

## The design and implementation of learning paths in a learning management system

<sup>ab</sup>Cindy De Smet, <sup>a</sup>Tammy Schellens, <sup>a</sup>Bram De Wever, <sup>c</sup>Pascale Brandt-Pomares, <sup>a</sup>Martin Valcke

<sup>a</sup>Ghent University, Department of Educational Studies, H. Dunantlaan 2, 9000 Ghent, Belgium

<sup>b</sup>University College Ghent, faculty of Education, Health and Social Work, K.L. Ledeganckstraat 8, 9000 Ghent, Belgium

<sup>c</sup>Aix\*Marseille Université, École Supérieure du Professorat et de l'Éducation Aix-Marseille, Rue Eugène Cas 32, 13248 Marseille cedex 04

This is an Accepted Manuscript of an article published by Taylor & Francis Group  
in Interactive Learning Environments on 15/09/2014, available online:

<http://dx.doi.org/10.1080/10494820.2014.951059>

### Please cite as:

De Smet, C., Schellens, T., De Wever, B., Brandt-Pomares, P., & Valcke, M. (2016). The design and implementation of learning paths in a learning management system. *Interactive Learning Environments*, 24(6), 1076-1096.

### Corresponding author:

Cindy De Smet

E-mail: [cindy@smetty.be](mailto:cindy@smetty.be)

Twitter: @drsmetty

## Abstract

Learning paths have the potential to play an important role in the way educators serve their learners. Empirical research about learning paths is scarce, particularly in a secondary education setting. The present quasi-experimental study took place in the context of a biology course involving 360 secondary school students. A 2 x 2 factorial research design was adopted. Learners were engaged in learning activities in a learning path. These learning activities (1) differed in design and (2) were either undertaken individually or collaboratively. Gender was considered as a critical co-variable given the focus on science learning. All learning paths were developed on the basis of visual representations, but in the experimental design conditions, learners worked with learning paths designed according to Mayer's multimedia guidelines (2003). Multilevel analyses were applied to study the impact on learning outcomes according to the design of learning paths, the individual/collaborative setting, and the co-variable gender. The study provides empirical evidence that both the design and the group setting (collaborative versus individual) have an impact on learning outcomes. Although there was no main effect, several significant interaction effects with gender were found. The results are helpful to direct research about the design and implementation of learning paths in a secondary school setting and underpin the relevance of representation modes in science learning.

**Keywords:** secondary education, learning management system, learning path, collaborative learning, science education, STEM

## Introduction

Earlier research by De Smet, Bourgonjon, De Wever, Schellens, and Valcke (2012) studied the rationale behind the technology acceptance of learning management systems (LMS) by secondary school teachers and also investigated the particular instructional use of LMS within this group of teachers. They found the “informational use of the LMS” or content published by the users (as defined by Hamuy & Galaz, 2010) can be considered a precursor for the “communicational use,” or all processes that foster the exchange of these contents, between LMS users. In other words, a basic usage level (e.g., document publishing or sending announcements) seems to be required before more advanced LMS functionalities can be adopted, such as a wiki (collaborative writing), a forum (moderated discussions) or learning paths (technology-enhanced road map).

De Smet and Schellens (2009) observed that from 376 Flemish secondary school teachers, only 10% actively used the learning path module. This low adoption level suggests that teachers do not know how to design and implement these learning paths. As a result, this study will focus on how learning paths could be appropriately designed and implemented.

Most literature on learning paths can be found within research for technology-enabled learning that studies algorithms for computer-adaptive systems (Capuano et al., 2009; Wong & Looi, 2012). Within this article, a “learning path” refers to the LMS functionality to order a number of learning objects in such a way that they result in a road map for learners. Within a

learning path, learning steps are pre-structured in a general way (as a navigation map or a table of contents) or in a very specific sequenced way (e.g., “complete first step 1 before moving on to step 2”). Learning paths can be created with authoring tools (e.g., eXe, Xerte, Uduku) or programmed by software developers. Central to the design of a learning path are the building blocks: the learning objects. Although the concept of “learning objects” is widely used, its definition is not always clear. According to Wiley (2000), the most cited definition of learning objects comes from the Learning Technology Standards Committee (also known as IEEE, 2005): “any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning.” In his review of definitions of learning objects, Kim (2009) concluded that most definitions include terms such as “learning,” “instructional,” “pedagogical,” or “educational.” In this article, we put forward the definition by Kay and Knaack (2007), who defined learning objects as “interactive web-based tools that support the learning of specific concepts by enhancing, amplifying, and/or guiding the cognitive processes of learners” (p. 6).

Learning objects have the potential to play an important role in the way teachers teach and learners learn. However, empirical research about learning objects is scarce, particularly in secondary education (Kay & Knaack, 2008). Cochrane (2005) found relatively little research reporting design principles for learning objects. Dalziel (2003) argued that e-learning usually has “a well-developed approach to the creation and sequencing of content-based, single learner, self-paced learning objects,” but added “there is little understanding of how to create sequences of learning activities” (p. 593). In addition, he emphasizes there is hardly any research addressing how to support learners with learning objects in a structured, collaborative environment. Given the lack of empirical research focusing on how learning paths should be designed, presented and implemented, and the lack of impact studies on student performance (Kay & Knaack, 2005; Nurmi & Jaakkola, 2006), we concentrated in this study on the impact of learning with learning paths that vary (1) in their design and (2) in the way they are studied, individually or collaboratively. In the next sections, we first present the theoretical basis underpinning design decisions for learning paths and the rationale in relation to collaborative versus individual study of the learning paths. Since our study is set up in the domain of science learning, we also focus on gender, a key variable in science education research.

## **Theoretical and empirical framework**

### *Visual representations*

Learning paths can differ in the way they are visually represented. The value of visual representations in the design of learning paths can theoretically be linked to Cognitive Load Theory (Sweller, 1988, 1994; Sweller, van Merriënboer & Paas, 1998; van Merriënboer & Sweller, 2005) and the Cognitive Theory of Multimedia Learning (Mayer, 2001, 2003, 2005). Cognitive Load Theory (CLT) is an instructional theory that focuses on the human cognitive architecture and its consequences for the design of instruction and learning materials. The

Cognitive Theory of Multimedia Learning (CTML) reiterates CLT's cognitive architecture but looks even more explicitly at design principles for multimedia learning.

### *Cognitive Load Theory*

Cognitive Load Theory is based on the assumption that the processing capacity of working memory (WM) of individual learners is limited, which is in contrast to an unlimited long-term memory (LTM). When new information is not well structured, too abundant, or not well represented, it will invoke extraneous cognitive load (see below) that will hinder the processing of new information, resulting in less successful storage in LTM (Baddeley, 1986). CLT also builds on the assumption that information is organized into schemas within WM, and are subsequently stored and retrieved more easily in/from LTM (Sweller, van Merriënboer & Paas, 1998). A *schema* is a cognitive structure that connects a large amount of information that can be processed as a single unit in working memory and stored in long-term memory. One frequently used example is that of a chess grand master who uses schemas to categorize board pieces and board moves into patterns. Information processing can occur automatically or consciously (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Automatic processing occurs after extensive practice and results in freeing up working memory, while conscious processing occurs in working memory itself and requires memory resources, thus invoking cognitive load. Therefore, a novice chess player who has few such schema available in LTM will need more time to execute a chess move than a professional player. In order to foster learning, schema construction is important, as it leaves working memory open for other tasks and stores information in LTM.

CLT distinguishes three types of cognitive load: *intrinsic*, *extraneous*, and *germane* cognitive load (Sweller et al., 1998). *Intrinsic cognitive load* is dependent on the intrinsic complexity of the information (number of elements and the interrelations between them). *Germane cognitive load* refers to the effort required to construct schemas. *Extraneous cognitive load* is the effort required to process information in view of schema construction. The latter is strongly dependent on the way information is represented.

CLT theory challenges instructional designers to design learning material that results in meaningful learning but does not put too heavy a burden on working memory (Sweller, 1999; van Merriënboer, 1997). Paas, Tuovinen, Tabbers, and Van Gerven (2003) state, "because intrinsic load, extraneous load, and germane load are additive, it is important to realize that the sum of intrinsic, extraneous, and germane cognitive load, should stay within working memory limits" (p.65). Given the fact that intrinsic load is intrinsic to the task, and germane cognitive load is required for schema construction, instructional designers should control extraneous load. Different techniques have been researched to handle extraneous cognitive load, among others, the Cognitive Theory of Multimedia Learning (Mayer, 2001, 2003, 2005).

### *Cognitive Theory of Multimedia Learning*

Instructional designers recognized the need for learning materials that are sensitive to cognitive load (Mayer & Moreno, 2003). A lot of research has been done based on the

Cognitive Theory of Multimedia Learning (CTML), as postulated by Mayer (2001, 2003, 2005). This theory represents a framework to direct instructional design of multimedia materials and results in the definition of practical guidelines to design multimedia learning materials.

CTML is based on three assumptions (Mayer, 2003): the dual channel assumption, the limited capacity assumption, and the active learning assumption. The *dual channel assumption* is derived from the research of Paivio (1978, 1991) and Baddeley (1992, 1995). Central to this assumption is that two separate information processing systems are active to process visual (e.g., text, images) and verbal (audio) representations. The *limited capacity assumption* also builds on the work of Baddeley (1992) and Baddeley, Gathercole and Papagno (1998). It states the amount of processing that can take place within the visual and auditory processing channel is limited. The *active learning assumption* is built on Wittrock's (1989) generative learning theory and implies the learner is actively engaged in processing information and mentally organizes it. Cognitive processes involved include selecting (visual/audio), organizing (mental representation), and integrating (visual, audio, and prior knowledge). In order to study the impact of learning path design, we build in the present study on CTML to differentiate between two learning paths, differing in the degree of elaboration and structure.

### *Collaborative learning*

In this article, the term “collaborative learning” refers to the engagement of all participants in solving a problem together (Roschelle & Teasley, 1995). Akkerman et al. (2007), building on the work of Valsiner and Van der Veer (2000), present both a cognitive and a socio-cultural view when focusing on group cognition. Within the cognitive perspective, the subject of learning is the individual who constructs knowledge about the surrounding world. Following the socio-cultural perspective, the learner is seen as a participant of a social entity where knowledge results from interaction and social activity. Akkerman et al. (2007) add that, within the cognitive view, the social aspect is not denied but rather “understood through its residence in the mind of the individual” (p.42).

Putting learners in a group does not guarantee spontaneous collaboration (Cohen, 1994) or effective learning behavior (Soller, 2001). As a result, instructional support is provided to scaffold or script the collaborative learning process (Kollar, Fischer & Hesse, 2006). Given the focus on learning management systems in the present article, the design of collaborative learning can strongly build on research in the field of Computer Supported Collaborative Learning (CSCL). Kollar, Fischer & Hesse (2006) put forward five minimum characteristics of scripting in a CSCL setting: scripts must 1) contain an objective, 2) engage learning activities, 3) sequence all required actions, 4) specify and distribute roles, and 5) contain a type of representation in which instructions are presented to the learners. In this research, we used teacher scenarios (see below) that were based on scripts.

Adopting collaborative learning in the context of learning paths, can – from a theoretical perspective – also be linked to cognitive load theory. Kirschner, Paas, and Kirschner (2009a) found that groups can be considered information-processing systems containing multiple

working memories, and as such, create a collective working space where cognitive load can be divided among the learners. In this view, groups are favored against individuals who can only rely on their individual working memory. Furthermore, when the group work is well structured (e.g., building on strongly elaborated and structured learning objects in the learning path), it reduces extraneous cognitive load and helps learners maximize cognitive processes that result in schema construction (Sweller, Van Merriënboer, & Paas, 1998), and thus, higher learning outcomes.

### *Science education and gender*

The present study takes place within the setting of STEM education (science, technology, engineering, and mathematics). Although STEM education leads to good jobs and a higher standard of living, today's youth seem to have little interest in science as a possible career path (European Commission, 2004, 2006; Organisation for Economic Co-operation and Development [OECD], 2007, 2008; U.S. Department of Education, 2007; National Governors Association, 2007). In addition, there is a clear gender gap in the STEM field. Several studies (European Commission, 2004, 2012) reveal that females are underrepresented in science careers. This comes in sharp contrast to the observation that girls are more successful at school, as they obtain higher grades and are less likely than boys to repeat a year (European Commission, 2006). In a recent publication, the European Commission (2012) presented the following reasons for this gender gap: stereotypes found in children's books and school manuals, gendered attitudes of teachers, gendered advice and guidance on courses to be followed, and different parental expectations regarding the future of girls and boys.

Research about gender differences does not always present a consistent picture. PISA 2012 (OECD, 2013) showed different levels of performance in science, reading, and mathematics between males and females, although differences were significantly larger within, rather than between, genders. Nevertheless, significant gender differences were observed for reading (in favor of girls) and mathematics (in favor of boys). They also found that for mathematics, girls are under-represented among the highest achievers in most countries and economies, and males have higher perceptions about their science abilities as compared to females. This is in line with research from Eccles (1994) and Lubinski and Benbow (2006), which stated that women are less likely to enter occupations linked to mathematics and physical sciences because they have less confidence in their abilities and place less subjective values on these fields compared to other occupations. Furthermore, Eccles (1994) argued that girls rate social values high and prefer to study academic subjects that have social implications, which, in the long term, enable them to do something worthwhile for society.

### *Learning outcomes based on gender*

We believe the main conditions under study (i.e., design decisions and the group setting) influence learning outcomes based on gender. When studying design conditions, we refer to Super and Bachrach (1957), as well as more recent follow-up research by Wai, Lubinski, and Camilla (2009), which focused on the critical role of spatial ability within STEM-education. The construct *spatial ability* was defined by Lohman (1994) as “the ability to generate, retain,



retrieve, and transform well-structured visual images” (p. 1000). Mayer and Sims (1994) found evidence that high-spatial learners had to dedicate fewer cognitive resources to build a representational connection between visual and verbal material, thus leaving more room for other processes. From their longitudinal findings, Wai, Lubinski and Benbow (2009) concluded that high levels of spatial visualization have a robust and highly relevant influence in approaching STEM domains. Ceci and Williams (2010) added that males excel in spatial ability and underline the fact that in large meta-analyses, the effect size for spatial ability is substantial: .50 to .75 for male superiority. As the second version of our learning path is optimized with Mayer’s guidelines (2003), leading to a better elaborated and structured course, we can postulate that this optimized version will offer better spatial visualization.

When researching group setting, we can build on group diversity literature. Harrison and Klein (2007) describe *group or unit diversity* as “the distribution of differences among the members of a unit with respect to a common attribute X” (p. 1200). They differentiate diversity as: separation (differences in opinion among members), variety (differences in knowledge and/or experience) and disparity (differences in status and/or power), and concluded that only variety has a positive impact on group effectiveness. As a result, gender diversity can be conceptualized as gender separation, gender variety, or gender disparity. Extending the work of Harrison and Klein (2007), Curşeu, Schruijer and Boroş (2007) and Curşeu and Sari (2013) found gender variety indeed has a positive outcome on group cognitive complexity, and mixed-gender groups achieve better results. Moreover, Curşeu and Sari (2013) stress that “the core argument in this line of research is that gender variety increases the pool of cognitive resources of groups because men and women have qualitatively different life experiences, therefore likely to have different task-related knowledge structures (Curşeu, Schruijer, & Boros, 2007; Rogelberg & Rumery, 1996)” (p. 1).

Slotta and Linn (2009) found web-based collaborative inquiry seems to be helpful in developing and maintaining positive attitudes towards science and science instruction. In a recent study, Raes, Schellens & De Wever (in press) found that low achievers, and more specifically, low-achieving girls, benefited from this type of intervention, especially with respect to the ability to participate in small group discussions.

On the basis of the group diversity literature and the positive impact that web-based collaborative inquiry has on girls, we expect that girls will benefit from working collaboratively.

## **Research design**

### *Research question and research hypotheses*

The main research question directing this study is whether additional investment in the design and implementation of learning paths will have a beneficial impact on learning outcomes. Gender is considered as a critical co-variable given the focus on science learning.

Building on the theoretical framework of CTL and CTML, we put forward the first hypothesis (H1): Students studying a learning path, optimized with Mayer's (2003) guidelines in mind, will attain significantly higher learning outcomes as compared to students studying a basic learning path with multimedia learning objects.

Building on the CSCL framework, we put forward the second hypothesis (H2): Learners studying the learning path collaboratively will attain significantly higher learning outcomes as compared to students studying the learning path individually.

Considering the empirical data in relation to gender and STEM, we put forward a third, twofold hypothesis. Given the critical role of spatial ability, we expect (H3a) a significant interaction effect with respect to gender, in favor of males, when studying the learning path optimized with Mayer's guidelines (2003). In view of the group diversity literature and the positive impact web-based collaborative inquiry has on girls, we expect (H3b) a significant interaction effect with respect to gender, in favor of females, when studying the learning path collaboratively.

### *Participants*

Secondary education in Flanders comprises six consecutive years of study, starting at the age of 12. We selected six secondary education schools in collaboration with a GO! staff member. GO! is one of the three dominant governing bodies that sets up schools in Flanders, the Dutch speaking area of Belgium. GO! schools comprise 15,27% of secondary school education in Flanders. Governing bodies have considerable autonomy to, among other things, develop school curriculum, recruit staff, choice of teaching methods, etc. As a consequence, the curriculum in the selected schools and classes is largely comparable. All participating schools are situated in urban areas, as Belgium, and Flanders in particular, is one of the world's most urbanized countries (United Nations World Populations Prospects, 2011).

All biology teachers (N = 8; 3 males, 5 females) teaching in the third grade of each of the six schools were willing to participate in the study. Twenty-nine different classes were selected at random to participate in the study. All students enrolled in these 9<sup>th</sup> grade classes (N= 360; 167 males and 193 females) participated in all consecutive activities during the study. Students were, on average, 15 years old (89,4%). Figure 1 shows the participant flow chart.

Prior to the study, informed consent to use the data for research purposes was obtained through the different school teachers.



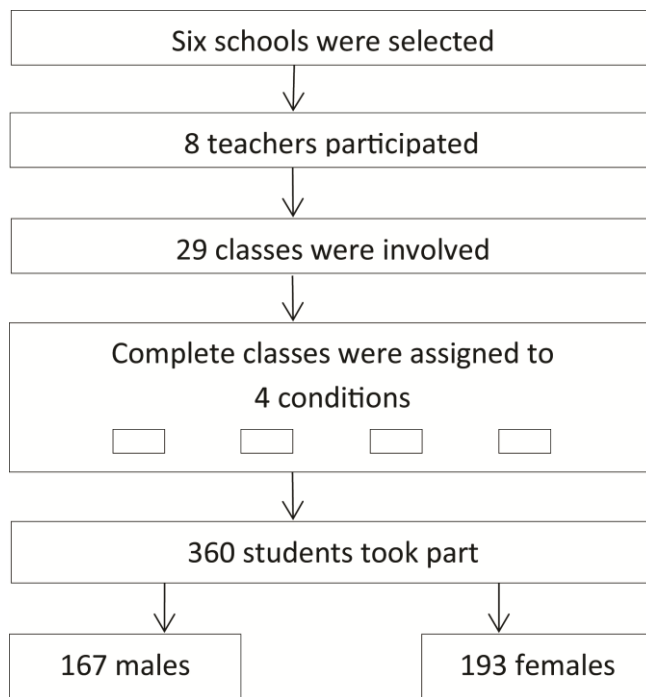


Figure 1. Participant flow chart.

### *The biology learning materials: Two versions of the “Bacteria” learning path*

In the present study, learning paths were developed using “eXe learning,” an open-source authoring tool. Resources authored in eXe can be exported as a website or imported in any SCORM (Sharable Content Object Reference Model) compliant Learning Management System. This gives teachers the opportunity to open learning paths via a browser (online or offline) or to integrate these learning paths within their school LMS.

From the biology curriculum, the topic “bacteria collection and growth” was selected in view of developing new learning materials. Two recently graduated biology teachers created learning materials following the official GO! biology curriculum. Next, these materials were reviewed and modified by 18 pre-service teachers majoring in biology under the supervision of their lecturer.

A first version of a learning path was elaborated, consisting of multimedia learning objects that build on text, schemes, pictures, and web-based exercises (see Figure 2). A second version of the same learning path was developed by applying Mayer’s multimedia guidelines (2003). Based on the handbook by Clark and Mayer (2007), learning objects in the second version of the learning path were optimized by applying, for example, the multimedia principle (adoption of both audio and graphs), the contiguity principle (alignment of the text and the corresponding graphics), the redundancy principle (explanations next to visuals were either with audio or text, not both), and the coherence principle (no extra interesting materials were added). The active learning assumption (Wittrock, 1989, Mayer, 2003) stresses the learning material should have a coherent structure and provide guidance to the learner on how to build knowledge structures. As a result, advanced organizers were included in the optimized learning path in order to help organize unfamiliar content (Ausubel, 1960, 1968).

For reading purposes, we will refer to the first version of the learning path as the “TSPW learning path” (Text, Schemes, Pictures and Web-based exercises) and to the second version as the “MGL learning path” (Mayer GuideLines).

Rijk	moneren	protisten	fungi (zwammen)	planten	dieren
Celltype					
complexiteit	eencellig	(meestal) eencellig	meercellig	meercellig	meercellig
bewegbaarheid	soms	soms	nee	nee	ja
voedingswijze	autotroof of heterotroof	meestal heterotroof	heterotroof	autotroof	heterotroof
bladgroenkorrels	soms	nee	nee	ja	nee

**levende wezens**

In de 17e eeuw werd door Antonie van Leeuwenhoek een eerste microscoop gebouwd. Daarmee zag hij kleine (bewegende) organismen in regenwater en in landplak, die organismen noemen we nu bacteriën.

**Bacteriën zijn net zoals ons levende wezens.**

In de eerste graad leerde je al wat levende wezens zijn en hoe ze ingedeeld worden in 5 rijken.

**Activiteit**

Ken je nog de kenmerken van levende wezens? Weet je nog hoe de levende wezens ingedeeld worden?

**Maak de oefeningen**

- wat zijn bacteriën 1
- wat zijn bacteriën 2
- wat zijn bacteriën 3

in de oefeningen module.



Figure 2. The uppermost image depicts an advanced organizer (Ausubel) on bacteria classification that was offered to all students following a MGL learning path before navigating to the rehearsal bacteria classification exercises (the image at the bottom). Students following a TSPW learning path were only exposed to the rehearsal bacteria classification exercises. No other information on the subject was given to these learners.

### *Individual versus collaborative study of the learning paths*

Along with a better multimedia elaboration of the learning path, we also studied the impact of the group setting. As defined by Kollar, Fischer and Hesse (2006), and as applied within this research, scripts contain several components, including a learning objective and a type of representation, in which instructions are presented to the learners. Scripts also engage learning activities and sequence all required actions.

We chose to implement scripts into teacher scenarios for two reasons. First, Flemish teachers are used to working with these scenarios on a daily basis. Pre-service teachers and in-service teachers use lesson preparation scenarios as part of their (sometimes obligatory) daily work routine. We used existing lesson preparation templates to create our teacher scenarios. Second, we wanted to guarantee the comparable nature of the activities under all research conditions. The collaboration scenarios did not result in differences in the content of what was studied about bacteria; they differed in the way students organized, shared, and carried out their work to guarantee that students – in whatever research condition – received the same learning opportunities and to monitor the way students followed the particular learning path.

### *Research instruments: Learning performance*

Students were offered knowledge tests at three consecutive moments: a pre-test, a post-test (immediately after completion of the learning path), and a retention test (one month after completion of the learning path). Each test consisted of 20 multiple choice and true/false questions. The study took, on average, between seven and nine weeks to be completed. However, since teachers were not able to refrain from monthly evaluation between the post-test and the retention test, we decided to focus on pre-/post-test differences in our study. Retention test scores are mentioned in Table 4; however, readers should keep in mind that these could be influenced by intermediate tests not taken into account in the present study.

All test items were created by two recently graduated biology teachers based on the official GO! biology curriculum. Six biology teachers tested all items within their classes. Based on the analysis of these tests and the teachers' item evaluation, some items were discarded and the remaining items were divided into three balanced tests (one test for each moment). Figure 3 shows how knowledge tests were created.

Item analysis was conducted to improve the quality and accuracy of the true/false items. A combination of item difficulty (p-value) and item discrimination (PBS or Point-Biserial correlation) was taken into account. Items with P-values above 0.90 and PBS-values near or less than zero were removed from the tests (Division of Instructional Innovation and Assessment, University of Texas at Austin, 2007). As a result, eight items were omitted from each test.

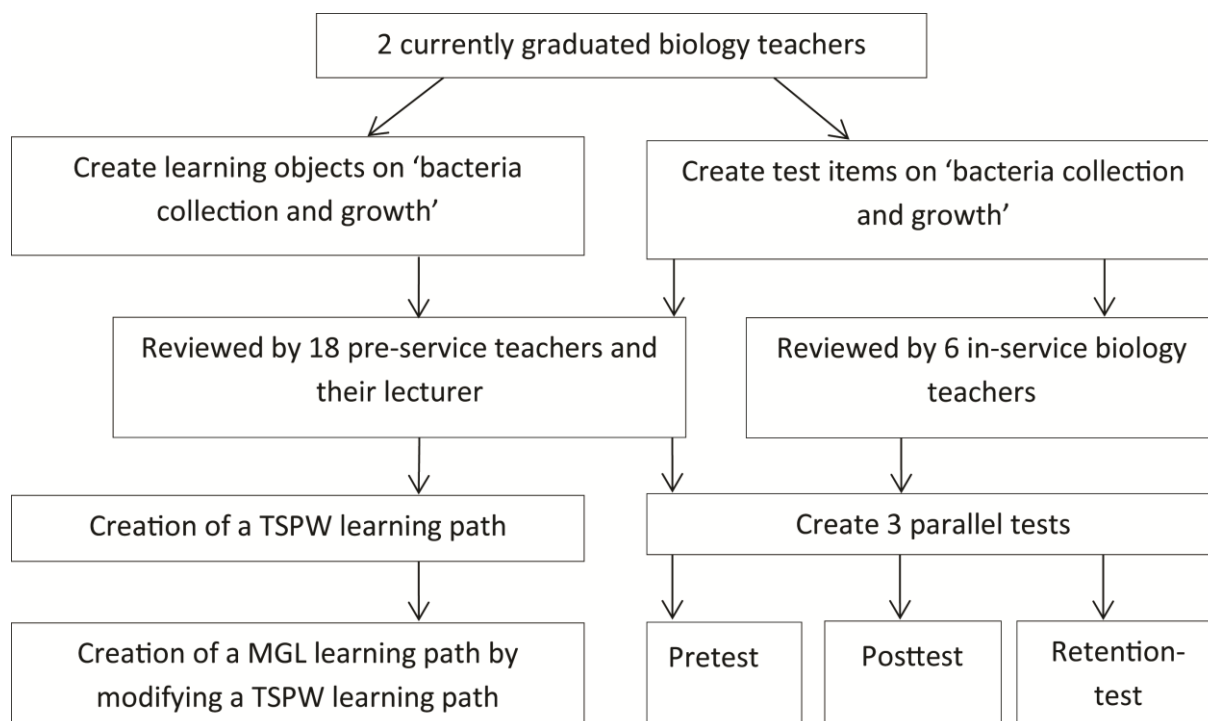


Figure 3. Creation process of the learning paths and the knowledge tests.

## Procedure

The researcher visited all teachers and gave a one-hour introduction. We briefed teachers on all the aspects of the research process. Other topics discussed included, amongst others, the proposed time schedule and technical information concerning learning paths within the Learning Management System. Complete classes (N = 29) were assigned to the four different conditions (see Table 1). It was mandatory that all lessons took place in computer classes.

As can be observed in Table 1, we did not reach a balanced number of students across all conditions. Two teachers assigned to the collaborative condition of the MGL learning path had to cancel their participation. Given the last-minute character of these events and the unfortunate timing in the middle of a semester, we were not able to recruit new teachers nor to redistribute the teachers over conditions.

Depending on the condition they were assigned to, all teachers received a digital (USB-stick) and/or a paper version of the following material: a research guideline, a comprehensive teacher scenario, the proposed time schedule, and two versions of the learning path (HTML and SCORM). At the same time, we provided a box containing paper versions of all the knowledge tests. We also sent teachers an e-mail address and telephone number by which they could contact three researchers. Only a few minor technical questions emerged.

Table 1. Number of participants across conditions.

	IndTSPW	ColTSPW	IndMGL	ColMGL
<u>Males</u>	59	63	37	8
<u>Females</u>	54	71	50	18
<u>Total</u>	113	134	87	26

*Note:* Ind = individual, Col = collaborative, TSPW = Text, Schemes, Pictures and Web-based exercises learning path, and MGL = Mayer GuideLines learning path.

## Statistical analysis

Our data have a clearly hierarchical structure (i.e., students in classes from different schools were offered knowledge tests at three consecutive moments). This leads to the conclusion that individual observations are not completely independent given the selection processes, common history, and experiences students share (Hox, 1994). Knowledge scores from students in the same classes might be dependent, and thus break the assumptions of a simple regression analysis. By doing so, we would ignore school-level and class-level inferences and

focus only on individual learning outcomes. In this respect, Multilevel Modeling is suggested as an alternative and adequate statistical approach (Diez-Roux, 2000, Nezlek, 2008), and most certainly in the case of repeated measures (Goldstein, 2003). Within multilevel analysis, the hierarchical nesting, dependency, unit of analysis, standard errors, confidence intervals, and significance tests are handled correctly (Goldstein, 1995) and, in general, even more conservative than a traditional regression analysis where the presence of clustering is ignored (Goldstein, 2003).

Following Van Der Leeden (1998), we consider repeated measures as a hierarchical structure where measurements are nested within individuals. Consequently, our knowledge tests are defined as the first level, students as the second level, classes as the third level, and schools as the fourth level. We used MLwiN software (Centre for Multilevel Modelling, University of Bristol) to analyze the hierarchical data (Nezlek, 2008, Rasbash, Steele, Browne, & Goldstein, 2009).

We followed a two-step procedure to analyze the effects of three independent variables (design decisions, group setting and gender) on the dependent variables (learning outcomes). The models built following this procedure are presented in Table 4 (in annex). First, we created a four-level conceptual null model (Table 4, Model 0) to serve as a baseline model. This unconditional null model (without any predictor variables) provides the overall pre-test, post-test, and retention scores across all students, classes, and schools. The second step concerned the input of the three main explanatory variables (visual representation, group setting, and gender) in the fixed part of the model and allowed cross-level interactions between student, class, and school characteristics. This resulted in Model 1 (Table 4).

## Results

### *Model building*

The models built following the two-step procedure are presented in Table 4.

Given our repeated measures approach, the conceptual unconditional null model (Table 4, Model 0) predicts the overall pre-test ( $M =$  the intercept, or 57.18 out of 100), post-test ( $M = 64.49 = 57.18 + 7.31$ ), and retention test scores ( $M = 71.93 = 57.18 + 14.75$ ) across all students, classes, and schools. Thus, in general, without taking into account visual representation, collaboration mode, and gender but controlling for the nested data structure, students score significantly higher on the post- and retention test as compared to the pre-test.

This null model also results in four variance estimates, as shown in the random part of the model: one for school level, one for class level, one for student level, and one for the measurement occasion. The variance in scores within this null model on the four levels are, except for the school level, significantly different from zero and significant at the  $p < .001$  level. As a result, we can state that 1.15 % of the total knowledge score variance lies at school level, 9.42% at class level, 14.26 % at student level, and finally, 75.17% at the measurement occasion.

Subsequently, based on the theoretical framework, visual representation, group setting, and gender were entered into the model as potential explanatory variables. All predictors were included in the models as fixed effects. Adding these variables to the null model resulted in a better model fit ( $X^2 = 55.59$ ,  $df = 21$ ,  $p < .001$ ). Model 1 (Table 4) shows the results of this factorial model with main and interaction effects added to the model. The reference category is a male working individually and following a TSPW learning path. In the random part of Model 1, all variance in scores are significantly different from zero and significant at the  $p < .001$  level, except for school level.

### *Student scores*

Table 2. Knowledge scores on pre- and post-test.

	Pre	Post
Male, Individ., TSPW	59.90	57.25
Male, Individ., MGL	61.29	76.52
Male, Collabor., TSPW	58.30	63.23
Male, Collabor., MGL	57.30	66.00
Female, Individ., TSPW	55.06	63.85
Female, Individ., MGL	55.92	72.22
Female, Collabor., TSPW	58.41	64.51
Female, Collabor., MGL	46.57	54.16

*Note:* Individ = individual; Collabor = collaborative; TSPW = Text, Schemes, Pictures and Web-based exercises learning path; and MGL = Mayer GuideLines learning path.

Figure 4 shows the drilled-down details of student scores, while Table 2 displays the knowledge scores on the pre- and the post-test. First, we notice students' scores are close together (between 55.05 and 61.29) at the pre-test measurement, except for females working collaboratively on a MGL learning path (46.57). However, this score on the pre-test was not significant ( $p = 0.23$ ). Second, we observe that the two steepest slopes (i.e., students who learned the most from the intervention) are the females and males within the individual MGL learning path condition. These students received the highest post-test scores: 76.52 for males (significant on the post-test and on the increase between the pre-test and post-test,  $p < .001$ ) and 72.22 for females (non-significant on the post-test and on the increase between the pre-test and post-test,  $p = .07$  and  $p = .29$ ). On the other hand, the lowest scores on the post-test can be found for males working individually on a TSPW learning path (non-significant on the post-test, i.e.,  $p = .30$ , but a significant decrease between the pre-test and post-test,  $p < .001$ ) and for females working collaboratively on a MGL learning path (non-significant on the post-test and on the increase between the pre-test and post-test,  $p = .46$  and  $p = .30$ ).

The remaining four scores are closely bundled together (between 63.22 and 66.00). Out of those four, only the condition where males worked collaboratively on a MGL learning path indicates a non-significant post-test score ( $p = .11$ ). When calculating the increase between



the pre-test and post-test scores within those four conditions, only the increase for females working individually on a TSPW learning path was found to be significant ( $p = .007$ ).

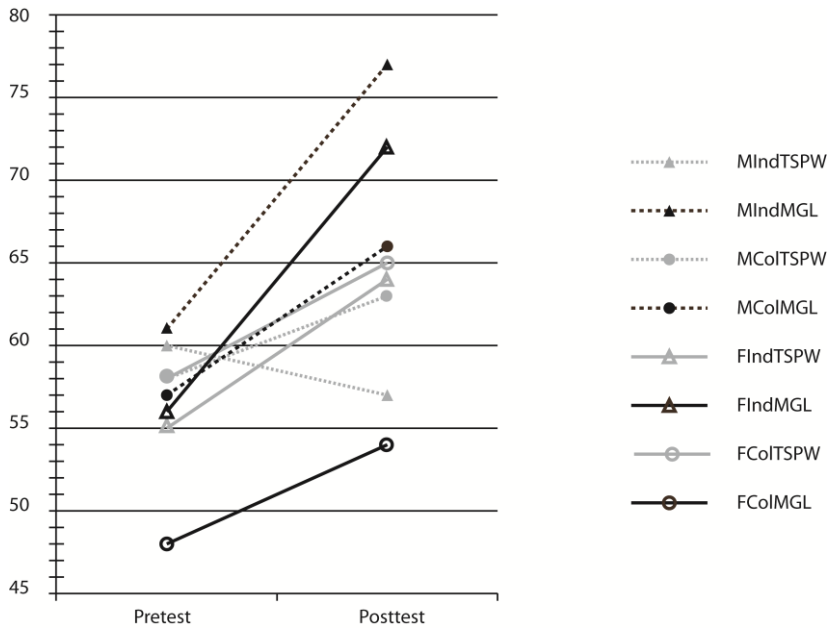


Figure 4. Knowledge scores in the pre-test and post-test for males and females.

*Note:* M = male; F = female; Ind = individual; Col = collaborative; TSPW = Text, Schemes, Pictures and Web-based exercises learning path; and MGL = Mayer GuideLines learning path.

### *Hypothesis testing*

Table 3. Differences between scores on the post-test and an overview of the significant increase between the pre- and post-test.

Relevant results found	<i>p</i>	Hypothesis
Male, Individ., TSPW < Male, Individ., MGL	.001	supports 1
		supports 3a
Male, Individ., MGL > Male, Collabor., TSPW	.045	supports 3a
Male, Individ., MGL > Male, Collabor., MGL	.019	contradicts 3a
Male, Collabor., TSPW < Male, Collabor., MGL	.130 <sup>a</sup>	contradicts 1
		contradicts 2
Female, Individ., TSPW < Female, Individ., MGL	.041	supports 1
Female, Collabor., TSPW > Female, Collabor., MGL	.420 <sup>a</sup>	contradicts 1
		contradicts 2
Female, Collabor., TSPW > Female, Individ., TSPW	.030	supports 3b
Female, Collabor., MGL < Female, Individ., MGL	.460 <sup>a</sup>	contradicts 3b
Overview of the significant increase between pre- and posttest		
	<i>p</i>	Hypothesis
Male, Individ., TSPW	.001	contradicts 2
Male, Individ., MGL	.001	supports 3a
Female, Individ., TSPW	.007	contradicts 2

*Note:* Individ = individual; Collabor = collaborative; TSPW = Text, Schemes, Pictures and Web-based exercises learning path; and MGL = Mayer GuideLines learning path.

<sup>a</sup> Given the non-significant difference, the direction of this difference was not established.

Given our first hypothesis (H1), we expected students following a MGL learning path to outperform students studying a TSPW learning path in their knowledge scores. As illustrated in Figure 4, the three highest knowledge scores on the post-test are attained by males and females following a MGL learning path within an individual setting (MIndMGL and FIndMGL), and by males in a collaborative setting (MColMGL). These findings suggest that optimizing a learning path with Mayer's Guidelines (2003) leads to better knowledge scores. However, when calculating the differences between the knowledge scores on the post-test (Table 3), this observation is only confirmed for students within the individual setting. MIndTSPW was significantly lower than MIndMGL ( $p < 0.001$ ) and FIndTSPW was significantly lower than FIndMGL ( $p = 0.04$ ). However, MColTSPW was not significantly lower than MColMGL ( $p = 0.13$ ) and FColTSPW was lower than FColMGL ( $p = 0.42$ ). Therefore, Hypothesis 1 can only be accepted for both males and females following the MGL learning path in an individual setting.

We also hypothesized students who collaborate in tackling the learning task would outperform students within an individual setting (H2). As depicted in Figure 4 and given the conclusion of Hypothesis 1, this was not the case. However, when observing the collaborative conditions (Figure 4), we notice a difference between males and females. On the one hand, males attain almost the same score under the collaborative condition, regardless of the version of learning path studied. This is confirmed when calculating the differences between the knowledge scores on the post-test (Table 3), as MColTSPW is not significantly lower than MColMGL ( $p = 0.13$ ). On the other hand, females have a higher score (Figure 4) on the post-test within the collaborative condition when they study with a TSPW learning path (as

compared to a MGL learning path). However, when calculating the differences between the conditions on the post-test, FColTSPW is not significantly higher than FColMGL ( $p = 0.42$ ). In addition, the superiority of studying in the individual setting for males and females – already concluded in relation to Hypothesis 1 for the MGL learning pathing – is now also confirmed for the TSPW learning path when calculating the increase in pre-test to post-test scores (Table 3) for MIndTSPW ( $p < 0.001$ ) and FIndTSPW ( $p = 0.007$ ). In view of these findings, we therefore have to reject Hypothesis 2.

Following our third hypothesis, we expected to observe a significant interaction effect of gender when studying the two versions of the learning paths in combination with group setting. For males, we hypothesized (H3a) that, given the critical role of spatial ability, males would benefit most from the – with Mayer’s guidelines (2003) – optimized learning path (MGL learning path). When testing Hypothesis 1, we found that males following a MGL learning path in an individual setting achieved better results than males following a TSPW learning path individually. We did not find similar results for males in the collaborative setting. Moreover, the superiority of the MIndMGL condition above MIndTSPW, MColTSPW and MColMGL is very obvious. When calculating (Table 3) the differences between knowledge scores on the post-test, MIndMGL was significantly higher than MIndTSPW ( $p < 0.001$ ), than MColTSPW ( $p = 0.045$ ) and MColMGL ( $p = 0.019$ ). When calculating the difference between the pre-test and the post-test for MIndMGL (Table 3), the increase in scores was significant ( $p = 0.001$ ). As a result, hypothesis H3a can only be partly accepted for males following the MGL learning path in the individual setting.

For females, we hypothesized (H3b), in view of the group diversity literature and the positive impact web-based collaborative inquiry has on girls, that learning outcomes would be significantly higher when girls work collaboratively. As seen in Figure 4 and Table 2, females following a TSPW learning path achieve slightly better scores on the post-test within the collaborative condition as compared to the individual setting. When calculating the difference between these conditions on the post-test, FColTSPW was significantly higher than FIndTSPW ( $p = 0.03$ ). However, females following a MGL learning path collaboratively achieved lower scores on the post-test as compared to girls under the individual MGL condition. When calculating the difference between these conditions on the post-test, FColMGL was not significantly lower than FIndMGL ( $p = 0.46$ ). Given the rather small difference between FColTSPW and FIndTSPW on the one hand and the problems arising from the unbalanced number of students in the FColMGL condition, we conclude there is no conclusive evidence to accept hypothesis H3b.

## Discussion

In this research, we focused on the impact of the way a learning path is designed, an individual versus a collaborative setting, and gender differences between boys’ and girls’ learning outcomes in the context of a STEM secondary education setting.

Our findings showing the superiority of an (with Mayer's guidelines, 2003) optimized learning path are in line with previous research on the critical role of spatial ability within STEM-education (Super & Bachrach, 1957; Wai et al. 2009; Mayer & Sims, 1994). A MGL learning path leads to a better elaborated and structured course, and thus, offers a better spatial visualization than a TSPW learning path. These findings help explain the superiority of a MGL learning path within this research, and more specifically, when students (both males and females) are working alone.

These results are important for different stakeholders. We present both practical and theoretical implications. In the first place, our results are important for teachers when they are designing learning paths to be implemented in an online learning environment. In addition, the results are also important for instructional designers creating learning materials to be used, for example, in addition to school manuals.

The importance of visual representation has theoretical implications within STEM-education. More specifically, the critical role of spatial ability (Mayer & Sims, 1994; Wai et al., 2009; Ceci & Williams, 2010) was reaffirmed. Empirical evidence from longitudinal research shows that spatial ability is an important psychological characteristic among adolescents in general, but particularly beneficial for those who go on to develop high levels of STEM-expertise in their future careers (Wai, Lubinski & Benbow, 2009). Lubinski (2004) even advocates the potential usefulness of spatial ability to identify women with genuine talent for and interest in math/science careers. Moreover, he stresses that, on the basis of individual differences in spatial ability, not only student selection, but also instruction and curriculum design should be taken into account.

Besides the strong impact of the way learning materials are visually represented, the impact of collaborative learning was less obvious. More specifically for females, the results demonstrate that collaboration does not automatically lead to better learning (Soller, 2001).

In their meta-analysis on the application of technology in support of collaborative learning, Resta and Laferrière (2007) refer to evidence that was found in favor of collaborative learning when groups are heterogeneous, including gender (see also Johnson & Johnson, 1996), but also to the tendency of women to be less active in learning groups (see also Felder, Felder, Mauney, Hamrin, & Dietz, 1995). Curşeu, Schrujjer, and Boroş (2007) and Curşeu and Sari (2013) postulated that gender variety has a positive outcome on group cognitive complexity and that mixed-gender groups achieve better results. However, as gender diversity can also be differentiated as gender separation and gender disparity, negative influences on group effectiveness may have taken place. Thus, several negative influences on collaborative learning may have played a role within in our instant research, where membership in a group was formed randomly. Despite our extensive teacher scenarios and comprehensive briefing of the teachers, these factors can explain, in combination with the strong impact of the way learning paths are visually presented, why the collaborative conditions underperform within this research. Given the importance of the female presence within STEM, further research should try to overcome these negative influences on collaborative learning.

Finally, our research reveals the same contradictory findings concerning gender differences as stated in the meta-analysis of Voyer and Voyer (2014): females seem to score differently than expected (or even underperform) on achievement tests, while research shows persistently that females outperform males in actual school performance (i.e., school marks) regardless of the material studied. According to the authors, a possible explanation can be found in the way research is generally designed, and more specifically, the fact that the particular achievement tests used in the studies are not based on teacher marks. The authors also refer to Lindberg et al. (2010), who reported that male advantages on achievement tests increase with age, with a peak in high school, but decline for college and adult learners. This helps explain inconsistent findings in gender scores.

All findings discussed above lead to the conclusion that, although we tried to fill in the gap in research about the design and implementation of learning paths with respect to gender within the STEM field, several areas need to be improved and should be further researched.

### *Limitations*

This quasi-experimental study took place in computer classes, involving 360 secondary school students. The fact that the study was performed in a regular school setting is advantageous for the ecological validity; however, there are some limitations.

Despite of all the advantages an authentic context has to offer, it also leads to uncontrolled and unexpected incidents. For instance, we asked teachers to refrain from any form of evaluation between the pre-test and the retention test, but due to a monthly evaluation system within the participating schools, teachers could not keep to this condition between the post-test and the retention test. As a result, we had to limit our focus to the pre- and post-test differences.

Another limitation was the unbalanced number of students across conditions, more specifically, within the collaborative condition of the MGL learning path (see Table 1). Due to a long-term illness, one teacher cancelled her participation; another teacher was fired. Given the last-minute character of these events, we were not able to recruit new teachers or to redistribute the teachers over conditions.

Third, within our research, complete randomization of students to conditions was not possible. As a result, complete classes were assigned to conditions. In this situation, multilevel analysis is the only appropriate statistical method, as ignoring group level (measurement occasions within students within classes within schools) would lead to overlooking the importance of group effects, and thus, violate the independence assumption (Nezlek, 2008). However, we would also like to note that the random assignment of individuals to particular conditions is sometimes impossible, impractical, or even unethical (Weathington, Cunningham & Pittenger, 2010).

Last, our results on collaborative learning indicate that follow-up research could benefit from more detailed information (Resta & Laferrière, 2007) on group composition of students (e.g., number of students within each group, same-sex vs. mixed-sex groups). Other aspects of

collaboration that need to be more closely studied are the degree of experience of our stakeholders and the interaction between the teacher and the students.

## *Conclusions*

Within this large-scale research, empirical evidence supported the importance of the actual design of a learning path and the impact of a collaborative versus individual learning setting on learning outcomes.

The importance of this study consists of, amongst others, (1) the implementation of learning paths, (2) in an LMS environment, (3) within the context of a STEM course, (4) involving 360 secondary school students and their teachers. This type of research is not only scarce (Kay & Knaack, 2008; De Smet & Schellens, 2009), but above all, important in a digitalizing world where the need for STEM education can be heard loud and clear within all levels of society.

Given the latest trends in online education and the focus on personalized learning and adaptive instruction; the initiatives undertaken in these fields by private grant-making foundations like the Bill and Melinda Gates Foundation to fill the education gap (e.g., their Adaptive Learning Market Acceleration Program, 2014), the rise of sophisticated adaptive learning software and platforms like Knewton (Time, 2013), and the partnerships between learning content publishers (e.g., Pearson, Sanoma Learning Solution) and software companies (e.g., Microsoft), we believe our research on learning paths can be an asset to help shape the future of learning and education.

## **Acknowledgments**

The Research Fund of University College Ghent financially supported this research.



## References

- Akkerman, S., Van den Bossche, P., Admiraal, W., Gijssels, W., Segers, M., Simons, R. J., & Kirschner, P. (2007). Reconsidering group cognition: From conceptual confusion to a boundary area between cognitive and socio-cultural perspectives? *Educational Research Review*, 2(1), 39–63.
- Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 51, 267–272.
- Ausubel, D. B. (1968). *Educational psychology: A cognitive view*. New York, NY: Holt, Rinehart & Winston.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Baddeley, A. D. (1992). Working memory. *Science*, 255, 556–559.
- Baddeley, A. D. (1995). *Working memory (Oxford Psychology Series)*. Oxford: Clarendon Press.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review*, 105, 158.
- Bill & Melinda Gates Foundation, Adaptive Learning Market Acceleration Program 2014, retrieved from <http://www.gatesfoundation.org/~media/GFO/Documents/How%20We%20Work/Adaptive%20Learning%20RFP%20Webinar%20Presentation.pdf>
- Capuano, N., Gaeta, M., Marengo, A., Miranda, S., Orciuoli, F., & Ritrovato, P. (2009). LIA: An intelligent advisor for e-learning. *Interactive Learning Environments*, 17, 221–239.
- Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions in Psychological Science*, 19, 275–279.
- Clark, R. C., & Mayer, R. E. (2007). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning (2 ed.)*. San Francisco: Wiley.
- Cochrane, T. (2005). Interactive QuickTime: Developing and evaluating multimedia learning objects to enhance both face-to-face and distance e-learning environments. *Interdisciplinary Journal of Knowledge and Learning Objects*, 1, 33–54. Retrieved from <http://www.ijello.org/Volume1/v1p033-054Cochrane.pdf>
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64, 1–35.
- Curşeu, P. L., & Sari, K. (2013). The effects of gender variety and power disparity on group cognitive complexity in collaborative learning groups. *Interactive Learning Environments*, (ahead-of-print), 1-12.

- Curşeu, P. L., Schruijer, S., & Boroş, S. (2007). The effects of groups' variety and disparity on groups' cognitive complexity. *Group Dynamics: Theory, Research, and Practice*, 11(3), 187.
- Dalziel, J. (2003). Implementing learning design: The Learning Activity Management System (LAMS). *Proceedings of the ASCILITE 2003 conference*, Adelaide, Australia, 593–596.
- De Smet, C., & Schellens, T. (2009). ELO's in het Vlaams secundair onderwijs: Nieuw of alweer achterhaald. *Advies & Educatie*, 26(5), 12–14.
- De Smet, C., Bourgonjon, J., De Wever, B., Schellens, T., & Valcke, M. (2012). Researching instructional use and the technology acceptance of learning management systems by secondary school teachers. *Computers & Education*, 58, 688–696.
- Diez-Roux, A. (2000). Multilevel analysis in public health research. *Annual Review of Public Health*, 21, 171–192.
- Division of Instructional Innovation and Assessment, University of Texas at Austin (2007). *Item analysis*. Retrieved from <http://www.utexas.edu/academic/ctl/assessment/iar/students/report/itemanalysis.php>
- Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly*, 18, 585–609.
- European Commission (Ed.). (2004). *Europe needs more scientists!* Brussels: European Commission, Directorate-General for Research, High Level Group on Human Resources for Science and Technology in Europe. Retrieved from [http://ec.europa.eu/research/conferences/2004/sciprof/pdf/final\\_en.pdf](http://ec.europa.eu/research/conferences/2004/sciprof/pdf/final_en.pdf)
- European Commission (Ed.). (2006). *Science education now: A renewed pedagogy for the future of Europe*. Brussels: European Commission Directorate-General for Research. Retrieved from [http://ec.europa.eu/research/sciencesociety/document\\_library/pdf\\_06/report-rocard-on-science-education\\_en.pdf](http://ec.europa.eu/research/sciencesociety/document_library/pdf_06/report-rocard-on-science-education_en.pdf)
- European Commission (Ed.). (2012). *She figures 2012. Gender in research and innovation*. Brussels: European Commission Directorate-General for Research. Retrieved from [http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/she-figures-2012\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/she-figures-2012_en.pdf)
- Felder, R. M., Felder, G. N., Mauney, M., Hamrin, C. E., & Dietz, E. J. (1995). A longitudinal study of engineering student performance and retention. Gender differences in student performance and attitudes. *Journal of Engineering Education*, 84, 151–163.
- Goldstein, H. (1995). *Multilevel statistical models*. London: Edward Arnold.
- Goldstein, H. (2003). *Multilevel statistical models. Third Edition*. London, Edward Arnold.
- Hamuy, E., & Galaz, M. (2010). Information versus communication in course management system participation. *Computers & Education*, 54, 169–177.

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organisations. *Academy of Management Review*, 32(4), 1199-1228.

Hox, J. J. (1994). *Applied multilevel analysis*. Amsterdam: TT-publikaties.

IEEE, Learning Technology Standards Committee (2005). *WG12: Learning object metadata*. Retrieved from <http://ltsc.ieee.org/wg12/>

Johnson, D. W., & Johnson, R. T. (1996). Cooperation and the use of technology. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 785–812). London: MacMillan.

Kay, R.H., & Knaack, L. (2005). Developing learning objects for secondary school students: A multi-component model. *Interdisciplinary Journal of Knowledge and Learning Objects*, 1, 229–254.

Kay, R. H., & Knaack, L. (2007). Evaluating the learning in learning objects. *Open Learning*, 22, 5–28.

Kay, R.H., & Knaack, L. (2008). A formative analysis of individual differences in the effectiveness of learning objects in secondary school. *Computers & Education*, 51, 1304–1320.

Kim, S. (2009). *The Conceptualization, utilization, benefits and adoption of learning objects* (Doctoral dissertation). Virginia Polytechnic Institute and State University, Blacksburg, VA . Retrieved from <http://scholar.lib.vt.edu/theses/available/etd-12212009-212450/>

Kirschner, F., Paas, F., & Kirschner, P. A. (2009a). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior*, 25, 306–314.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts—A conceptual analysis. *Educational Psychology Review*, 18, 159–185.

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, 136(6), 1123.

Lohman, D. F. (1994). Spatial ability. In R. J. Sternberg (Ed.), *Encyclopedia of intelligence* (Vol. 2, pp. 1000–1007). New York, NY: Macmillan.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "General intelligence, objectively determined and measured". *Journal of Personality and Social Psychology*, 86(1), 96.

Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science*, 1, 316–345.

- Mayer, R. E. (2001). *Multimedia learning*. Cambridge, MA: University Press.
- Mayer, R. E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction*, 13, 125–139.
- Mayer, R. E. (Ed.) (2005). *The Cambridge handbook of multimedia learning*. Cambridge: University Press.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38, 43–52.
- Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 86, 389.
- National Governors Association. (2007) *Innovation America: Building a science, technology, engineering and math agenda*. Washington, D.C.
- Nezlek, J. B. (2008). An introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass*, 2, 842-860.
- Nurmi, S., & Jaakkola, T. (2006). Effectiveness of learning objects in various instructional settings. *Learning, Media, and Technology*, 31, 233–247.
- Organisation for Economic Co-operation and Development. (2006). *Assessing scientific, reading and mathematical literacy. A framework for PISA 2006*. Paris: OECD.
- Organisation for Economic Co-operation and Development. (2007). *PISA 2006. Science competencies for tomorrow's world*. Paris: OECD.
- Organisation for Economic Co-operation and Development. (2008). *Encouraging student interest in science and technology studies*. Paris: OECD.
- Organisation for Economic Co-operation and Development. (2013). *PISA 2012 results in focus. What 15-year-olds know and what they can do with what they know*. Paris: OECD.
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38, 63–71.
- Paivio, A. (1978). A dual coding approach to perception and cognition. In J. H. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 39-51). Hillsdale, NJ: Erlbaum.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255–287.
- Raes, A., Schellens, T., & De Wever, B. (2013). Web-based collaborative inquiry to bridge gaps in secondary science education. *Journal of the Learning Sciences*, (ahead-of-print), 1-32.

- Rasbash, J., Steele, F., Browne, W., & Goldstein, H. (2009). *A user's guide to MLwiN, version 2.10*. Bristol: Centre for Multilevel Modelling, University of Bristol.
- Resta, P., & Laferrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review*, 19, 65–83.
- Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69-97). Berlin: Springer-Verlag.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127.
- Slotta, J. D., & Linn, M. C. (2009). *WISE science: Web-based inquiry in the classroom. Technology, education connections*. New York, NY: Teachers College Press.
- Soller, A. (2001). Supporting social interaction in an intelligent collaborative learning system. *International Journal of Artificial Intelligence in Education*, 12, 40–62.
- Super, D. E., & Bachrach, P. B. (1957). *Scientific careers and vocational development theory*. New York, NY: Bureau of Publications, Teachers College, Columbia University.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12, 257–285.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295–312.
- Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251-295.
- U.S. Department of Education. (2007). *Report of the Academic Competitiveness Council*. Washington, DC.
- Van der Leeden, R., 1998. Multilevel analysis of repeated measures data. *Quality and Quantity*, 32, 15–19.
- United Nations. (2011). *Population division. World population prospects - The 2010 revision*. New York, NY: United Nations. Retrieved from <http://esa.un.org/wpp/>
- Valsiner, J., & Van der Veer, R. (2000). *The social mind: Construction of the idea*. Cambridge: University Press.

- Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills*. Englewood Cliffs, NJ: Educational Technology Press.
- 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.
- Voyer, D., & Voyer, S. D. (2014). Gender Differences in Scholastic Achievement: A Meta-Analysis. *Psychological Bulletin*. Advance online publication. <http://dx.doi.org/10.1037/a0036620>
- Wai, J., Lubinski, D., & Benbow, C.P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101, 817–835.
- Weathington, B. L., Cunningham, C. J., & Pittenger, D. J. (2010). *Research methods for the behavioral and social sciences*. New Jersey, NJ: John Wiley & Sons.
- Webley, K. (2013, June 06). The adaptive learning revolution. *Time*. Retrieved from <http://nation.time.com/2013/06/06/the-adaptive-learning-revolution/>
- Wiley, D.A. (2000). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In Wiley, A. (Ed.), *The instructional use of learning objects* (pp.1-35). Retrieved from <http://reusability.org/read>
- Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist*, 24, 345–376.
- Wong, L.H., & Looi, C.K. (2012). Swarm intelligence: new techniques for adaptive systems to provide learning support. *Interactive Learning Environments*, 20, 19–40.



## Annex

Table 4. Multilevel parameter estimates for the four-level analyses of learning outcomes.

Parameter	Model 0	Model 1
<i>Fixed part</i>		
<b>Pretest</b>	<b>57.18*** (1.50)</b>	<b>59.90*** (2.59)</b>
Collaborative setting (MColTSPW)		-1.60 (3.57)
Woman (FIndTSPW)		-4.85 (3.03)
Collaborative setting*Woman (FColTSPW)		4.95 (4.13)
MGL LP (MIndMGL)		1.39 (4.02)
MGL LP*Collaborative setting (MColMGL)		-2.43 (7.66)
MGL LP*Woman (FIndMGL)		-0.53 (4.57)
MGL LP*Collaborative setting*Woman (FColMGL)		-10.27 (8.54)
<b>Posttest (MIndTSPW)</b>	<b>7.31*** (1.09)</b>	-2.66 (2.59)
<b>Post*Collaborative setting (MColTSPW)</b>		<b>7.59* (3.62)</b>
<b>Post*Woman (FIndTSPW)</b>		<b>11.45** (3.74)</b>
<b>Post*Collaborative setting*Woman (FColTSPW)</b>		<b>-10.28* (5.09)</b>
<b>Post*MGL LP (MIndMGL)</b>		<b>17.88*** (4.18)</b>
Post*MGL LP*Collaborative setting (MColMGL)		-14.07 (8.78)
Post*MGL LP*Woman (FIndMGL)		-10.37 (5.73)
Post*MGL LP*Collaborative setting*Woman (FColMGL)		8.05 (11.01)
<b>Retention test</b>	<b>14.75*** (1.13)</b>	<b>12.93*** (2.59)</b>
Retention*Collaborative setting		2.04 (3.64)
Retention*Woman		-0.43 (3.74)
Retention*Collaborative setting*Woman		1.48 (5.11)
Retention* MGL LP		-1.03 (4.25)
Retention*MGL LP*Collaborative setting		0.00 (0.00)
Retention*MGL LP* Woman		6.58 (5.88)
Retention*MGL LP*Collaborative setting*Woman		0.00 (0.00)
<i>Random Part</i>		
Level 4: School		
Intercept/Intercept	3.19 (6.60)	0.00 (0.00)
Level 3: Class		
<b>Intercept/Intercept</b>	<b>26.15*** (10.86)</b>	<b>18.91*** (7.68)</b>
Level 2: Student		
<b>Intercept/Intercept</b>	<b>39.57*** (9.62)</b>	<b>42.24*** (9.40)</b>
Level 1: Knowledge test		
<b>Intercept/Intercept</b>	<b>208.57*** (11.42)</b>	<b>195.89*** (10.73)</b>
<i>Model fit</i>		
-2*loglikelihood:	8568.51	8512.92
$\chi^2$		55.60
df		21
<i>p</i>		<.001
<i>Reference model</i>		Model 0

*Note.* Reference information on parameters and standard errors for Model 0 and Model 1 are in parentheses. M = male; F = female; Ind = individual; Col = collaborative; TSPW = Text, Schemes, Pictures and Web-based exercises learning path; and MGL = Mayer GuideLines learning path.

\*  $p < .05$  \*\*  $p < .01$  \*\*\*  $p < .001$