With United Forces:

How **Design-Based Research** can Link **Theory and Practice** in the Transdisciplinary Sphere of **CLIL**

SILVIA BAUER-MARSCHALLINGER

UNIVERSITY OF VIENNA

silvia.bauer-marschallinger@univie.ac.at

This paper intends to exemplify how a design-based research (DBR) methodology can be used to put theory into practice by reporting on a first research cycle of a larger DBR project in the context of upper secondary CLIL history education in Austria.

The project aims to identify design principles of teaching techniques and materials which both support the acquisition of subject-specific competences and language. To this end, this study draws on Dalton-Puffer's (2013) construct of Cognitive Discourse Functions (CDFs), comprising seven key categories of academic language functions which have also been shown to be closely linked to historical competences.

In the course of this study, the researcher and a collaborating teacher systematically developed CDF-based history materials, which were then applied in the classroom and continuously evaluated, using interviews, observations, and written tasks for data collection.

Results of the first research cycle suggest that students lack awareness of possible connections between content and language learning and struggle with expressing complex historical content. Both teacher and students responded positively to the intervention on a general level, but pointed out a number of potential refinements, such as a more continuous and balanced intertwining of content and language.

Die vorliegende Studie soll veranschaulichen, wie Design-Based Research (DBR) im transdisziplinären Raum, der für CLIL bezeichnend ist, dazu beitragen kann, Theorie und Praxis besser zu verschränken. Diese Studie stellt den ersten Forschungszyklus eines Dissertationsprojekts dar, welches die Integration von Sprach- und Fachlernen im englischsprachigen Geschichtsunterricht der Sekundarstufe II näher beleuchtet.

Genauer gesagt versucht dieses Forschungsprojekt Design Prinzipien für CLIL Lehr- und Lernmaterialien zu identifizieren, welche sowohl die Entwicklung fachlicher Kompetenzen als auch sprachlicher Kompetenzen fördern. Zu diesem Zweck bedient sich diese Studie an Dalton-Puffers (2013) Konstrukt kognitiver Diskursfunktionen (CDFs), welches sieben zentrale akademische Sprachfunktionen definiert und sehr eng mit historischen Kompetenzen verbunden zu sein scheint.

Im Zuge der Studie entwickelten ein Team bestehend aus Lehrperson und Forscherin systematisch CDF-basierte, kompetenzorientierte Materialien, welche im regulären Unterricht der Lehrperson angewandt und kontinuierlich evaluiert wurden. Zur Evaluierung dienten Unterrichtsbeobachtung, Interviews mit Schülerinnen und Lehrperson sowie schriftliche Aufgabenstellungen.

Die Ergebnisse dieses ersten Forschungszykluses suggerieren, dass SchülerInnen oftmals nicht bewusst ist, wie Sprach- und Fachlernen verbunden sein kann, obwohl sie durchaus Probleme damit haben, komplexere, geschichtliche Inhalte auszudrücken. Im Allgemeinen reagierten die Lehrperson und auch die Schülerinnen durchwegs positiv auf die durchgeführte Intervention, sie zeigten aber auch Raum für Verbesserung auf, wie etwa eine durchgängigere und ausgewogenere Verflechtung von Sprach- und Fachlernen.

KEYWORDS:
CLIL history; design-based research; Cognitive Discourse Functions; integrated pedagogy; subject literacy.

SCHLÜSSELBEGRIFFE:
CLIL Geschichte; Design-Based Research; kognitive Diskursfunktionen; integrierte Pädagogik; Fachsprache.

1. Introduction

Without doubt, the present CLIL research scene can be described as lively and well established within applied linguistics and education (albeit to a somewhat lesser degree), and also in practice, CLIL has secured its place in Europe's educational landscape. Looking at the current research agenda more closely, it seems that, after many years of either focusing on language or content learning in CLIL, one major concern is now how we can truly integrate content and language learning, as the label would suggest. One notion assumed to allow a genuine integration of language and content learning is the concept of Cognitive Discourse Functions (CDFs, cf. Dalton-Puffer, 2013). Section 2.1 conceptualizes the integration of subject-specific and language-didactic perspectives in CLIL from a wider perspective, Section 2.2 then discusses the operationalization of content-and-language-integrative approaches, and Section 2.3 reviews the concept of CDFs in more detail.

Another observation regarding CLIL research is that some of the initial enthusiasm concerning CLIL and its benefits, as put forward by numerous studies, has recently been questioned (Pérez Cañado, 2018). Previously, it was often assumed that teaching in a foreign language would provide a "language bath" that would naturally benefit the learning process (Dalton-Puffer, 2007, p. 3). Now, many researchers in the field take a more differentiated stance. For instance, Meyer et al. (2015), argue that "adopting a CLIL approach does not automatically lead to effective learning" (p. 44). Furthermore, knowing and understanding a lot about CLIL in theory, including notions of integration, does not necessarily translate into classroom practice. As Meyer et al. (2015) argue "[a] deeper understanding of how the integration of content and language can actually be conceptualised has only recently begun to emerge […], so there may be too few resources and materials for teachers to draw on" (p. 45).

What seems to be the missing element is a stronger link and interrelation between research and practice. In other words, very often, educational research does not carry into the classrooms and thus, one could argue, CLIL is not reaching its full potential yet. To improve the reality of CLIL teachers and learners, transdisciplinary research combining the expertise of researchers and practitioners needs to find its place on the research agenda. Working towards an operationalization of genuinely integrated approaches seems to be a crucial step towards more successful CLIL practice, but also towards a better (and more ecologically valid) understanding of content and language integration. This is the main aim of the study presented here, which forms a first research cycle of a larger, doctoral study. To be more precise, this project intends to systematically develop teaching materials based on CDFs for the context of upper secondary CLIL history education. This endeavour should also drive forward theoretical conceptualizations of content and language integration.

Aiming at bridging theory and practice effectively, a design-based research (DBR) methodology, joining the expertise of a practitioner and a researcher, has been adopted. Since one aim of this paper is to exemplify how DBR can be used to put theory into practice in the context of CLIL, the methodological section (Section 3), including a general outline of the approach, takes up more space than would normally be the case. Another major part of this paper is dedicated to the presentation and discussion of sample materials developed in this research cycle (Section 4). Then, the results of the first cycle (pilot cycle) are outlined in section 5. Finally, insights and implications are summarized and discussed in section 6.

2. Integration of subject-specific and language-didactic perspectives in CLIL

2.1 Conceptualizing the relation of content and language in CLIL

As already hinted at in the introduction, contrary to its label, research on CLIL has traditionally made a distinction between studies focused on language or content, respectively, with the former receiving considerably more attention than the latter (Dafouz et al., 2014). Most of these language-focused CLIL studies examined the effect of CLIL on the learners' foreign language proficiency and many reported some effectiveness for the different aspects of language acquisition, of course to varying degrees. Research on aspects of content learning in CLIL, however, has been comparatively sparse (Meyer et al., 2015). It seems that content educationalists tend to be sceptical of potential benefits of CLIL for subject-specific learning, suspecting that the use of a foreign language could overburden students and teachers and therefore impede content learning (Dalton-Puffer, 2007; Maset, 2015).

For all intents and purposes, however, a strict division between language and content learning misses the point of content and language integrated learning and it seems that the CLIL research community agrees on this now. For example, Nikula et al. (2016) published an entire edited volume on the theory of integration in CLIL education. Similarly, Cenoz and Ruiz de Zarobe (2015) quite dramatically state that integration is "the way forward in […] CLIL for the rest of the twenty-first century" (p. 90).

Of course, there are already a number of insightful studies integrating content and language perspectives, working with notions such as *subject-specific discourse, genre* or *subject literacies* (cf. Systemic Functional Linguistics, e.g. Llinares et al., 2012). These studies are very much rooted in linguistics and might therefore be inaccessible for teachers, especially if they are content-subject specialists rather than trained language teachers (Dalton-Puffer & Bauer-Marschallinger, 2019). Furthermore, previous studies tended to focus on written language, but classroom interaction is overwhelmingly in the oral mode.

Another aspect overlooked in previous CLIL studies is the conceptualisation of *content*. Very often, content learning is

conceptualized as declarative knowledge, i.e. knowing facts and figures. Examples for these studies would be Badertscher and Bieri (2009), Gablasova (2014) or Dallinger et al. (2016), which all suggest that attaining factual knowledge via an L2 is indeed (satisfactorily) possible, but the actual learning effect has yet to be determined. Looking at current European history curricula, declarative knowledge, however, does not seem to be the prime learning goal any longer (Heil, 2012; van Drie & van Boxtel, 2008). Instead, most curricula focus on procedural knowledge by defining historical skills and competences which target critical analysis of historical sources rather than a mere memorization and recount of historical events (Körber et al., 2007; Heil, 2012). Accordingly, the ulterior goal of most competence models is to educate learners to become mature, responsible, and historically aware citizens who are capable of historical reasoning.

2.2 Operationalizing content and language integration via Cognitive Discourse Functions

Irrespective of the conceptualization of *content* learning, most research insights in relation to content and language integration have not been made viable for classroom implementation (Meyer et al., 2015). This entails that teachers lack suitable materials integrating content and language perspectives (Morton, 2013) as well as conceptual understanding, which might also be a result of insufficient possibilities for CLIL training (Pérez Cañado, 2016). Addressing these related issues, Dalton-Puffer and Bauer-Marschallinger (2019, p. 33) point out that a conceptualisation is needed "that makes language a natural concern of non-language educators" (see also Meyer et al., 2015). This is now where the concept of Cognitive Discourse Functions (CDFs) comes in, which will be presented in more depth in the following section. CDFs can be defined as communicative patterns that are routinely used to express thought processes and appear to be an integral part of subject pedagogies, too (Dalton-Puffer et al., 2018). These language patterns very often overlap with curricular

subject goals, which tend to be expressed with the help of performative verbs, as exemplified by the following samples taken from the Austrian curriculum for upper secondary history education (Austrian Federal Ministry for Education, 2014):

The students can

- **describe** the influence of historical developments on individuals, society, and state

- **outline** and **analyse** social developments and **evaluate** their significance in a historical context

- **analyse** and **discuss** causes, motives, and impact of war[1].

Because of these interconnections, Dalton-Puffer (2013) argues that CDFs are able to bridge content and language pedagogies in a way that is acceptable and tangible for (content) teachers.

2.3 Dalton-Puffer's (2013) Construct of Cognitive Discourse Functions

Since Bloom's (1956) Taxonomy of Thinking Skills, a great number of frameworks of cognitive or academic discourse functions have been published. Some of these can be considered to be more general-cognitive, like Anderson and Krathwohl's (2001) revision of Bloom's taxonomy, while others are more subject-specific, such as Beacco et al. (2010). Dalton-Puffer (2013) reviewed a great number of these frameworks, identifying over 50 different functions. She then tried to systematize them according to communicative intentions about dealing with knowledge (cf. functional pragmatics, Ehlich & Rehbein, 1986) and ended up with a condensed construct comprising only seven main types, making it somewhat more viable for subject educators.

An updated version of the construct can be found below in **Table 1**:

Communicative Intention	Type	Members
I tell you how can we cut up the world according to certain ideas	CATEGORIZE	*Categorize, classify, compare, contrast, exemplify, match, structure, subsume*
I tell you about the extension of this object of specialist knowledge	DEFINE	*Define, identify, characterize*
I tell you details of what I can see (also metaphorically)	DESCRIBE	*Describe, label, identify, name, specify*
I tell you what my position is vis a vis X	EVALUATE	*Evaluate, argue, judge, take a stance, critique, comment, reflect, justify*
I tell you about the causes or motives of X	EXPLAIN	*Explain, reason, express cause/effect, deduce, draw conclusions*
I tell you something that is potential (i.e., non-factual)	EXPLORE	*Explore, hypothesize, predict, speculate, guess, estimate, simulate*
I tell you something external to our immediate context on which I have a legitimate knowledge claim	REPORT	*Report, inform, summarize, recount, narrate, present, relate*

Table 1. The construct of Cognitive Discourse Functions (Dalton Puffer & Bauer-Marschallinger, 2019, p. 35)

In the left column, communicative intentions of the individual types are defined while the right column presents a list of possible members. This list, of course, is not exhaustive and boundaries are not clear-cut.[2]

In a previous study by Dalton-Puffer and Bauer-Marschallinger (2019), these seven types were mapped against the competences underlying the Austrian history curriculum, indicating rather strong links between CDF-types and history skills. These connections were also empirically investigated in that study, looking at their realization in class and in testing situations. The results suggest that CDFs play an integral role in competency-oriented history teaching in Austria.

The next logical step, and the overall aim of this PhD project, would now be to develop pedagogical tools and materials based on the CDF construct to examine more closely how we can integrate content and language learning in real classrooms which also addresses the teachers' need for suitable content-and-language-integrative materials.

3. Methodology

As mentioned already, the present paper is part of a larger, ongoing doctoral study aiming at the systematic, cyclical development of CDF-based history materials within a design-based research framework. This paper presents the results of one of the two pilot cycles of this doctoral research project pursuing the following four research questions:

RQ1:

What kind of CDF-based pedagogical measures and materials do students need to improve and elaborate their verbalization of cognitive processes?

RQ2:

Which features should characterize language-and-content-integrative materials?

RQ3:

How do students respond to explicit teaching of CDFs in the history CLIL classroom?

RQ4:

How does CDF-oriented teaching affect the learners' performance of historical competences and academic language skills?

The outcomes of the pilot cycle will not be able to provide complete answers but first insights to these questions. Furthermore, this pilot cycle can shed light onto the usefulness of design-based research for transdisciplinary endeavours like this PhD project.

3.1 Design-based research (DBR)

DBR is a fairly novel approach aimed at bridging theory and practice by being dual-focused, meaning that DBR intends to further develop theory while also improving local educational settings via the production of model didactic tools (Barab & Squire, 2004). Putting it differently, DBR is not limited to the development and testing of interventions, since its interventions incorporate a set of theory-based assumptions concerning the learning process (The Design-Based Research Collective, 2003). As such, DBR is not primarily interested in "what works"; instead it predominantly focuses on "how we can make something work and why" (McKenney & Reeves, 2014, p. 143).

3.1.1 Characteristics of DBR

Regarding terminology, there are various labels besides *design-based research* such as *(educational) design-research, development(al) research or formative research,* all embodying a number of core characteristics. McKenney and Reeves (2012: 13-15) identified the following:

- **Theoretically oriented:** DBR is theory-based, striving towards new insights while considering practical issues as well.

- **Interventionist:** Interventions are core elements of DBR, intended to initiate improvement.

- **Collaborative:** In DBR, teacher and researcher work closely together. By doing so, it is hoped to bridge theory and practice, which addresses the research gap of operationalization as outlined above.

- **Responsively grounded:** DBR is set in naturally occurring test beds to create interventions that can succeed in real classrooms.

- **Iterative:** DBR is typically organized in research cycles to be able to adjust and continuously improve the design.

3.1.2 A typical DBR process:

McKenney et al. (2006) argue that authentic contexts require a flexible research design which can respond to emerging factors and thus, including several (typically qualitative) data collection methods is helpful. **Figure 1** below exemplifies a generic DBR process. This representation of a typical DBR process was created by drawing on McKenney and Reeve's (2012) conceptualisation of DBR outlined above as well as Fraefel's (2014) visualisation, which stresses the cyclicality and the close connection of design and implementation as core elements of DBR.

At the beginning of a typical DBR process, the local situation, (curricular) requirements, as well as theory are thoroughly
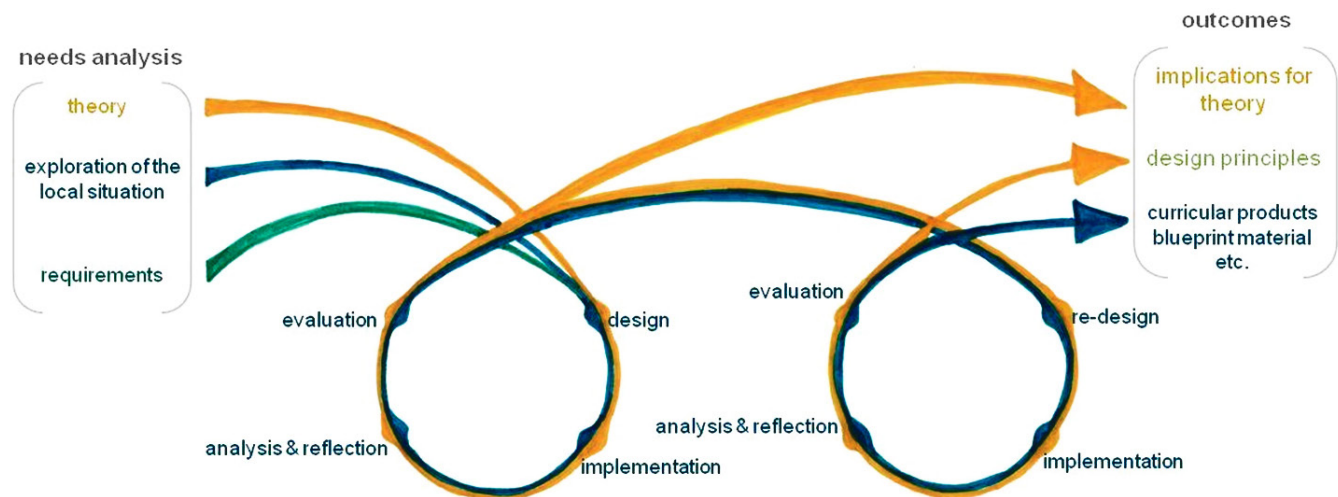
Figure 1. DBR as a cyclical research process, based on Fraefel (2014) and McKenney & Reeves (2012)

analysed. Taking these insights into account, teacher and researcher jointly design the intervention, which is then implemented by the teacher and observed by the researcher. Subsequently, process and product are examined and formatively assessed to improve and refine the intervention, which then undergoes another research cycle. Depending on the scale of the intervention and the goal of the study, these cycles can be repeated several times. At the end of a larger project, one should also evaluate the intervention summatively, i.e. checking effectiveness of the intervention.

3.1.3 Outcomes of DBR

Throughout the whole process, implications for theory should be continuously specified. Due the small scale of most DBR projects, these results cannot be generalized to other populations, but they might be generalized to theoretical models (i.e. analytic generalization, cf. Firestone, 1993). Practical outcomes of DBR are the curricular products, blueprint materials etc. developed in the course of the study. *Design principles* are another type of result in DBR, which can be defined as the following:

> *"If you want to design intervention X for the purpose/ function Y in context Z, then you are best advised to give that intervention the characteristics A, B, and C, and to do that via procedures K, L, and M, because of arguments P, Q, and R"* (van den Akker, 1999, p. 5)

As we can see, design principles do not strive to be independent from context. As such, it is the teacher's (or researcher's) task to adapt these principles for their own context (McKenney et al., 2006). This also means that DBR results permit case-to-case generalization. In other words, (parts of) the results can also be used in other contexts but might require some adaption first.

3.2 Context of the present study and participants

The first cycle of this larger project was set in a Viennese public, vocational upper secondary school focusing on business in which student population is predominately female. The school offers a five-year-long bilingual programme, which upon a successful graduation, gives students access to both professional labour market and university education. In the bilingual programme, at least 50% of schooling should be in English. Very often, English-natives co-teach with content-subject teachers who are not English teachers themselves, or native speakers who hold teaching qualifications for content subjects teach on their own. Non-native content teachers who are qualified English teachers as well usually teach without native co-teachers, which is also the case in this particular study. The teacher collaborating in this project is a German native speaker and has been teaching English and history (in the mono- and bilingual programme) for over 18 years now. However, she has not completed any formal in-service trainings on CLIL due to their non-availability in her first years of bilingual teaching[3].

The student group of this cycle attended grade 11 during data collection and consisted of 12 female learners. Their teacher described their competence- and motivation-levels for both English and history education as rather low, which was the prime reason for choosing this group for the intervention since it was considered that this group could benefit most from a new approach. Their language backgrounds were very multilingual: three listed more than one L1, three others reported not speaking German as their L1(s), and all students were learning a third language at school (Spanish or French).

3.3 Research process

As can be seen in **Figure 2**, at the beginning of a cycle, the needs of students and teachers were identified by conducting semi-structured interviews with students (focus group) and the teacher (individually) as well as via competency-based written tasks. These written tasks were modelled after the Austrian final history exam. For these tasks, the students were asked to analyse historical visual sources on various levels (deconstruction competence), and make connections to the present or their own historical identity (orientation competence). The prompts of these tasks involve performative verbs, such as 'explain', 'describe' or 'evaluate', and always target the same historical skills (on pre-defined levels) and CDF-types. Therefore, these task sets can be 'filled' with different topics (see **Figure 4** in Appendix for a sample template). The tasks prior to the intervention featured a topic previously dealt with and the post-intervention tasks dealt with the topic of the intervention, which, in this case, was ideologies of the 19th century. In light of the insights of the needs analysis, the researcher and the teacher collaboratively designed six lessons covering ideologies of the 19th century and incorporating CDF theory. The teacher then implemented these pedagogical tools in her classrooms (observed by the researcher in the pilot cycle and video-taped in the subsequent main cycles 1-3). As a next step, the process and product were formatively evaluated, using retrospective semi-structured interviews with the teacher (individually) and students (in groups) as well as competency-based written tasks once more.
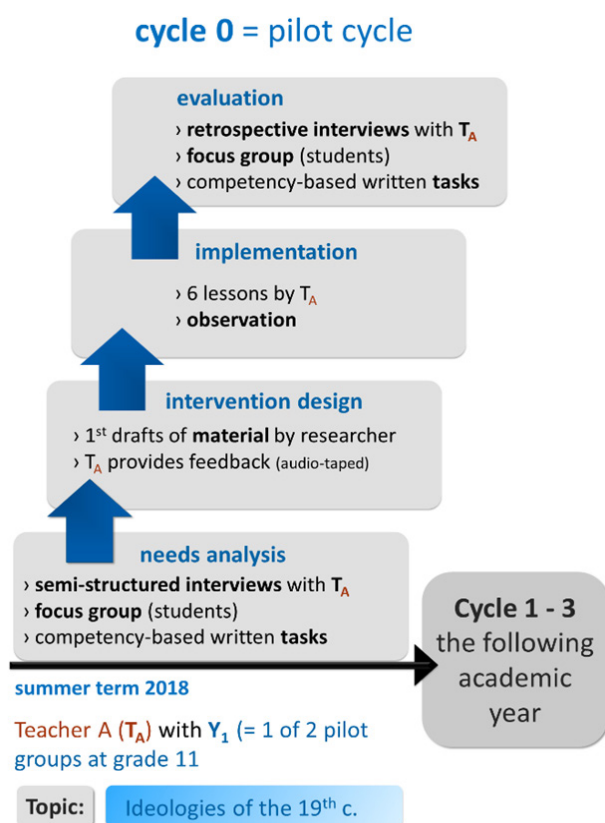
Figure 2. Data collection

In accordance with the results of the pilot cycle, the didactic materials were adapted, re-implemented, and evaluated in the following winter and summer term (main cycle 1, 2, and 3).

3.4 Data analysis

The semi-structured interviews were analysed according to qualitative content analysis, which allows to systematically structure and interpret (spoken and written) text data, usually interview data, via a process of coding recurring themes and categories (cf. Kuckartz, 2016). The competency-based written tasks were analysed by the researcher with the help of two different types of rubrics (see **Figures 5 and 6** in Appendix for samples), which have been designed for the purpose of the doctoral study. One rubric looks into the performance of history skills and is based on the history testing guidelines used for the design of the tasks. As such, it focuses on the following criteria:

- **Target level:** includes the achievement of the defined minimum level, the accuracy of the content, and the systematicity of the answers.

- **Target competence:** includes the performance of the intended competence, the degree of interaction with the source, and the amount of detail.

The second rubric is based on the CDF-construct, paying attention to the following aspects:

- Choice and (logical) composition of CDF-types.

- Appropriateness of linking in terms of form and function.

- Two typical features of historical discourse, which are also included in the intervention, namely nominalisation and hedging.

Both rubrics defined three levels plus a zero level, with 3.0 being the best result. This 3-point scale somewhat reflects the three stages of historical learning defined in the history testing guidelines. Technically speaking, this would mean that this rating scale is ordinal. However, similar to school grades, one can treat them as metric since the distance between the individual ratings is theoretically constructed to be the same and includes a zero level as well. A metric scale presents the advantage of allowing calculations of means and t-tests when comparing results. Yet, due to the small sample, t-tests are likely to point towards insignificance, meaning that all changes could also be, statistically speaking, just random. Therefore, and also considering the overall thrust of the study, it should be noted though that any quantifying calculations are not the focus of the study and not representative. Instead, they simply provide some descriptive, additional information, showing how CDF-based teaching might affect the performance of history skills and academic language.

4. The intervention and sample material

The intervention of this research cycle focused on economic and political ideologies of the 19th century in accordance to the respective curriculum. In terms of language focus, the material featured all seven CDF-types, linking (in relation to certain CDF-types), as well as two typical features of historical discourse, namely nominalisation and hedging.

The intervention was scheduled for three consecutive double sessions (i.e. three times 100 min). After a brief introduction, including a teacher-led brainstorming on the term *ideology* and its connection to politics and economy, the procedure for this unit was outlined and students were allocated into groups of four. In these groups, the students were given a text and a worksheet on one of the four major ideologies of the 19th century (liberalism, conservatism, Marxism/ socialism/ communism, capitalism). The texts were taken from the internet and since they were not produced for language learners, they were slightly adapted in terms of complexity and length. Furthermore, a glossary was added for special terminology and unfamiliar words (according to the teacher's and researcher's judgment). The first part of the worksheet was basically designed so that students can work through their topic (mainly based on the text provided) by working on CDF-types.

In the second part, they were asked to analyse two historical sources related to their topic with the help of general, but explicit guidelines and examples on how to use language when analysing historical sources (see **Figure 3**). These guidelines are based on the initial needs analysis.

The students then had to utilize their output to prepare a presentation and a handout on 'their' ideology. Finally, the students presented their results in front of their peers, who all worked on different ideologies.

The main historical competences featured in this unit are *reconstruction, deconstruction* and *orientation competence* (cf. Körber et al., 2007). Deconstruction competence can be defined as the ability to critically analyse historical sources on various levels. Reconstruction competence refers to the ability of using historical knowledge (ideally based on the results of source analysis) to create historical narratives. Orientation competence describes the skills related to the management of one's historical awareness, i.e. understanding connections between the past and the present and how this affects our perceptions of the world. Below, a number of exercises used in this pilot cycle illustrate our approach.

In **Exercise 1** below, the students were given explicit information concerning the CDF-type DEFINE. They were then asked to apply this knowledge first passively in the text provided and then actively by coming up with their own definition of a different term (output-scaffolding).

Language tips for source analyses:

1. When you need to *summarize*, make sure to convey the central message and do not yet interpret.
2. When you are asked to *describe*, only describe what you see, do not yet interpret.
3. Try to work closely with the source and do not write about something unrelated to the source (unless it's explicitly stated in the task).
4. Always back up your interpretations, opinions, and judgments. Give reasons, examples and/ or refer to facts.
5. Distinguish between fact, opinion, or interpretation.
6. Use language to mark whether something is **fact**, *opinion,* or interpretation.
   e.g. we know from other sources/ historians seem to agree that/ it's a well known fact that...
   From my point of view/ I am of the opinion/ I believe, etc.
   Looking at this quote it seems that/ we assume/ this indicates that/ this could mean that...
   → Show that your answer is not the only interpretation! You can use modals like could or would or you could add adverbs/ adjectives like probably, possible etc.

Figure 3. Language tips for source analyses.

1) Spot the *definitions* of Conservatism in paragraph A and C.

Usually, definitions consist of the following parts: <u>term</u> = **broader category** + *specifics/ description* (+ further information/ limitations/ examples); e.g. <u>A cat</u> is a **small animal** with *fur, four legs, a tail, and claws*, usually kept as a pet or for catching mice.

    a. Now, highlight these parts in the two definitions given in paragraph A and C, using different colours.
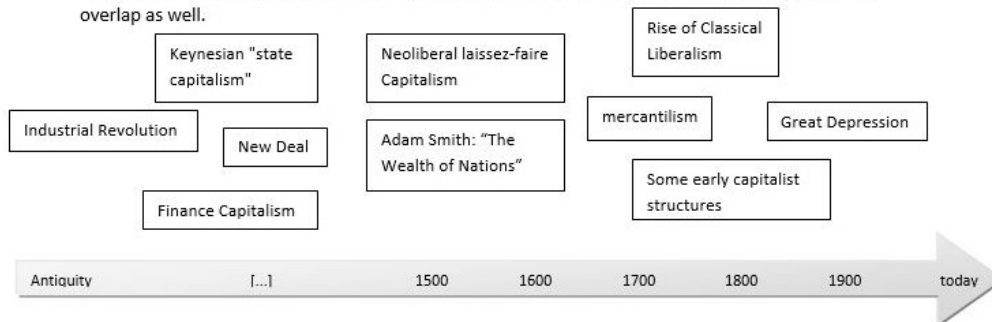
Exercise 1. DEFINE

In **Exercise 2**, the students were asked to put key steps of the historical development of capitalism into the right sequence. This was intended to help them make sense of a rather dense and complex text. As a next step, the students had to report this historical development in their own words, which was aimed at the CDF-type REPORT (and EXPLAIN and DESCRIBE as sub-elements) as well as *reconstruction competence*. They were also provided with useful phrases to describe developments and cause-and-effect relations, as these were the elements of this summary. By using the key steps in nominal form, the students' attention was drawn towards the use of nominalisation in historical narratives. This was explained in the box below. In this exercise, the students were made aware of typical features of historical

discourse and, at the same time, could implement them in their own production in a guided way.

Finally, **Exercise 3** shows an analysis task of a written source for the group working on liberalism. Using their previously acquired knowledge on the topic as well as the language tips for source analyses (**Figure 3**), the students were asked to analyse historical sources, which they were supposed to present either in their class presentation or handout. Difficult or archaic vocabulary was provided in a glossary. In the teacher's version, target competences, activity type (as defined in the guidelines for competency-based history testing) and target CDF-type (with the main type in bold) were provided in square brackets.

2) *Reporting* historical development:

Here are key steps in the historical development of capitalism. Put them in the right order by indicating where they would be on the timeline. Please be aware that sometimes two boxes could overlap as well.

Keynesian "state capitalism"		Neoliberal laissez-faire Capitalism		Rise of Classical Liberalism	
Industrial Revolution	New Deal	Adam Smith: "The Wealth of Nations"	mercantilism		Great Depression
	Finance Capitalism			Some early capitalist structures	

| Antiquity | [...] | 1500 | 1600 | 1700 | 1800 | 1900 | today |

Now, try to concisely[2] *report* this development, using the terms from the boxes above as well as the linking words below:

Useful words to describe cause-and-effect and development: *These phrases help you to express reasons and consequences and describe how and why something developed.*

lead to	in the course of	another important development/ factor	as a consequence	result in
eventually	in the light of	in particular	cause	rooted in	based on	therefore
due to	because of	for this reason	subsequently

By using the terms from the boxes above the timeline, you've written in a **nominal style**, meaning that you used nouns rather than verbs. This is a very common strategy for formal, written English, especially in history when we talk about concepts rather than individual actions.

Exercise 2. REPORT, EXPLAIN, DESCRIBE, and nominalisation.

SOURCE ANALYSIS

A. Have a look at the quote and do the following tasks below:

"We cannot doubt that self-interest is the mainspring[2] of human nature. It must be clearly understood that this word is used here to designate[3] a universal, incontestable[4] fact, resulting from the nature of man, and not an adverse[5] judgment, as would be the word selfishness." [6]

Frédéric Bastiat, French politician and economist, 1801.

a) Summarize the quote in your own words. Make sure to include the difference between self-interest and selfishness, according to Bastiat. [*deconstruction,* Reproduktion (I), **REPORT**]
b) Explain how this quote is related to liberalism. [*deconstruction*, Transfer (II), **EXPLAIN,** CLASSIFY]
c) Do you think that Bastiat can be considered a liberal? Give reasons for your judgment. [*deconstruction*, Reflexion (III), **EVALUATE,** EXPLORE, CLASSIFY,]
d) Do you agree with Bastiat? Argue your opinion. [*orientation*, Reflexion (III), **EVALUATE,** EXPLAIN]

Exercise 3. Source analysis. Teacher version.

5. Findings

This section provides an overview of the main insights gathered in this research cycle. The results of the written tasks indicate how CDF-based teaching affects the learners' output, but also – if not mainly – provide information on which aspects should be featured more prominently in future versions of the intervention. The interviews also provide information on the learners' and teacher's needs as well as a first evaluation of the intervention.

5.1 Written competency-based tasks[5]

Table 2 provides an overview of the written task results, comparing overall pre- and post-intervention results. Without getting too much into detail, we can see that both content and language results improved from the first to the second testing with substantial change on the content scale and only minor change in terms of language. On average, both dimensions started at roughly the same value, but content results increased considerably more and also tended to be better than the respective language results for most students. Interestingly, while the average range decreased for content results, it widened for language outcomes from T1 to T2.

student code	Content		Language	
	T1	T2	T1	T2
ADP10	2.07	2.18	1.67	2.00
ARE09	2.50	2.50	2.33	2.33
ARG03	0.82	1.50	1.00	1.00
ART11	1.15	2.27	1.17	1.83
ATQ11	1.40		2.00	
AUM06	1.55	2.42	1.33	1.67
ETE10	0.98	1.97	1.50	1.67
IAP07	2.23	2.50	1.50	2.00
ICA08	1.25	1.70	1.67	0.83
SIH07	1.62	1.62	1.67	1.67
SOA09	1.33	2.27	1.33	1.67
SOG09		1.43		1.67
Average	**1.54**	**2.03**	**1.56**	**1.67**
Range	**1.68**	**1.07**	**1.33**	**1.50**

Table 2. Overall results of written tasks (pre- and post-intervention)

	T1 average	T2 average
choice of CDF-types	2.18	2.20
composition of CDF-types	2.00	2.10
linking in terms of function	1.91	1.80
linking in terms of form	1.91	2.40
use of hedging	0.73	0.90
use of normalisation	0.64	0.60
overall (language)	**1.56**	**1.67**

Table 3. CDF-based rating (pre- and post-intervention)

Turning to the language results (**Table 3**), the most obvious result is that there is hardly any evidence of (appropriate) use of hedging and nominalisation. Consequently, this is something that should be allocated more time in future interventions. Secondly, linking within and between CDF-types was not always successful, especially in terms of function, meaning that sometimes, students used linking devices that were not appropriate for the CDF-type they employed, as exemplified by **extract 1** [6]:

Extract 1

Yes because in some countries such as China there is still hard labour <u>because</u> workers work very long and to much at bad working conditions. They also get less paid for their work. And these products are still exported then.

Here, the student was asked to argue whether (or in which ways) a caricature[7] might still be relevant in the 21st century. The second because does not fit as it introduces a purely descriptive report rather than a real explanation.

Another observation is that whenever CDF-types were chosen outside the intended target range, students often missed the main point of a task, which also affected content-related outcomes, mostly in the form of low results for accuracy and relevance of content. **Extract 2** provides an example:

Extract 2

The men only want to get money, capitalism is only about money without looking after your environment.

The task here was to assess the connection of the caricature and capitalism. In this sample, the learner does not EVALUATE, i.e. taking a stance and providing any reasons for her claims or even referring to the source somehow. Instead, she only provides a superficial statement of something she remembers from class, which constitutes an example of REPORT, resulting in lower content-related results, too, due to lack of relevance and reference to the source material.

Finally, it should be mentioned that there are only slight gains from the first to the second testing. Unsurprisingly, this suggests that intervening just for one unit does not significantly affect the students' use of CDFs[8].

	T1 average	T2 average
target level	1.95	2.15
accuracy of content	1.53	2.24
systematicity	1.41	2.10
overall level	**1.70**	**2.14**
target competence	1.76	2.11
justification/ comprehensibility	1.25	1.89
scope	1.08	1.79
overall competence	**1.37**	**1.93**
overall	**1.54**	**2.03**

Table 4. History-based rating (pre-and post-intervention)

Turning to the history-related results (**Table 4**), it can be seen that the results of the second sitting were considerably better. Level *results* were generally better than competence results. Good level *results* mean that the students tended to perform the intended historical-cognitive action, presenting accurate and relevant content in a systematic way. The biggest issues seem to revolve around justifications and amount of details.

Many answers were vague and superficial, like dropping big phrases (e.g. "the rich getting richer, the poor getting poorer") without really connecting them to the concept or source at hand. For instance, when asked about the connection of the caricature and capitalism, one student wrote (**Extract 3**):

Extract 3

Yes, I think it is connected to capitalism because people in capitalism had to work a lot and rich people got even richer.

Again, this student reports bits and pieces that she remembers from class about capitalism, but what she is reporting here is not really what is depicted in the source. Furthermore, many answers, like the one presented above, were very general, vague, and lacking detail. Connected to that, the learners often did not perceive the source in their contemporaneous context. For example, when asked about possible intentions of the artist who produced the caricature, they often wrote answers like "to show us how it was". Again, this is very general and does not consider the source in its temporal context.

However, it should be noted that after the intervention, all of these observations were considerably less frequent and grave, as **Table 4** shows.

5.2 Interviews with students

Looking at learner needs from the students' perspective, it becomes quite apparent that students were not very aware of their own language-related learning needs, in general and in terms of CDF use. The only problematic aspect they reported was limited vocabulary, as expressed by this student:

English translation

`ETE10: […] You can't recall another word and then you say it the wrong way and then it's wrong, and people will think you don't know anything. But you do know, you just don't know how to express it.`

Original quote

`ETE10: = [...] Man findet kein anderes Wort, und dann sagt man es falsch, und dann ist es halt falsch, und dann glaubt der, du kannst es nicht, aber du kannst es ja, du weißt nur nicht, wie du es ausdrücken sollst.`

However, even though they felt like they did not have any language needs other than lexical ones, when asked more specifically or confronted with some previous observations, some did say that they struggled with complex historical content:

English translation

`ADP10: 'You'd normally think 'yes', I understand, so I should be able to get it across, but this doesn't always work for me.'`

Original quote

`ADP10: man denkt sich, ja ich verstehe es, dann kann ich's auch rüberbringen, so wie ich es verstehe, aber das funktioniert bei mir manchmal nicht.`

In general, though, it seems that students were not very aware of possible links between content and language learning. Initially, they did not really see the point of focusing on language in content subjects since they considered language and content as separate entities. As the discussion developed, however, some re-evaluated some of their own learning experiences and started to question this assumption. For instance, initially, they all said that they always knew what they needed to do when confronted with performative verbs in testing situations, but during the discussion, they started to view it more nuancedly or even changed their mind, like in this extract:

English translation

SOG09: Well, I find 'analyse' always so (.) so difficult, because it means that one should write a lot about it [=the topic] (.) a whole es-, yes I have to write a whole essay about it now, or I don't know, or I have to just briefly say what it is about, or definitions? I really don't know what and how exactly.

Original quote

SOG09: *Also ich finde 'analyze' is immer so (.) so schwer, weil das heißt man muss jetzt ur extrem viel drüber schreiben (.) einen ganzen Auf-, ja ich muss jetzt einen ganzen Aufsatz darüber schreiben, oder ich weiß nicht, oder eben muss ich nur sagen, kurz worum es geht, oder Definitionen? also ich weiß nicht was, wie genau.*

Moving on to the students' evaluation of the intervention, it seems that they enjoyed this unit, especially because of its hands-on approach, the guiding structure, the different ways of scaffolding as well as the language boxes. In general, though, the historical topic was perceived to be rather challenging due to its abstractness and conceptual complexity.

The students reported that they were more engaged and active than usual. They also believed that they learned a lot about their own topic but had difficulties grasping what other groups presented. They felt that the step-by-step structure as well as several scaffolding techniques helped them understand when working through the content. For example, one student said:

English translation

ADP10: 'I really liked working on the poster [= historical source on socialism] because then I [...] understood socialism better.'

Original quote

ADP10: *Ich fand das mit dem Poster total gut, weil dann habe ich [...] Sozialismus*

besser verstanden.

As scaffolding elements were obviously missing in the students' presentations, one could argue that the lack of these elements in the presentations made it difficult to understand these rather complex concepts.

Generally, the structure of this unit put the teacher more in the role of a 'coach'. However, as students were not really used to focusing on language in history and the topics were rather challenging, a more prominent teacher role would have been necessary. Especially the weaker students struggled throughout the project and would have needed more teacher support. Therefore, the overall structure of having groups presenting their topic to the other students should be reconsidered.

Other necessary changes involve modification of individual tasks as some prompts were somewhat ambiguous or their execution was not very practical (e.g. in Exercise 2, Section 4, as many of the developments to be put onto the timeline overlapped). Another point to keep in mind is that language boxes are only helpful if they are purposeful as otherwise they might distract the learners and overcomplicate tasks. Finally, tasks only focusing on lexico-grammatical aspects do not suffice, e.g. only underlining certain grammatical features does not really promote the understanding of a historical concept. So, either one could add another activity focused on the content or the activities should be more content-and-language-integrative from the start. For example, the task "underline all verbs used to report Marx's views" could be changed to "underline all verbs used to report Marx's views and explain how the choice of verbs influences the message". Of course, the exact parameters and features of these content-and-language-integrative activities will be further explored in the main research cycles of this PhD project.

> "Another point to keep in mind is that language boxes are only helpful if they are purposeful as otherwise they might distract the learners and overcomplicate tasks."

> "Tasks only focusing on lexico-grammatical aspects do not suffice, e.g. only underlining certain grammatical features does not really promote the understanding of a historical concept."

5.3 Interviews and design sessions with the teacher

In the needs analysis interview with the teacher, she agreed with the students that vocabulary (academic and field-specific) is indeed a central issue for them, but she also stated that responding to performative verbs was a huge challenge for most students. She argued that the students:

English translation

```
... simply don't know, because they can't
associate anything with performative verbs.
If it says 'explain, describe', they don't
know what to do.
```

Original quote

```
... nicht wissen, weil sie können einfach
nix anmachen äh nix äh verbinden mit den
Operatoren (.) nichts anfangen, das wollte
ich sagen (.) Wenn da steht erkläre,
beschreibe, dann wissen sie nicht was sie
tun sollen (.)
```

Another, more general issue identified by the teacher was lack of motivation and engagement. She therefore stressed the importance of student engagement for the design of the intervention.

As for teacher needs, two major aspects were discussed, namely the need for more (engaging) materials and resources suitable in terms of content- and language-level for CLIL history education and for clear guidelines in terms of assessment and the role of language. In this respect, the teacher reported:

English translation

```
TA: During my first years, I was inculcated
not to teach language in content subjects.
I was always told 'content teaching is not
language teaching' and at the beginning,
this was really hard for me, but I got used
to it. […] And now I don't know where we are
headed at.
```

Original quote

```
TA: Nein gar nicht. Ja weil ich des, das
ist mir so eingetrichtert worden in den
ersten Jahren (.) Hat immer geheißen
sprach- ah nein der Fachunterricht ist
kein Sprachunterricht (.) und das ist
wirklich schwer am Anfang. Aber wie gesagt
mittlerweile habe ich mich daran gewöhnt (.)
und jetzt schauen wir mal.
```

Obviously, this need for clearer guidelines also affects materials selection and design. She further stated that she created a lot of material herself and most resources online required some sort of modification to fit the students' needs.

In the design sessions, organisational matters were discussed and first drafts of exercises were scrutinized, rearranged, and specified. The teacher commented on the feasibility of tasks and changes were made, including the exclusion of some tasks. The texts were adapted too, and agreement was reached on which words should be included in the glossary. The teacher and the researcher realized that including some more explanatory comments in the lesson plan and the teacher's version of the resources would help teachers implement the materials. Furthermore, it was agreed that comprehensive, annotated solutions would be helpful for teaching this unit.

In the post-intervention interview, the teacher's assessment of and reflection on the intervention was mainly positive. She thought that the didactic tools were successful as they were engaging and centred on students. She also appreciated the structure and scaffolding techniques. To address the problems concerning the overall structure of groups presenting 'their' topic to their peers and overload in general, it was decided to have more teacher guidance in the next cycle. This does not mean that there would be a turn towards teacher-centred teaching but rather that there should be more frequent breaks between group-/pair work to discuss individual steps in class[10]. Furthermore, to ensure in-depth treatment of tasks and contents, it was agreed to sacrifice some bits of the scope of content. This also entails that fewer language-focused tasks should be included to allow a more thorough treatment of the exercises. Additionally, to avoid the negative connotations of grammar, it might be a good idea to further stress the functional aspect of language in the future. As students were not used to focusing on language in content subjects yet, it was assumed that this change in perspective might need more time to really make students aware of language within specific subjects. Another point mentioned in the interview with the teacher was the need to include a more extensive closing activity to facilitate potential uptake.

6. Conclusion

From a methodological perspective, one aim of this pilot study was to explore to what extent DBR is living up to the expectation of being able to effectively bridge theory and practice. Notwithstanding the organisational complexity, DBR turned out to be a methodological framework conducive to practice-oriented yet theoretically-

> **"Working on language should not be self-contained but always needs to be connected to the content via the linguistic choices that are being discussed."**

grounded educational research. By discussing lesson plans and materials in detail, the researcher and practitioner could successfully link theoretical and practical perspectives. However, DBR requires teamwork, commitment, and clear communication. But, as the teacher put it:

English translation

```
It's really nice when you're doing it in
pairs, then you'll realize all the things
that need adjustment.
```

Original quote

```
Das is aber auch das Nette, wenn man das
zu zweit macht, man kommt drauf da und da
happerts.
```

Turning to research insights, this research cycle has shown that students are often not aware of possible connections between content and language learning and consequently, they are also not used to focusing on language in content subjects. This entails that it is the teacher's task to first raise awareness in this respect. Yet, it should not stop with raising awareness. Especially at the beginning of language-based content teaching, teachers might need to support their learners more continuously. This also affects the overall structure of this intervention in the sense that more frequent breaks between group- or pair-work, in which individual steps are discussed, are necessary. For that, teachers need comprehensive materials and annotations to draw on, especially if they are not language experts.

As for the materials, one thing to keep in mind is to avoid tasks that are only concerned with lexico-grammatical aspects without thematising the content. These tasks tend to be rejected by the learners as they feel that content is neglected. Instead, one should either add another content-focused layer or create activities that integrate content and language more genuinely. This means that working on language should not be self-contained but always needs to be connected to the content via the linguistic choices that are being discussed.

Finally and most importantly, this research cycle has shown that working with and on CDFs seems to be an approach accepted by teachers and students to work on language in content subjects. In other words, CDFs are a notion conducive to language-based content teaching, which should ensure that learners do not just focus on subject-specific competences but also work on academic language connected to these content-related skills. ■

Notes

1 Translated by the author.

2 For a full discussion of the construct, see Dalton-Puffer (2013).

3 Now, there are some in-service teacher trainings available. Yet, they are not mandatory and come in different forms (some are only very short and/or focused on one aspect, others are more comprehensive, accompanying teachers for more than one semester).

4 For more information on historical competences, see section 4. The intervention and sample material.

5 As this is a pilot study, one purpose of this research cycle was also to check the usefulness of the individual task items, meaning that two tasks were cut from the template after analysis. All results presented here refer to results disregarding discarded items. The template provided in Appendix does not contain discarded items.

6 All extracts contain the learners' original mistakes and errors.

7 The caricature is available here: *http://www.independentlabour.org.uk/main/2010/07/12/ equality-of-sacrifice/*

8 In the subsequent (main) cycles, one group experienced two interventions and this group indeed showed great improvements on both language and content scales.

9 In Austria, performative verbs, so-called "Operatoren", were introduced for the new standardized final exams in all subjects and therefore, most learners are frequently confronted with those. Previous studies, e.g. Dalton-Puffer and Bauer-Marschallinger (2019), suggest that students do not always respond accordingly.

10 This structure was tried out in a second pilot cycle and worked well.

References

Anderson, L. W., & Krathwoh, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives* (Complete ed.). New York: Longman.

Austrian Federal Ministry for Education. (2014). *Lehrplan der Handelsakademie [Curriculum of the Commercial College]*. Retrieved from *https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10008944*

Badertscher, H., & Bieri, T. (2009). *Wissenserwerb im Content and Language Integrated Learning [Knowledge acquisition in Content and Language Integrated Learning]: Empirische Befunde und Interpretationen [empirical evidence and interpretations]* (1st edition). Bern, Stuttgart, Wien: Haupt.

Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *Journal of the Learning Sciences, 13*, 1–14.
https://doi.org/10.1207/s15327809jls1301_1

Beacco, J.-C., Coste, D., van de Ven, P.-H., & Vollmer, H. J. (2010). *Language and school subjects. Linguistic dimensions of knowledge building in school curricula: Document prepared for the Policy Forum The right of learners to quality and equity in education – The role of linguistic and intercultural competences:* Language Policy Division. Directorate of Education and Languages.

Bloom, B. S. (1956). *Taxonomy of educational objectives; the classification of educational goals.* New York: Longmans, Green.

Cenoz, J., & Ruiz de Zarobe, Y. (2015). Learning through a second or additional language: content-based instruction and CLIL in the twenty-first century. *Language, Culture and Curriculum, 28,* 1–7.
https://doi.org/10.1080/07908318.2014.1000921

Dafouz, E., Camacho, M., & Urquia, E. (2014). 'Surely they can't do as well': a comparison of business students' academic performance in English-medium and Spanish-as-first-language-medium programmes. *Language and Education, 28,* 223–236.
https://doi.org/10.1080/09500782.2013.808661

Dallinger, S., Jonkmann, K., Hollm, J., & Fiege, C. (2016). The effect of content and language integrated learning on students' English and history competences – Killing two birds with one stone? *Learning and Instruction, 41,* 23–31.
https://doi.org/10.1016/j.learninstruc.2015.09.003

Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning (CLIL) classrooms. Language learning and language teaching: Vol. 20.* Amsterdam: Benjamins.
https://doi.org/10.1075/lllt.20

Dalton-Puffer, C. (2013). A construct of cognitive discourse functions for conceptualising content-language integration in CLIL and multilingual education. *European Journal of Applied Linguistics, 1*(2), 216-253.
https://doi.org/10.1515/eujal-2013-0011

Dalton-Puffer, C., Bauer-Marschallinger, S., Brückl-Mackey, K., Hofmann, V., Hopf, J., Kröss, L., & Lechner, L. (2018). Cognitive discourse functions in Austrian CLIL lessons: towards an empirical validation of the CDF Construct. *European Journal of Applied Linguistics, 6*(1), 5-29.
https://doi.org/10.1515/eujal-2017-0028

Dalton-Puffer, C., & Bauer-Marschallinger, S. (2019). Cognitive Discourse Functions meet Historical Competences: Towards an integrated pedagogy in CLIL history education. *Journal of Immersion and Content-Based Language Education, 7*(1), 30-60.
https://doi.org/10.1075/jicb.17017.dal

The Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry, 32, 5–8. Retrieved from *http://www.designbasedresearch.org/reppubs/DBRC2003.pdf*
https://doi.org/10.3102/0013189x032001005

Ehlich, K., & Rehbein, J. (1986). *Muster und Institution [Pattern and institution]: Untersuchungen zur schulischen Kommunikation [Studies on school communication]. Kommunikation und Institution: Vol. 15.* Tübingen: Gunter Narr.
https://doi.org/10.1017/s0272263100005696

Firestone, W. A. (1993). Alternative Arguments for Generalizing From Data as Applied to Qualitative Research. *Educational Researcher, 22,* 16–23.
https://doi.org/10.3102/0013189X022004016

Gablasova, D. (2014). Issues in the assessment of bilingually educated students: expressing subject knowledge through L1 and L2. *The Language Learning Journal, 42,* 151–164.
https://doi.org/10.1080/09571736.2014.891396

Heil, W. (2012). *Kompetenzorientierter Geschichtsunterricht [competency-based history education]* (2nd edition). Stuttgart: Kohlhammer.

Körber, A., Schreiber, W., & Schöner, A. (Eds.). (2007). *Kompetenzen historischen Denkens: ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik [Competences of historical thinking: A structural model contributing to competency-based history didactics]*. Neuried: ars una.

Kuckartz, U. (2016). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung [qualitative content analyis: methods, practice and computer assistance]* (3rd, revised ed). Weinheim Basel: Beltz Juventa.

Llinares, A., Morton, T., & Whittaker, R. (2012). *The roles of language in CLIL. Cambridge language teaching library.* Cambridge, New York: Cambridge University Press.

Maset, M. (2015). *Bilingualer Geschichtsunterricht: Didaktik und Praxis [bilingual history education: didactics and practice]* (1st edition). *Geschichte im Unterricht.* Stuttgart: Verlag W. Kohlhammer.

McKenney, S., Nieveen, N., & van den Akker, J. (2006). Design research from a curriculum perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 67–90). Hoboken: Taylor & Francis Ltd.
https://doi.org/10.1007/bf02504982

McKenney, S. E., & Reeves, T. C. (2012). *Conducting educational design research.* New York: Routledge.

McKenney, S. E., & Reeves, T. C. (2014). Methods of evaluation and reflection in design research. In D. Euler & P. F. E. Sloane (Eds.), *Zeitschrift für Berufs- und Wirtschaftspädagogik Beiheft. Design-based research* (pp. 141–155). Stuttgart: Franz Steiner Verlag.

Meyer, O., Coyle, D., Halbach, A., Schuck, K., & Ting, T. (2015). A pluriliteracies approach to content and language integrated learning – mapping learner progressions in knowledge construction and meaning-making. *Language, Culture and Curriculum, 28,* 41–57.
https://doi.org/10.1080/07908318.2014.1000924

Morton, T. (2013). Critically Evaluating Materials for CLIL: Practitioners' Practices and Perspectives. In J. Gray (Ed.), *Critical Perspectives on Language Teaching Materials* (pp. 111–136). London: Palgrave Macmillan UK. Retrieved from *http://link.springer.com/10.1057/9781137384263_6*

Nikula, T., Dafouz, E., Moore, P., & Smit, U. (Eds.). (2016). *Bilingual education and bilingualism. Conceptualising integration in CLIL and multilingual education.* Bristol, Buffalo: Multilingual Matters.

https://doi.org/10.21832/9781783096145

Pérez Cañado, M. L. (2016). Are teachers ready for CLIL? Evidence from a European study. *European Journal of Teacher Education, 39,* 202–221.
https://doi.org/10.1080/02619768.2016.1138104

Pérez Cañado, M. L. (2018). CLIL and pedagogical innovation: Fact or fiction? *International Journal of Applied Linguistics, 28*, 369–390.
https://doi.org/10.1111/ijal.12208

Van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, N. Nieveen, T. Plomp, K. Gustafson, & R. M. Branch (Eds.), *Design Approaches and Tools in Education and Training* (pp. 1–14). Dordrecht: Springer Netherlands. Retrieved from *http://nbn-resolving.de/urn:nbn:de:1111-201108211273 https://doi.org/10.1007/978-94-011-4255-7*

Van Drie, J., & van Boxtel, C. (2008). Historical Reasoning: Towards a Framework for Analyzing Students' Reasoning about the Past. *Educational Psychology Review, 20*, 87–110.
https://doi.org/10.1007/s10648-007-9056-1

Appendix

[Topic]

[Historical source/s]

[reference]

Look at this [source type] and do the task below. Argue with what you see on the picture and what you know from the lessons. Please answer in full English sentences. If you don't know an English word, just put the German translation in [brackets].

1. [reproduction (I): introduction] Describe the picture. What do you see? (DESCRIBE, *de-con*)

2. [transfer (II): analysis - motivation/intention]
   Explain why the artist might have produced this [source/type]. (EXPLAIN-episode, EXPLORE-basic *de-con*)?

3. [reorganisation (II); analysis - concept]
   Explain how (if at all) this source is connected to [concept]. (EVALUATE-episode, REPORT-basic, DESCRIBE-basic, EXPLAIN-basic, *de-con*)

4. [problem-solving (III): source evaluation]
   Considering what you know about [topic], do you think the sources are reliable? Give reasons for your evaluation. (EVALUATE-episode, REPORT-basic, DESCRIBE-basic, *de-con*)

5. [reflection (III): discussion - relevance]
   Argue whether (or in which ways) *this [source type]*/[issue depicted] is still relevant in the 21st century. (EVALUATE-episode, REPORT-basic, CATEGORIZE-basic, (DESCRIBE-basic), *orientation*)

Figure 4. Written task template

		Level 0	Level 1	Level 2	Level 3	individual levels	Levels of sub-categories	overall level
Use of CDF-Types	Choice of CDF-Types	no/hardly any CDF-Types are target CDF-Types	some CDF-Types are target CDF-Types	most CDF-Types are target CDF-Types	all epidsode and most basic CDF-Types are target CDF-Types			
	Composition of CDF-Types	composition of CDF-Types is illogical/ unclear	composition of CDF-Types is partly logical/ clear	composition of CDF-Types is generally logical/ clear	composition of CDF-Types is logical/ clear throughout			

Figure 5. Sample of CDF-based assessment rubric

stage 3	stage 2	stage 1	stage 0
the source is deconstructed; including direct reference to the source and its content	the source is deconstructed, including indirect reference to the concepts displayed in the source	deconstruction of the source is not visible, but related historical concepts are included (yet without reference to the picture itself)	there is no reference to the source/ no deconstruction of the source
the connection between source and answer is fully comprehensible/ justified	the connection between source and answer is tangible/ justified with some effort (i.e. the answer is very comprehensible/ reasonable but the justification is vague or superficial)	the connection between source and answer is comprehensible/ justified with great effort (i.e. the answer is somewhat comprehensible/reasonable but there is no (explicit) justification)	there is no (justified) connection between source and answer/ no deconstruction of the source
great amount of detail, all parts covered	most parts covered, sufficient amount of details	some parts covered, supported by some details; or a good amount of details but central point is not included	no details, substantial parts missing

Figure 6. Sample of the history-related assessment rubric: deconstruction competence