applying the 4C-ID model to the design of a digital educational resource for teaching electric circuits: Effects on student achievement. In H. M. Fardoun & J. A. Gallud (Eds.), *Proceedings of the 2014 workshop on interaction design in educational environments* (pp. 8-14). doi:10.1145/2643604.2643605

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development, 50*(3), 43-59.

Partnership for 21st Century Skills (2009). *Framework for 21st century learning.* Retrieved from: http://www.p21.org/storage/documents/P21_Framework_Definitions.pdf

Redecker, C., Leis, M., Leendertse, M., Punie, Y., Gijsbers, G., Kirschner, P. A., Stoyanov, S., & Hoogveld, B. (2011). *The future of learning: Preparing for change.* Sevilla, Spain: European Commission, Institute for Prospective Technological Studies. Retrieved from: http://ftp.jrc.es/EURdoc/JRC66836.pdf

Sarfo, F. K., & Elen, J. (2007). Developing technical expertise in secondary technical schools: The effect of 4C/ID learning environments. *Learning Environments Research, 10*(3), 207-221.

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296. doi:10.1023/A:1022193728205

Van Merriënboer, J. J. G., Jelsma, O., & Paas, F. (1992). Training for reflective expertise: A four-component instructional design model for complex cognitive skills. *Educational Technology, Research and Development, 40*(2), 23-43. doi:10.1007/BF02297047

Van Merriënboer, J. J. G., & Kirschner, P. A. (2013). *Ten steps to complex learning: A systematic approach to four-component instructional design* (2nd ed.). New York, NY: Routledge.

Van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist, 38*, 5-13. doi:10.1207/S15326985EP3801_2

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*, 147-177. doi:10.1007/s10648-005-3951-0

Walton, M., & Archer, A. (2004). The Web and information literacy: Scaffolding the use of web sources in a project-based curriculum. *British Journal of Educational Technology, 35*(2), 173-186.

Wilson, T. D. (1999). Models in information behavior research. *Journal of Documentation, 55*(3), 249-270.

Wopereis, I., Brand-Gruwel, S, & Vermetten, Y. (2008). The effect of embedded instruction on solving information problems. *Computers in Human Behavior, 24*(3), 738-752.

Wopereis, I., Frerejean, J., & Brand-Gruwel, S. (2015). Information problem solving in higher education: A case study on instructional design. *Communications in Computer and Information Science, 552,* 293-302. doi:10.1007/978-3-319-28197-1_30

Wopereis, I., Frerejean, J., & Brand-Gruwel, S. (2016). Teacher perspectives on whole-task information literacy instruction. In: S. Kurbanoğlu et al. (Eds.), *Information literacy: Key to an inclusive society. ECIL 2016* (pp. 678-687). doi:10.1007/978-3-319-52162-6_66

Appendices

Appendix A

Blueprint of the English Semester 4 Course

Task class: in order to learn information literacy skills, students carry out six learning tasks. Topic familiarity is above average. Students conclude each learning task with a conventional task that varies in length with the sixth task (group research report) serving as the summative test.	
Supportive information: *present cognitive strategies*	
• Systematic approaches to problem solving (SAPs) from Bailey (2015) for writing summaries, writing a paraphrase, writing quotations, and using academic style. • SAPs about writing contrasting phrases, organizational structures, and writing research questions. • SAPs surrounding the evaluation of sources: a flowchart and a checklist.	
Supportive information: *present mental models*	
• Conceptual models from Bailey (2015) about plagiarism, summaries, paraphrasing, quotations, reference, and citations as well as own designs about Google search results and topic sentences. • Causal models from Bailey (2015) for writing summaries, paraphrases, and punctuating quotations. • Structural model showing the steps in the research report writing process with for each learning task highlighted which steps will be taken in that task.	
Supportive information: *cognitive feedback*	
• During each individual session • During exercises, e.g., when comparing their own judgment about source trustworthiness with an expert's and when having to investigate further pages of the Google search results.	
Learning task 1: *one paragraph about one source* Source is provided, accessible topic and language use. Emphasis is on understanding plagiarism, summarizing, and citations in APA style.	*Procedural information presentation:* • Procedures for using in-text APA style • Procedures for using reporting verbs
Learning task 2: *one paragraph about three sources* Sources are provided, more complex topic and language. Emphasis is on writing topic sentences and organizing structures.	*Procedural information presentation:* • More procedures for using in-text APA style • More procedures for using reporting verbs • Procedures for writing topic sentences
Learning task 3: *one paragraph about three sources* Six sources are provided; students need to select the three trustworthy ones. Topic and language are technical, but not too complex. Emphasis is on evaluating trustworthiness and writing contrasting information.	*Procedural information presentation:* • Procedures for writing topic sentences (fading) • Procedures for using in-text APA style (fading) • Procedures for using reporting verbs (fading)
Learning task 4: *two paragraphs about one source* Students need to find a suitable source in the Marketline database. Fairly complex topic and language. Emphasis is on reading and understanding high level texts and the structure of the Marketline database.	*Procedural information presentation:* • Procedures for using Marketline

Learning task 5: *two paragraphs about two sources* Students need to find suitable sources themselves using Google. Accessible topic, language complexity is student's choice. Emphasis is on executing Google searches and using EndNote for their bibliography.	*Procedural information presentation:* • Procedures for using Google *Part-task practice:* • Creating a bibliography using EndNote
Learning task 6: *group research report + individual two pages with at least three sources each* Students write a research report with two-three classmates. They choose one of three cases or their own topic. Each individual completes two subquestions, maximum one page, with at least three unique text sources. The report's main question and subquestions need to be approved and an extensive outline needs to be made.	*Procedural information presentation:* • Procedures for writing an extensive outline • Procedures for using Google (fading) • Procedures for using Marketline (fading)

Appendix B

Detailed Overview of Error Type Classification

	APA Application	Source Integration	Trust	Example error	Example correction
	Error type				
Misplaced period	X			Sentence. (Walmart, 2017)	Sentence (Walmart, 2017).
Authors in separate brackets	X			(Ikea, 2017) (The Towel Shop, 2017)	(Ikea, 2017; The Towel Shop, 2017)
Initial or first name used	X			J. Jones (2016) (Jack Jones, 2016)	Jones (2016) (Jones, 2016)
Title or publication used	X			(Carrefour our products, 2017) (Bloomberg, Susan Berfield, 2013)	(Carrefour, 2017) (Berfield, 2013)
Author in brackets or repeat	X			According to (Jurevicius, 2017) According to Euromonitor (Euromonitor, 2017)	According to Jurevicius (2017) According to Euromonitor (2017)
"and" in brackets	X			(Mun and Yazdanifard, 2012)	(Mun & Yazdanifard, 2012)
& outside brackets	X			Mun & Yazdanifard (2012)	Mun and Yazdanifard (2012)
nd, ND or sd[1] for no date	X			(European Commission, nd) (European Commission, ND) (European Commission, sd)	(European Commission, n.d.)
No paragraph or page number in quotations	X			"Sentence" (Walmart, 2017).	"Sentence" (Walmart, 2017, para. 3).
Mismatch with bibliography	X			[Citation is not listed in bibliography or student cites publication title instead of author.]	
Comma missing author + year	X			(Gehrmann 2001)	(Gehrmann, 2001)

---

[1] When students use Microsoft Word to create their bibliographies, they may fail to set the language for English, which then produces the Dutch sd instead of n.d.

et al. at first mention	X		(Hofstede et al., 1990)	(Hofstede, Neuijen, Ohayv & Sanders, 1990)
Misusing a/b after year	X		Euromonitor (2017) … Euromonitor (2017)	Euromonitor (2017a) … Euromonitor (2017b)
Abbreviation at first mention	X		RVO (2016)	Rijksdienst voor Ondernemend Nederland (2016)
No space between author and year	X		Dawson(2009) (Marketline,2016)	Dawson (2009) (Marketline, 2016)
Publisher used as author when author is known	X		McKinsey identified… (Breuer & Spilleche, 2010). McKinsey says…	McKinsey identified… (Breuer & Spilleche, 2010). Breuer and Spilleche say…
Invalid source specified	X		[When using Word for citations, this appears when Word cannot find the automatic citation in the bibliography of that file]
Year not in brackets	X		Euromonitor 2010	Euromonitor (2010)
Only first author mentioned	X		Kotler (2016)	Kotler and Armstrong (2016)
Repeating (author, year)		X	Sentence (Hofstede, 2001). Sentence (Hofstede, 2001). Sentence (Hofstede, 2001).	Hofstede (2001) states… He… Moreover…
Misplaced citation		X	It can be concluded that Walmart should not enter Spain (Hofstede, 2001).	… (Hofstede, 2001). It can be concluded that Walmart should not enter Spain.
Misplaced year		X	Euromonitor reports that … (2017).	Euromonitor (2017) reports that …
Source missing		X	Half of the Brazilian population owns a credit card.	Half of the Brazilian population owns a credit card (Bruha, 2015).
Year missing		X	Marketing School has…	Marketing School (2016) has…
Author missing		X	Half of the Brazilian population owns a credit card (2015).	Half of the Brazilian population owns a credit card (Bruha, 2015).
Unclear second citation		X	The same author/article… [when unclear to which it refers]	Johnson also states…
Providing incorrect source		X	According to Debenhams (2017)…	According to the Debenhams case study (Mülders, 2017)…

Untrustworthy source	X	[Source used does not meet general standards of trustworthiness.]
Plagiarism	X	[Student has copied directly from the original without quotation.]

Questionnaire Distributed to Students in Experimental Group

Please complete this questionnaire about English semester 4. **Your answers will not affect your grade.** The data will be used by IBL teacher Chantal Mülders for her Master's thesis in Educational Sciences. By completing this questionnaire you agree that the data be used for this purpose and this purpose only.

Student number: _____   Class: _____

Question 1: The written part of the course included five practice booklets. Did you complete them?
0   Yes, I completed at least one booklet
0   No, I did not do any of the practice booklets (go to question 3)

Question 2: Exactly what parts of the booklets did you do? Circle the relevant answers for each booklet.

Booklet	Exercises	Final paragraph	Feedback session
1	Yes / No	Yes / No	Yes / No
2	Yes / No	Yes / No	Yes / No
3	Yes / No	Yes / No	Yes / No
4	Yes / No	Yes / No	Yes / No *
5	Yes / No	Yes / No	Yes / No

* Either the peer review in class or individually with your teacher

Question 3: What were your reasons for doing or not doing the practice booklets?
_____
_____

Question 4: Did you receive feedback from your teacher on your subquestions? If not, why?
0   Yes, on both subquestions
0   Yes, but only on one of them, because _____
        _____
0   No, because _____
        _____

*Question 5: How much time did you spend on **the written part** of the course?*
IMPORTANT: The written part of this course involved four tasks: (1) working on the booklets, (2) attending feedback sessions, (3) attending the booklet 6 kick-off class, and (4) writing the research report.
0   < 11 hours
0   Between 11 and 22 hours
0   Between 23 and 34 hours
0   Between 35 and 46 hours
0   Between 47 and 58 hours
0   Between 59 and 70 hours
0   > 70 hours

Question 6: Did you submit your final report with your group on June 6? If not, why not?
0   Yes
0   No, because _____
        _____

This section of the questionnaire contains 35 statements. Please circle the correct number for each one. You may notice that some statements reappear in slightly different wording; this is intentional.

1 = strongly disagree; 2 = agree; 3 = undecided; 4 = agree; 5 = strongly agree

1	I performed a series of increasingly complex authentic tasks in the written part of the course.	1 5	2	3	4
2	In the written part of the course I engaged in experiences that then helped me learn ideas or skills that were new and unfamiliar to me.	1 5	2	3	4
3	My instructor demonstrated skills I was expected to learn in the written part of the course.	1 5	2	3	4
4	My instructor detected and corrected errors I was making when solving problems, doing learning tasks, or completing assignments in the written part of the course.	1 5	2	3	4
5	I had opportunities in the written part of the course to explore how I could personally use what I learned.	1 5	2	3	4
6	I did **not** do very well on most tasks in the written part of the course, according to my instructor's judgment of the quality of my work.	1 5	2	3	4
7	Compared to what I knew before I took this course, I learned a lot about writing with sources.	1 5	2	3	4
8	I am very satisfied with how my instructor taught the written part of the class.	1 5	2	3	4
9	Overall, I would rate the quality of the written part of the course as outstanding.	1 5	2	3	4
10	I solved authentic problems or completed authentic tasks in the written part of the course.	1 5	2	3	4
11	In the written part of the course I was able to remember, describe or apply my past experience so that I could connect it with what I was expected to learn.	1 5	2	3	4
12	Media used in the written part of the course (booklets, illustrations, graphics, audio, video, computers) were helpful in learning.	1 5	2	3	4
13	I had opportunities to practice or try out what I learned in the written part of the course.	1 5	2	3	4
14	I frequently did very good work on projects, assignments, problems and/or activities for the written part of the course.	1 5	2	3	4
15	I learned a lot about writing with sources in this course.	1 5	2	3	4
16	I am dissatisfied with the written part of the course.	1 5	2	3	4
17	Overall, I would recommend this instructor to others.	1 5	2	3	4
18	My instructor provided a learning structure that helped me to mentally organize new knowledge and skills in the written part of the course.	1 5	2	3	4
19	My instructor gave examples and counter-examples of concepts that I was expected to learn in the written part of the course.	1 5	2	3	4
20	My course instructor gave me personal feedback or appropriate coaching on what I was trying to learn in the written part of the course.	1 5	2	3	4
21	I spent a lot of time doing tasks, projects and/or assignments for the written part of the course, and my instructor judged my work of high quality.	1 5	2	3	4
22	I see how I can apply what I learned in the written part of the course to real life situations.	1 5	2	3	4

		1 = strongly disagree; 2 = agree; 3 = undecided; 4 = agree; 5 = strongly agree				
23	I learned very little about writing with sources in this course.	1	2	3	4	5
24	In the written part of the course I solved a variety of authentic problems that were organized from simple to complex.	1	2	3	4	5
25	In the written part of the course I was able to connect my past experience to new ideas and skills I was learning.	1	2	3	4	5
26	I was able to publicly demonstrate to others what I learned in the written part of the course.	1	2	3	4	5
27	I am very satisfied with the written part of the course.	1	2	3	4	5
28	My instructor did **not** demonstrate skills I was expected to learn in the written part of the course.	1	2	3	4	5
29	I did **not** learn much about writing with sources as a result of taking this course.	1	2	3	4	5
30	The written part of the course was a waste of time and money.	1	2	3	4	5
31	In the written part of the course, I was able to reflect on, discuss with others, and defend what I learned.	1	2	3	4	5
32	I put a great deal of effort and time into the written part of the course, and it has paid off—I believe that I have done very well overall.	1	2	3	4	5
33	In the written part of the course I was **not** able to draw upon my past experience nor relate it to new things I was learning.	1	2	3	4	5
34	My instructor provided alternative ways of understanding the same ideas or skills in the written part of the course.	1	2	3	4	5
35	Overall, I would rate this instructor as outstanding.	1	2	3	4	5

Thank you for completing this questionnaire!

Appendix D

Interview Questions for Teachers

<u>Mid-way interviews</u>

How many years of teaching experience do you have?

How many years have you taught this type of student (hbo)?

How many years have you taught writing?

How many years have you taught this specific type of writing: research based on sources?

What is your opinion on the booklets? Too long/short, easy/difficult, useful/useless? Why?

How do you see the students working with these booklets? What comments do they have? About length, difficulty, usefulness, etc.?

What is your view on the effect the booklets have on the students? Motivation, writing skills

What is your opinion on the on/off schedule? Cutting class time in half

How do the feedback sessions work for you? (Calendly or not?)

What suggestions do you have for improvement, on the booklets, on the feedback sessions, on the on/off schedule?

The plan is to implement this in year 1 of the new curriculum. What adjustments do you think are necessary?

<u>Final interviews</u>

How would you rate your experience with the writing course? Why?

How would you rate your class's performance? Why?

What aspects of the writing course were effective? Which were not? (booklets, final paragraph, feedback sessions, final report, on/off schedule)

What changes would you make to the booklets?

What kind of feedback did you give most frequently to students? cognitive vs. corrective (structure, topic sentence, writing style, grammar, spelling, APA style, sources used)

Appendix E

95% CI Tables

Table E1

95% CI for All Classes

Scale	All classes	
	LL	UL
First Principles		
Authentic Problems	2.99	3.30
Activation	3.37	3.61
Demonstration	3.56	3.78
Application	3.91	4.07
Integration	3.11	3.41
Total 5 Principles	3.38	3.60
Other scales		
Academic Learning Time	3.49	3.97
Learning Progress	3.42	3.84
Satisfaction	3.38	3.75
Global Quality	3.66	3.94

Table E2

95% CI for the Two Streams

Scale	Dutch		English	
	LL	UL	LL	UL
First Principles				
Authentic Problems	3.24	3.57	2.67	3.15
Activation	3.56	3.83	3.12	3.49
Demonstration	3.59	3.88	3.43	3.78
Application	3.98	4.24	3.57	4.00
Integration	3.34	3.64	2.81	3.28
Total 5 Principles	3.56	3.78	3.15	3.50
Other scales				
Academic Learning Time	3.54	3.80	3.35	3.72

Learning Progress	4.09	4.43	2.74	3.36
Satisfaction	3.81	4.15	2.90	3.48
Global Quality	3.83	4.15	3.42	3.84

Table E3

95% CI for Dutch Classes

	201N		202N		203N	
Scale	LL	UL	LL	UL	LL	UL
First Principles						
Authentic Problems	3.15	3.65	3.32	3.93	2.83	3.54
Activation	3.42	3.86	3.67	3.96	3.29	4.02
Demonstration	3.35	3.89	3.50	3.89	3.68	4.21
Application	3.89	4.30	3.90	4.32	3.80	4.48
Integration	3.13	3.62	3.22	3.80	3.32	3.96
Total 5 Principles	3.41	3.80	3.55	3.87	3.47	3.99
Other scales						
Academic Learning Time	3.52	3.93	3.41	3.99	3.32	3.82
Learning Progress	3.91	4.54	3.98	4.49	3.96	4.69
Satisfaction	3.67	4.21	3.54	4.18	3.82	4.50
Global Quality	3.65	4.21	3.67	4.29	3.77	4.38

Table E4

95% CI for English Classes

	211E		213E		214E	
Scale	LL	UL	LL	UL	LL	UL
First Principles						
Authentic Problems	3.10	3.75	2.63	3.49	2.04	2.77
Activation	3.62	4.05	2.98	3.63	2.61	3.13
Demonstration	3.66	4.14	3.06	3.71	3.15	3.79
Application	4.03	4.47	3.54	4.16	2.97	3.77
Integration	3.51	4.02	2.62	3.38	2.14	2.80
Total 5 Principles	3.66	3.98	3.05	3.57	2.65	3.19
Other scales						

Academic Learning Time	3.63	4.13	3.14	4.04	2.93	3.53
Learning Progress	3.40	4.23	2.72	3.92	1.90	2.68
Satisfaction	3.91	4.48	2.45	3.46	2.13	2.81
Global Quality	3.66	4.27	2.94	3.86	3.09	3.83