Lost in Latin translation

Teaching students to produce coherent target texts

Luger, S.

[Link to publication](#)

LOST IN LATIN TRANSLATION

Teaching Students to Produce Coherent Target Texts

Suzanne Luger

LOST IN LATIN TRANSLATION

Suzanne Luger

LOST IN
LATIN TRANSLATION

TEACHING STUDENTS
TO PRODUCE
COHERENT TARGET TEXTS

ISBN: 978-94-028-1901-4
NUR: 846

Cover design by Monique van Hootegem
Printed by Ipskamp, Enschede

Lost in Latin Translation

Teaching Students to Produce Coherent Target Texts

ACADEMISCH PROEFSCHRIFT

Ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Aula der Universiteit

op vrijdag 7 februari 2020, te 13.00 uur

door

Suzanne Luger

geboren te Amsterdam

Promotiecommissie:

Promotores:	Prof.dr. C.H.M. Kroon	Universiteit van Amsterdam
	Prof.dr. G.C.W. Rijlaarsdam	Universiteit van Amsterdam
Copromotor:	Dr. S.M. Adema	Universiteit van Amsterdam
Overige leden:	Dr. I. Schrijver	Universiteit van Antwerpen
	Prof.dr. J.J. Beishuizen	Vrije Universiteit
	Dr. P.S. Gerbrandy	Universiteit van Amsterdam
	Prof.dr. I.J.F. de Jong	Universiteit van Amsterdam
	Prof.dr. C.A.M. van Boxtel	Universiteit van Amsterdam

Faculteit der Geesteswetenschappen

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
TRANSLATING LATIN IN CLASS

Key Concepts and Problem Analysis

1   GENERAL INTRODUCTION

Translating Latin into Dutch is considered one of the cornerstones of Latin secondary education in the Netherlands (Kroon & Sluiter, 2010). Since the early 1970's[1] this cornerstone has become more and more of a stumbling block because of the poor quality of the target texts Dutch students produce. This constitutes a problem, as the aim of the translation assignment is for students to demonstrate their comprehension of the source text by translating it. This aim is reflected in the programme for the Central Final Examination in Latin:

> Domain A: Reflection on Classical texts.
>
> Sub-domain 1: The candidate can demonstrate his/her comprehension of Latin and translated Classical texts by:
> - translating an unseen passage;
> - analysing and interpreting a passage from a linguistic, literary and cultural-historical perspective;
> - comparing a passage from a linguistic, literary and cultural-historic perspective with other cultural utterances from Antiquity or later periods.[2] (translated by SL).

The equivalence of source text and target text in students' translations is often inadequate and the target text itself is usually incomprehensible and incoherent (Kroon & Sluiter, 2010; Luger, 2015; Raap, 1992). To offer some background as to what I mean by a coherent and an incoherent target text, I present the translation of a fable (Phaedrus I,16) by two 10th grade[3] students in *Table 1*. Text A offers an example of an incoherent target text, text B of a more coherent text.

Some students show the ability to produce a target text that can be understood without the aid of the source text, indicating it is not an impossible task.

---

[1] *For an exhaustive overview of the discussion see http://www.stilus.nl/examen/#10 (Dutch) and Derix, Felten, Van Gessel, Van Mourik, Tielens, and Verhoeven (1983:46).*
[2] *www. examenblad.nl: examenprogramma Latijnse taal en cultuur.*
[3] *5th (pre-final) year of the pre-academic track of secondary education in the Netherlands.*

What do these students do right? Can what they do be taught to other students who are struggling?

Table 1: Examples of Incoherent and Coherent Translations

A. Example of an incoherent translation by a student [English translation below by SL]	B. Example of a coherent translation by a student [English translation below by SL]
Een schaap, een hert en een wolf Een oplichter roept schurken op om borg te staan om, niet de zaak op te lossen, maar slechts verlangt hij te beginnen. Het Hert vroeg een schepel van een schaap, iemand die borg staat aan een wolf. Maar zij, vrezend voor een list, Altijd is de wolf gewend te roven en daarna weg te gaan. Om uit jouw zicht te vluchten in volle vaart, Waar zal ik jullie terugvinden, wanneer zal de dag zijn gekomen.	Een schaap, een hert en een wolf Wanneer een oplichter schurken oproept om borg te staan, verlangt hij niet om de zaak op de lossen, maar om het slechte te beginnen. Ooit vroeg een hert een schaap om een schepel tarwe, met de wolf als degene die borg stond, Maar zij, omdat ze vreesde voor een list, (zei):'De wolf is gewend om altijd te roven en weg te gaan; jij om in volle vaart uit het zicht te vluchten. Waar zal ik jullie terugvinden, als de dag zal zijn gekomen?'
A sheep, a deer and a wolf A fraud summons villains to warrant to, not solve the case, but only longs he to start. The Deer asked a bushel from a sheep, someone who warrants for a wolf. But she, fearing for a ruse, The wolf is always accustomed to plunder and to leave after that. To flee from your sight, where will I find you, when will the day have come?	A sheep, a deer and a wolf When a fraud summons villains to warrant, he does not wish to solve the case, but to start evil. Once a deer asked a sheep for a bushel of wheat, with the wolf as a warrant, But she, because she feared for a ruse, (said): 'The wolf is used to plunder always and then leave; you are used to flee from sight at full speed. Where will I find you, when the day has come?'

The research question this dissertation addresses is how to improve the target-text coherence in texts upper-secondary students produce when translating Latin. I intended to design lessons that will enable students to produce more coherent texts when they translate Latin into Dutch and test these lessons in an experimental study. The road towards the design and realisation of an experimental study which addresses this problem led to sub-questions that I will discuss in chapters two to five. First, I wanted to know how coherence of a target text could be assessed reliably (Chapter 2). Secondly, I wondered what translation activities those students use who achieve coherent target-text production (Chapter 3). After establishing the translation activities of successful student translators, I needed to operationalise these activities and design lessons teaching those activities to other students (Chapter 4). I wanted to test these lessons before performing the experimental study, so that I could improve the lessons if necessary (Chapter 5). All these chapters lead up to the

report on the main experiment and its effects in chapter six. Chapter seven summarises the main findings of this dissertation. It also discusses the validity of the experiment, the implications for classroom practice and the future of the translation assignment as I picture it.

This dissertation combines the humanities and the educational sciences, which raises some readability issues, considering the diverse backgrounds of its intended readers. I will try to accommodate all. As I was working on a crossroads of three disciplines, translation studies, Latin teaching practice and educational sciences, it was necessary to define the concepts I used from those disciplines in this introductory chapter. Therefore, the present chapter provides both the theoretical framework for this dissertation, discussing key concepts borrowed from each field, and the problem analysis of students seeming unable to produce coherent target texts.

Following this introductory section (Section 1), section two will briefly present the state of affairs in translation studies. It provides definitions of the terms *translating, translation assignment*, *translation competence*, the *translation process*, various *types of target texts,* and *target-text coherence.* Translation studies as a field is primarily concerned with the activities and competences of professional translators, as opposed to those of the student translators who are the focus of this dissertation.

The third section applies the concepts from translation studies to the specific Latin classroom-context: what translation assignment, target text producer, target text reader, and source texts do we deal with in Latin class? This section also describes the practice of teaching Latin translation in the Dutch classroom providing important elements of the problem analysis, as the current translation assignment and instruction method may not stimulate students to translate a Latin source text into a coherent target text.

Section four looks into educational sciences to establish key concepts for the design of the experiment. This section covers the basic concepts learning aims and learning activities, as well as models for learning and instruction of complex skills and how these models may be applied in the teaching of Latin translation.

2   INTRODUCTION IN TRANSLATION STUDIES

The main elements of translation studies are 1) the source text, 2) the translation assignment,[4] 3) the translator/target-text producer, 4) the target text and 5) the target-text reader. In translation studies translator and target-text reader are never identical, as the act of translation would be superfluous if the target-

---

[4] *'Translation assignment' is the term used in translation studies to identify the purpose, function and audience of a translation.*

text reader could understand the source text in its original form. The following three subsections offer a brief introduction into translation studies and are based on several influential publications. For the description of the historical background of translation studies (2.1) the publications by Nord (1997) and Pym (2003) were valuable. The subsection concerning the translation process (2.2.) relies on Göpferich, who provided an overview of the study of the translation process in *Translationprozessforschung, Stand, Methoden, Perspektiven* (2008) as well as on Breedveld (2002), who researched the translation process in time. Through the years, focus in translation studies seems to have gradually moved from source text to target text and simultaneously from translation production to translation process. Section 2.3 discusses the focus-shift from source text to target text, which has led to an increased emphasis on *target-text* production (as opposed to *translation* production) and consequently on writing skills in the target language, as Schrijver argued in her dissertation *The translator as a text producer* (2014).

2.1 Historical Background of Translation Studies

Translation studies developed as an academic field in the 1970s. Earlier, translating was mostly seen as a simple result of proficiency in two languages and as such a purely linguistic endeavour (Pym 2003:483). In the seventies, the German philologist Wilss, one of the founding fathers of translation studies as an independent field, added "understanding of the content and style" of the source-text to the definition of translation, saying the following: "Translation leads from source-language text to a target-language text which is as close an equivalent as possible and presupposes an understanding of the content and style of the original" (Wilss, 1977:70 quoted by Nord, 1997:7). Although in the second part of his definition of translation Wilss explicitly included understanding of the content and style of the original, the effect was small as translation studies remained predominantly focused on the first part of this definition: the equivalence of source text (ST) and target text (TT). In refining the term 'equivalence,' a distinction was made between *formal equivalence*, i.e. faithfully rendering all formal elements of the ST providing a literal, word for word translation, and *dynamic equivalence*, i.e. aiming at equivalence in extra-linguistic characteristics such as the communicative intentions of the *author* (Nord, 1997:4-5). This communication was still regarded as a two-way street from source text (writer) to target text (producer), while the target-text *reader* was left out. Even if we include Wilss' "understanding of content and style" in the definition of translation, the target-text *reader* has no part in the definition. In other words, a *communicative* function of translation was still not in focus.

During the 1970s and into the 1980s, the equivalence-based approach to translation was gradually replaced by a more functional view, represented by the 'German School' of Vermeer, Reiss and Holz-Mänttäri. The German school viewed translation as more than a linguistic act. Translation came to be perceived as a form of communication, in which the cultural domains of both source language and target language had to be taken into account. Translating became a form of intercultural communication, the translator a mediator: "We see that this approach essentially involves viewing translation as an intentional, interpersonal, partly verbal intercultural interaction based on a source text" (Nord:18).

Consequently, the focus moved from source text to target text, and the cultural aspect of translating became leading. To distinguish what translators do to gain understanding of the source text and what they do to write a target text Holz-Mänttäri (1984) distinguished *translational action* (i.e. what translators do) and *translation* (i.e. target-text production). This marks an important development, as a translator's mental activities and text production were considered to be virtually the same up until then.

Inspired by the functional approach Reiss and Vermeer defined translating as a purpose-oriented activity, developing the *skopos-theory* as a theoretical framework for translation studies (Reiss & Vermeer, 1984). According to this theory the target-text situation was seen as more important than the source text: "The prime principle determining any translation process is the purpose (*skopos*) of the overall translational action" (Nord, 1997:27). A translation, according to Reiss and Vermeer (1984), can be much more than a parallel rendition of ST in TT: the *skopos*, i.e. function and audience as defined by the translation assignment, defines the method of translation. A clarifying example is given by Christiane Nord in her introduction to *Translating as a purposeful activity* (1997): a tourist in Indonesia is lost and asks a police officer for directions through a local interpreter. The officer's answer takes a long time and eventually the interpreter translates and says: 'he doesn't know.' In the specific cultural context where admitting that you *don't know* something means losing face, the long and winding answer implies that the officer does not know and may even be giving wrong directions. Therefore, the translation by the interpreter, who correctly assumes that the officer does not know, is a correct rendition of the source text, according to the *skopos*-theory. Whether it can be considered a correct translation depends on the paradigm: in the functionalistic paradigm it can, in the linguistic paradigm it cannot.

We can say that the *translation assignment* determines the *skopos* of the translation and the *skopos* determines the required translation method. However, not all translation assignments are as clear cut as the example of the Indonesian interpreter. More often translators deal with a variety of *skopoi* and

face the task of arguing the hierarchy of their choice. *"Skopos* may require a 'free' or a 'faithful' translation, or anything between these two extremes, depending on the purpose for which the translation is needed" (Nord, 1997:29). Nord points out that the *skopos* could also be a literal or even word for word translation, if this is what the translation assignment entails. In addition to the term *skopos* Vermeer used other words to describe the aim of translation, such as purpose, aim, function and intention (Vermeer, 1990:93ff). This led to a certain amount of confusion in the field, and in an attempt to create some order Nord (1997) proposed, in addition to *skopos,* a distinction between *intention* and *function* of the translation. Thus she distinguishes what the two parties on both sides of the translator (mediator) want to achieve: the source text writer (the sender) *intends* something and the text has a *function* to be established for the target text reader (receiver). Ideally the intention and function converge. However, when sender/source text writer and receiver/target text reader have very different cultural backgrounds - e.g. a time gap of 2000 years - *intention* and *function* may widely diverge. Translators have to find their way in this complexity of concepts concerning the *skopos, intention* and *function* of the translation.

The development in translation studies from considering translation a simple linguistic one-way street to conceptualising it as a complex activity concerning sender/writer, translator/target-text writer, target-text reader/receiver, taking into consideration the *skopos* of the translation assignment, *intention* and *function* of the translation, led to the question *how* translators work and *how* a translation comes into being. A logical consequence of this development is the shift in translation studies to research of translation competence, translation process, and target-text production.

2.2 *Translation Competence and Translation Process*

Translation process research aims to develop a model representing the translation process. Following the old[5] assumption that translation is a purely linguistic act involving proficiency in two languages, the translation process can simply be described as a process of decoding and encoding (*Figure 1*, based on Pym, 2003[6]). This model clearly oversimplifies what goes on in a translator's mind.

As translation involves knowledge of at least two languages and two cultural domains, as well as a number of competences in other fields, the study

---

[5] *Before the 1970s, see 2.1.*
[6] *This simplified model is also used in Van Amerongen, Budé, Derix, Van Gessel, Goossens, Mak, (…) and Zwart (1986).*

of the translation process has been undertaken from various angles: sociolinguistics, second language acquisition and cognitive psychology.

Figure 1: Simplified Translation Process[7].
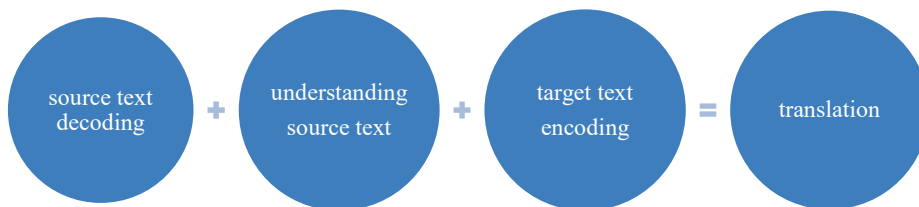
Depending on the chosen angle, different research methods have been used in the search for a model of translation processes. Göpferich (2008; Chapter 3) provides an overview of these methods. Initially, translation process studies were mostly theory-based, as the tools for a more empirical approach were absent. Research on the translation process requires observations of what translators do, which is a difficult endeavour as the translation process happens largely inside the translator's mind. In the 1980's think-aloud protocols became accepted as a research method in empirical studies in cognitive psychology.[8] This acceptance paved the way for a more empirical approach to studying the translation process. In addition to these think-aloud protocols, technical developments - such as the use of computers - further advanced this type of research, as translation behaviour could be monitored through keystroke logging and screencast software. Empirical research into the translation process was boosted recently by another technical development, the eyetracker, which provides detailed data of the translator's eye movements, such as the attention switches between the source text and the target text. The eyetracker was used in earlier studies concerning reading behaviour, and results regarding the meaning of fixations and eye movements from these studies can apply to translation studies as well (Just & Carpenter, 1980; Krings, 2005; Göpferich, Jacobsen & Mees, 2009; Hvelplund, 2014).

Theoretical and empirical research on the translation process has resulted in models for translation processes as well as translation competence. Translation competence looks at *what* translators do while translating, translation process regards *when* translators do it. Translation competence and process are closely related, as part of the competence of translating is to know *when* to do *what*.

---

[7] *Based on "1+1 Concept" Described in Pym 2003:483.*
[8] *Ericsson and Simon (1984), Protocol Analysis: verbal reports as data (Göpferich 2008: 4).*

The accepted view on translation competence today is that it involves several sub-competences, although the number of identified sub-competences varies. Göpferich (2008) discusses various models that aim at showing what translation competence consists of and what sub-competences translators use. She appreciates two models developed by the PACTE research group in Barcelona, as these models aim at clarifying the relation of sub-competences to each other as opposed to merely listing the various sub-competences. Göpferich also developed a model herself, based on these two models. She used this model in TransComp, a longitudinal study of the development of translation competence (Göpferich, 2008; 2009). In the present section I discuss Göpferich's model at some length, as it is the most complete model I have found. The translation competence model as schematised by Göpferich is shown in *Figure 2*.

Three elements, the dark grey rectangles at the bottom, are the foundation of Göpferich's model as they direct the use of the sub-competences:

1. *Translation norms / assignment*, determining the *skopos* of the translation as discussed in 2.1.
2. *Translator's self-concept / professional ethos*, depending on the training the translator received.
3. *Psycho-physical disposition*, depending on the translator's ambition, stamina, and self-awareness.

These three elements influence the translator's use of the strategic competence, which guides the translation process, moving from one sub-competence to another, as indicated by the arrows in the model. The effective use of strategic competence is further influenced by motivation, also situated in the model's core. Motivation can be either intrinsic, e.g. the translator enjoys the act of translating, or extrinsic, e.g. the translator is translating for a grade or other type of reward.The five sub-competences Göpferich distinguishes are:

- *Communicative competence in at least two languages*. Lexical, grammatical and pragmatic knowledge of both source language and target language, including knowledge of both cultural domains (cf. Pym, 2003). While the source language competence can be mostly receptive,[9] the target language competence must be productive.[10] In Göpferich's model, target-text writing is not considered a competence in itself, it is merely a part of *Communicative competence in at least two languages*.

---

[9] *The competence to speak or write the source language is not necessary to understand the source text.*

[10] *The competence to actively write the target language is necessary to write the target text.*

Figure 2: Translation Competence[11].
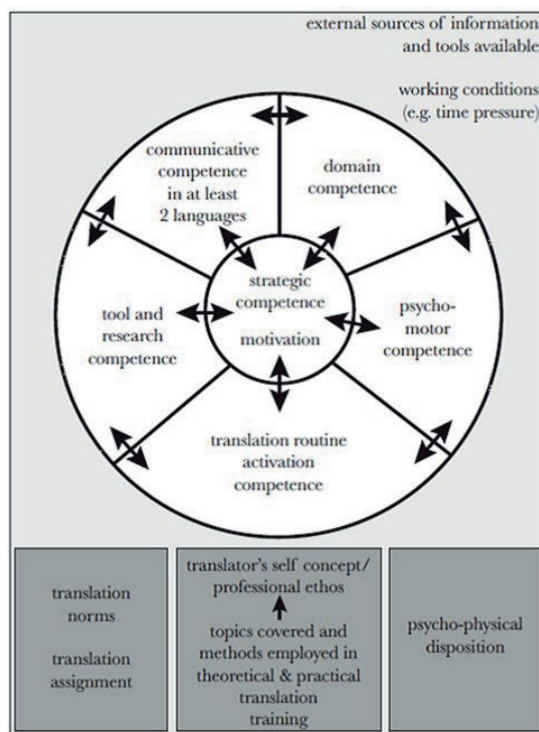
- *Domain competence.* By this Göpferich means knowledge of the world beyond the language and cultural world of the source and target languages. This knowledge is used to understand the source text in a broader sense and to formulate the target text. It also entails the capacity to determine when other sources of information need to be consulted to fill knowledge gaps. For instance, a translator who is translating a text on the subject of statistics may lack the required expertise and decide to consult statistics experts.
- *Tool and research competence.* The knowledge and skills of using dictionaries, electronic appliances, encyclopaedias, data bases or machine translation to solve translation problems.
- *Translation routine activation competence.* The knowledge of generally accepted translations for frequent problems and the skill to activate this knowledge.

---

[11] *Göpferich, 2008.*

- *Psycho-motor competence.* This refers to psycho-motoric skills necessary to read and write. The more developed these skills are, the less they occupy the translator's working memory.

Göpferich's model reflects the translation competence, stressing the importance of the strategic sub-competence to guide the decisions to switch from one sub-competence to another, but does not indicate when these switches occur. The model does not provide a timeframe or sequence for translation activities.

The translation *process* can be seen as translation competence with the added dimension of time. In research on the translation process, the focus moves from the competence to the activities translators perform and the chronological order in which they perform them. Breedveld (2002) studied the translation process of five translators in time using think-aloud protocols and found that the translation process consists of four phases: orientation, writing target text (first draft), writing target text (second draft), and revision of target text. She distinguished six translation activities in the translation process: 1) read source text, 2) read target text, 3) formulate,[12] 4) write, 5) evaluate source text/target text, 6) evaluate target text /target text. The translator performs ST-TT evaluation to assess ST-TT equivalence and intertextual coherence, and TT-TT evaluation to assess target-text fluency and intratextual coherence. Additionally, Breedveld observed that the frequency of these activities differs according to the moment in the translation process of translator T[13] (*Table 2*).

Table 2: Activities during Runs-Through in Translator T

Run	Read ST	Read TT	Formu-late	Write	Evaluate ST-TT	Evaluate TT/TT	Oth-ers	N
I	25.0	0.0	62.5	0.0	0.0	0.0	12.5	56
II	4.5	20.6	23.4	20.8	0.8	6.1	23.8	884
III	1.8	36.0	24.6	9.2	2.2	6.6	19.5	272
IV	4.8	47.6	9.5	11.9	0.8	7.9	17.5	126
Over-all	4.9	25.4	24.0	16.7	1.1	6.1	21.8	1338

Adapted from Breedveld, 2002:228; Percentage of Total Occurrences per Run

The first phase, also named run-through, concerns mostly the source text. In the second run-through, *formulating, writing,* and *reading TT* are dominant, indicating that this phase concerns target-text production. In the third phase, the frequency of *TT reading* and *formulating* increase, *writing* decreases and

---

[12] *By which she means mentally formulating the meaning of the source text.*
[13] *The five translators differed in the number of run-throughs and frequency of activities. The division of the process in orientation, text production and revision phases is present in all translators.*

in the fourth and final phase, *formulating* drops while *TT reading* and *writing* increase again. This shows a general pattern of attention movement from source text to target text.

In chapter three of this dissertation I will rely on Göpferich's translation-competence model (2008) and the distribution of translation activities across the translation process as described by Breedveld (2002) as a frame of reference to study the translation competence and process of proficient students.

2.3 From Source Text to Target Text

The attention shift from source text to target text in the translation process observed by Breedveld, suggests we take a closer look at target-text production. As target-text production aims at producing a coherent text in the target language, we need to look into how target-text coherence is conceptualised. Nord (1997) quotes Reiss and Vermeer (1984) on the coherence of the produced target text, saying that in target-text coherence the *skopos*-theory[14] is leading:

> Intertextual coherence is considered subordinate to intratextual coherence and both are subordinate to the *skopos*-rule (Nord, 1997: 32/33; Reiss & Vermeer, 1984:139).

That is, to fulfil the *skopos* the target text must be a coherent text by itself (intratextual coherence) while maintaining coherence between the target and source text (intertextual coherence). Moreover, a target text should always have meaning in the target-culture, as supported by the following quote:

> (..) what a translator can do, and should do, is to produce a text that is at least likely to be meaningful to target-culture receivers. In Vermeer's terms: the target text should conform to the standard of 'intratextual coherence' (Nord, 1997:32 quoting Reiss & Vermeer 1984:109ff)

Target-text writing is the core of Schrijver's dissertation *The Translator as a Text Producer, The Effects of Writing Training on Transediting and Translation Performance* (2014). Schrijver does not provide a definition of text coherence as such. However, she describes coherent text as the aim of writing and translating: '(…) writing and translation are both processes aimed at producing a coherent written text for a target audience (…)' (Schrijver, 2014:12).

In this dissertation I use the minimalist yet functional definition of (target)text coherence that a coherent text is understandable and meaningful in itself. This means a coherent target text is comprehensible to the reader

---

[14] Skopos-*theory: function and audience as defined by the translation assignment defines the method of translation (see 1.2.1).*

without prior knowledge of the source text. Obviously, it should also convey the main message of the source text to qualify as its translation.

How is a target text that answers to this minimalist definition of coherence produced? Target-text writing is not specifically addressed by Göpferich's model of translation competence, although it seems to be an integral part of the observable activities in the translation process. Schrijver comments on the position of writing in translation-competence models such as Göpferich's:

> When linguistic knowledge is addressed in translation-competence models, the distinction between receptive competence in the SL and productive competence in the TL usually remains implicit (2014:23).

Schrijver follows earlier research by distinguishing three basic sub-competences in translating, similar to Göpferich's first three sub-competences and also used in other and earlier translation process studies (cf. Pym, 2003) .

1. *Linguistic sub-competence,* the sub-competence concerning grasp of language. As such, it is related to Göpferich's *communicative competence in at least two languages*.
2. *Extra-linguistic sub-competence,* which is similar to Göpferich's *domain competence*.
3. *Instrumental sub-competence,* which is similar to the *tool and research competence* in Göpferich's model.

To these three sub-competences, Schrijver added the aspect that Göpferich omitted: the competence of writing a (target) text: "this text-productive competence of translators is not explicitly addressed, let alone defined, in the literature in Translation Studies" (2014:43). She aims at filling this gap and values this sub-competence as a central element in translation. Schrijver[15] suggested a flower shaped model for translation tasks (*Figure 3*). This model is a Venn diagram, with writing at the heart of the action, always partly overlapping with other sub-competences.

The aim of this dissertation is to find a way to teach upper secondary students to improve target-text coherence when they translate Latin. Here, the translation task flower model could be of great use to me, as it emphasises the importance of target-text writing and editing. The model focusses on target-text production, to the extent that translating becomes a type of target-text writing assignment, based on a specific source text. In such a writing assignment, target-text revision is part of the translation process and various types of editing and revision (transediting, self-revision, revision and post-editing) form petals of the flower of translation tasks (*Figure 3*). Schrijver states "If we conceptualise translation and writing as forms of text production, it is likely that the

---

[15] *In personal communication through e-mail (2014).*

writing competence necessary for translation will be used for that aspect of translation that is most similar to writing: the production of a written text." (Schrijver, 2014:12). Studies of revision in the writing process support the importance of revision for writing (Van den Bergh, Rijlaarsdam & Breetvelt, 1994; Rijlaarsdam, Couzijn, & Van den Bergh, 2004). I argue that translation itself is a process of target-text editing and revising.

Figure 3: Translation Task Component: The Flower Model[16].
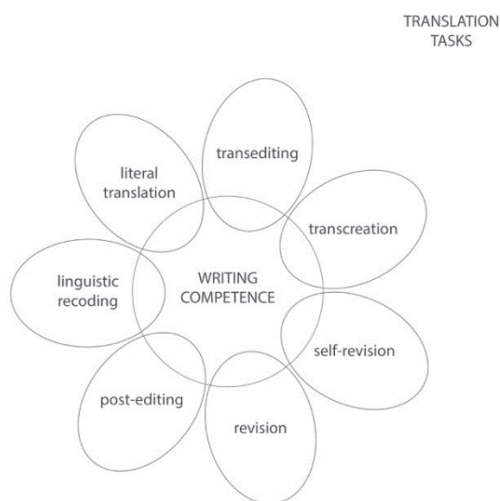
Schrijver's concept of translating compels me to have a further look into research on revision in translation. In addition to the translation process in time, Breedveld studied the *revision phase* in translation (Breedveld & Van den Bergh, 2002). An important starting point for this study was the observation of a correlation between time spent on revision activities and target-text quality. Translators who produced high-quality target texts spent more time on revision than those who produced low-quality target texts (Gerloff, 1988 as cited in Breedveld & Van den Bergh, 2002:330). Another relevant observation by Gerloff was that the length of text units translators processed increased with the number of run-throughs they performed. This would suggest that revision starts on word level and ends on text level. Consequently, it was expected that different revision activities would be observed related to the moment in the revision phase. Breedveld and Van den Bergh analysed the revision activities of five translators through think-aloud protocols. This study again focused on the distribution of activities over time. The observed

---

[16] *Iris Schrijver, personal communication (2014)*

translators revised the target text one or more times before they considered it finished. With Gerloff's observations in mind, Breedveld and Van den Bergh hypothesised that translators would change their revision activities each time they went through their target text. However, the five translators appeared to apply revision activities randomly during the entire process, and no specific order for the revision activities was observed.

Robert and Van Waes (2014) studied the efficiency of four specific revision procedures[17] in sixteen professional revisers to assess if any of these four procedures were significantly more efficient than the others. They did not clearly define what type of efficiency they aimed at, but it seems that they were looking for a short (time efficient) procedure that delivered high quality target texts. According to their definition of revision, it concerns *examining the target text*, *comparing the source text and the target text*, and *recommending changes* (Robert & Van Waes, 2014:306). The procedures they studied differed in number of re-readings (once/twice) and whether revision was performed re-reading the target text (monolingual), re-reading source text and target text (bilingual) or comparing source text and target text. They focused on efficiency and accuracy of the procedures, and wanted to answer three questions:

1. Do revisers need the source text to revise the translation?
2. How many times should a reviser read the translation?
3. When reading the text twice (for example once together with the source text and once without) does the order in which this is performed matter?

They studied four revision procedures:

1. single monolingual re-reading (referred to as the 'M-procedure'),
2. single bilingual re-reading (referred to as the 'B-procedure'),
3. comparative re-reading followed by a monolingual re-reading (referred to as the 'BM-procedure'), and
4. monolingual re-reading followed by a comparative re-reading (referred to as the 'MB-procedure') (Robert & Van Waes, 2014: 306-307).

They distinguished four types of revision setting: a *loyal* setting, taking only content and transfer into account; a *functional* setting, taking language and readability into account; a *minimal* setting, taking only grammar, spelling and transfer into account, or a *full revision*, which is a setting combining all of the above (Robert & Van Waes, 2014: 308). The sixteen professional revisers were asked to revise four similar target texts, using one of the four revision procedures in various revision briefs each time. Robert & Van Waes used a product-based revision analysis and process-based analyses through keystroke logging and think-aloud protocols to analyse the data. They concluded that the

---

[17] *(1) Monolingual, (2) Bilingual, (3) Bilingual followed by Monolingual, and (4) Monolingualfollowed by Bilingual.*

answers to the three questions depend on the setting. For this dissertation, I was interested in the most efficient procedure in a *functional setting,* which aims at improving the readability of the target text. Robert and Van Waes found that in a *functional* revision setting monolingual revision is as efficient[18] and as fast as bilingual revision, so revisers are free to choose whether they use the source text, and that monolingual revision is faster than a *two-step procedure* (BM/MB) and as efficient as bilingual revision (B). These conclusions suggest that *monolingual revision* is an efficient way to revise target texts in my project, where improving the readability of the target text is the goal.

To summarise, in this dissertation I follow Schrijver in considering translation to be a type of text production. I use a minimalist approach to text coherence: the target text is considered coherent when it is a coherent text in itself (intratextual coherence). Additionally, I adhere to the view of Breedveld (2002) and Breedveld and Van den Bergh (2002) that revision is essential in the translation process and is part of the translation process in time. In my experiment I aim to improve language and readability of students' target texts (intratextual coherence). This setting equals the *functional setting,* which suggests that a monolingual revision procedure is most efficient for the improvement of target-text readability (Robert & Van Waes, 2014).

3   INTRODUCTION TO TRANSLATING LATIN IN CLASS

Translation has always been an important part of the practice of teaching Latin in Dutch secondary education. The belief that translating a Latin source text leads to its full and complete understanding and as such is the purest form of connection with antiquity is still widespread (Kroon & Sluiter, 2009, 2010). This belief is reflected in Dutch teaching practice, where upper secondary Latin lessons are centred around 'reading'[19] source texts, as well as in Dutch testing practice, where translating a Latin source text is always part of the final exam.

The emphasis on *translating* in Latin education has not resulted in attention for developments in translation studies, which may be explained by the fact that these insights from translation studies cannot directly be applied in the context of secondary education. Translation studies focuses on the translation of source texts by translators whose proficiency in both source language and target language is fluent, while students translating Latin in upper secondary education are still developing their proficiency in both languages. The

---

[18] *Here 'efficient' seems to mean 'accurate.'*
[19] *Reading does not really describe the slow process of deciphering the Latin source text, hence the quotation marks.*

teaching of Latin translation in the Netherlands is mainly aimed at developing the students' linguistic sub-competence in the source language. In other words: teachers are teaching their students Latin, not Dutch or translation.

The complexities of the act of translating as described in Section 2 suggest that the focus on the development of the linguistic sub-competence alone may not be enough to teach students to translate Latin into coherent Dutch.

In school, the sub-competence that Schrijver (2014) calls *extra-linguistic* and Göpferich (2008) calls *domain competence* is to be developed, as well as the *instrumental* (Schrijver) or *tool and research* (Göpferich) competence, as students are allowed to use a dictionary with a morphologic appendix (Pinkster, 2018). Obviously the instrumental sub-competence could compensate for students' shortcomings in other sub-competences. However, after a brief introduction on the use of the dictionary the development of this sub-competence usually has no further role in the curriculum. Furthermore, the use of these instruments differs between beginner and more experienced translators: beginners use dictionaries and reference books mainly to solve problems in understanding the source text, whereas experienced translators use them to solve problems in target-text production (Schrijver, 2014:18). As students in upper secondary education are beginning translators, they use the dictionary only as a tool to help understand the source text.

3.1 *Historical Background of Teaching Latin Translation*

The Dutch national final exam in secondary education contains a translation task as language test. Translation is widely considered to be the cornerstone of Latin education and has become an almost sacrosanct item, as is illustrated by the uproar that followed the presentation of the preliminary report of the Exploratory Committee on Classical Languages (Kroon & Sluiter, 2009). This committee explored the problems in teaching Latin, which culminated in worrying results for the Latin final exams in the Netherlands, particularly for the translation assignment (Goris, 2002, 2007). The preliminary report suggested that the translation assignment that is part of the national final exam in Latin (and Greek) would be replaced by some other type of assignment to assess students' linguistic grasp of the source text. In the final report (Kroon & Sluiter, 2010) the committee had not dropped that suggestion. The Secretary of Education, however, possibly influenced by a strong lobby of teachers who favoured the traditional translation assignment, did not follow the suggestion: the final exam remained unchanged.

How has the translation assignment in Latin education obtained this sacrosanct status? The traditional method of teaching Latin was the grammar-translation method. This method was used in Dutch secondary school from the

beginning of Latin teaching until the late 1960s. Students learned paradigms by heart as a preparation for the translation assignment, which primarily functioned as a test for their knowledge of these paradigms. Through a thorough morphologic analysis of each word, they built meaning per word group and sentence. As preparation for the translation of full texts, themed sentences were offered. This method is best illustrated by the textbooks[20] that were used before 1968, the year that the educational reform bill[21] fundamentally changed both the structure and the content of Dutch secondary education. Before this bill, all foreign languages, both modern and classical, were taught in Dutch secondary education by means of the grammar-translation method. After 1968 however, education in the modern languages became more focused on communication. Dutch classicists tried to follow in this development by adapting the school curriculum,[22] aiming at a "confrontation of the self and the other" through reading and comprehending Latin and Greek source texts (Verhoeven,1996: Chapter 3). However, teaching practice of the classics remained primarily translation-oriented, as the desired communication with Romans and ancient Greeks through reading and comprehending authentic texts proved to be a difficult enterprise.[23]

In the 1970s insights from functional grammar and reading research started to seep into the textbooks. Themed sentences were replaced by simple Latin texts and teaching translation was approached from a more syntactic and semantic angle, as is illustrated by the method *Redde Rationem*.[24] This approach was demonstrated more recently by students of the VU Amsterdam[25] who developed the PSOLMO strategy. The acronym PSOLMO stands for the constituents in the order in which students should identify them while translating: P = Persoonsvorm (predicate), S = Samengesteld (complex sentence), O = Onderwerp (subject), L = Lijdend voorwerp (object), M = Meewerkend voorwerp (indirect object) and O = overig (other). Some problems arise when this strategy is followed too rigorously. The translator's focus remains mainly on sentence level and students may experience difficulty in understanding the source text and identifying its intratextual coherence. Moreover, when PSOLMO dictates the sequence of translating instead of the source text's word order, the sentence is ripped apart while translating, as the predicate, step one in PSOLMO, is often placed at the end of the sentence in Latin. Thus, applying the PSOLMO strategy increases the gap between translating and reading.

---

[20] *Such as Tirocinium Latinum (1955).*
[21] *Bill on secondary education, generally known as 'mammoetwet'.*
[22] *Commissie Modernisering Leerplan (1971).*
[23] *Verhoeven (1996: Chapter 6), Van Oeveren (2019).*
[24] *De Man and Te Riele (1971), Redde Rationem, Wolters Noordhoff.*
[25] *Adema, S.M., Faber, T., De Groot, R., De Jonge, F., Langedijk, I., Langerak-Wakker, M., ... Winterstein, D. (2008).*

The linear approach, competing with the PSOLMO strategy in teaching translation, addresses the difference between reading and translating. In this approach the source text is read as well as translated from left to right. This strategy requires students to read and translate more or less simultaneously, while asking the appropriate questions to identify each word in full, that is to say identifying its semantic, morphologic and syntactic meaning (Goris, 1999; Verbaal, 2015). The linear approach is the basis for the positional method developed by Verbaal (2015), which adds semantic structure: elements of each word, each word group, and each sentence are positioned in a significant order. The neutral order of each constituent is *colour | core | clarification.*[26] Authors can deliberately change the neutral order to highlight certain aspects of the text. The positional method was adapted and tested for secondary education by Hulstaert (2016), but it cannot solve our problem with translation because it is aimed at reading and understanding Latin, while translation concerns text production as well.

Teaching 'Living Latin' is a trend in Latin education (Hunt, 2018; Adema, in preparation; SLO, 2015:55). Through the Ørberg method[27] students learn to read, understand and speak Latin as a living foreign language. The Dutch final exam, however, contains a translation assignment as language test. Teaching 'living' Latin does not prepare for this type of test, although the teachers applying this method believe that translating will be easier if fluency in the source language is improved and that learning Latin as a living language improves fluency in Latin. The results of this method have not yet been subject to research, so we simply do not know if students in general improve their grasp of Latin through this method as compared to more traditional methods. Although submersion and immersion are proven methods for learning a second language, learning 'living Latin' cannot be seen as either of those. The time in the Dutch curriculum spent on teaching and learning Latin[28] is simply too short to accomplish the effects attributed to these methods.

Throughout and despite these developments in the teaching of Latin, from the grammar-translation method to PSOLMO-strategy, to reading Latin to, in some schools, Living Latin, the translation assignment has remained the dominant form of linguistic test in the Netherlands. At this point it seems important to clearly distinguish teaching *Latin* from teaching *Latin translation.* Learning a language and translating it are different enterprises, as Pym so aptly phrased:

---

[26] *For instance the noun* consuetudinem *is divided in* con | sue | tudin-em, *where the root* 'sue' *means 'habituality', prefix* 'con' *adds colour (many people do this together), the suffix –* tudin *clarifies that the word concerns an independent entity:* a habit, *and the case ending –*em *further specifies the syntactic function of the word in the sentence (Hulstaert, 2016:19).*
[27] *https://lingualatina.dk/wp/ visited on 25/10/2018.*
[28] *Generally, ca 150 minutes per week.*

A minimalist approach should ideally enable a clearer distinction between translator training and language learning. The latter should be at least to some extent analytical, rule-bound and grammar-oriented, whereas the training of translators should be relatively non-analytical, context-bound, and example-oriented. (Pym, 2003:492).

This dissertation focusses on the question of how students can learn to write more coherent target texts when they are translating Latin into Dutch.

3.2 The Translation Assignment in Latin Class

All translation starts with the translation assignment. It is to be expected that the translation assignment that students receive is instrumental for their execution of the translation task. Sicking formulated the aim of the translation assignment in Dutch education as follows: (the production of) "a text that is completely understandable and acceptable as Dutch for someone not familiar with the original text" [29] (1968:167). This formulation creates a (fictional) intended reader. However, more often students in the Netherlands are requested to translate as "literally as possible and as freely as necessary." [30] The reason for this is that the translation task is used as a test to assess students' *linguistic* knowledge of Latin in the final exams.

In Latin class, a somewhat unnatural situation occurs as the person who provides the assignment (the teacher) is more proficient in the source language than the student. The teacher does not need the translation to understand the source text, which raises the question: who is to be considered as the intended reader of the target text? In class, no authentic communication situation exists, since the teacher is usually the only reader of the target text. Therefore, students of Latin may not perceive their task as a genuine translation assignment, as a result of which it becomes hard for them to determine the *skopos* of the target text. Using the terms presented in Section 2.1, a considerable gap between *intention*[31] and *function*[32] exists in the case of translating Latin, as source language and target language are centuries apart and the cultural differences are considerable. Additionally, there is an age gap because the *sender* (the Latin author) is a highly educated adult and the *target text producer* is an adolescent with a limited understanding of the world the *sender* comes from. Moreover, the *receiver* or teacher assesses the quality of the target text by looking mainly at the linguistic equivalence of source text and target text. He

---

[29] *Translation by SL.*
[30] *Definition by Simon Veenman, as quoted by Huisman, Rijpstra, de Rooij, and Simons (2012:7) (translated by SL)*
[31] *Intention: what the source text writer (the sender) intends.*
[32] *Function: the text has a function for the target text reader (receiver).*

or she cannot really function as an authentic reader of the target text, because he or she is familiar with the content of the source text.

The problem of the poor quality of the target texts Dutch students produce, presented in the introduction, is closely related to the translation assignment students receive. Students are asked to produce target texts that are lexically and syntactically as similar to the source text as possible. When the *skopos* of the Latin classroom translation is the production of a literal, word for word translation and the gap between *intention* and *function* is large, the tendency to sacrifice the coherence of their target text students are perceived to have in their attempts at a literal translation seems understandable. Furthermore, it can be assumed that students' motivation[33] to perceive the translation assignment as a meaningful task and do a good translation job suffers through this type of presentation: what is the use of translating a source text which the intended reader (the teacher) can read and understand better than the translator?

3.3 *Source Texts and Target Texts in Latin Class*

What Latin source texts are presented to students and what target texts are produced by students in Latin class? In the Netherlands, the first three to four years of learning Latin in secondary school are centred around translating text-book stories, often adaptations of original texts by modern classicists. In the fourth year, students are expected to have mastered a basic vocabulary and to possess sufficient linguistic knowledge[34] to start reading and translating orig-inal Latin source texts with increasing complexity. The educational text edi-tions provide lexical, cultural, and some grammatical annotations to support understanding of the source text. In addition, students can use a dictionary with a morphologic supplement (Pinkster, 2018) in support of their efforts.

In upper-secondary Latin class students generally produce two types of target texts: 1) a sort of draft text for personal use and 2) a text as closely related to the source text as possible for a test. The aim of these texts differs. As a preparation for classroom discussion, students are supposed to translate or read the source text and understand its formal aspects. This may lead to the draft target text or just to notes on the source text. The teacher usually does not assess each target-text draft individually. The source text and target text are discussed in class sentence by sentence on linguistics, stylistics and

---

[33] *Motivation is an important aspect of the translation competence, situated in the heart of Göpferich's model (2.2).*

[34] *As defined in the list of* basic morphologic and syntactic knowledge of Latin *formulated by the Dutch Board of Tests and Examinations (College voor Toetsen en Examens, or CvTE). https://www.examenblad.nl/examenstof/syllabus-2020-latijn-vwo/2020/f=/latijn_ver-sie_2_vwo_2020.pdf.*

content.[35] This discussion usually plays a crucial role in the students' understanding of the source text's content.[36]

The second type of target text is used for assessment. Students receive an annotated Latin source text they have never studied before, with an introduction in Dutch on its content, and then translate the source text into Dutch using annotations and the dictionary (Pinkster, 2018). The literal instruction for the translation task of the final exam 2018 was: "Study the introduction and annotations to text 3. Translate line 1-13 into Dutch".[37] This assignment has a set time limit. In the national final exam students have approximately ninety minutes to translate a source text of approximately one hundred and twenty words. In other tests the length of the source texts depends on the time available for the test. The quality assessment of this target text is performed by the teacher[38] using an analytic scoring model. Chapter two discusses quality assessment of target texts using this analytic model and using holistic models in more detail.

3.4 The Translation Process of Latin students

The difficulties students experience when translating Latin are not limited to the Netherlands. For instance, Balbo (2011) discusses translation from Latin to Italian at high school and in introductory courses of Latin at universities and argues that "Translating is a point of arrival, not of departure. From the didactic point of view it is an activity which encounters increasing difficulties in school" (2011: 371). Balbo proposes a translation process for students consisting of nine steps[39]:

1. Reading of the introductory information and of the text
2. Consciousness raising
3. Lexical reflection
4. Examination of the significant syntactic elements
5. Syntactic analysis of sentences and clauses
6. Identification of the more significant pragmatic elements
7. Analysis of the pragmatic elements
8. Work translation
9. Revision and definitive translation.

---

[35] A written 'work' translation can be provided by the teacher before or after classroom discussion of the source text.

[36] SLO (2015), Van Oeveren (2019).

[37] CE Latijn 2018, opgavenboekje, p.6, https://static.examenblad.nl/9336118/d/ex2018/VW-1110-a-18-1-o.pdf.

[38] The assessment is checked by a teacher from another school, assigned by the government.

[39] Adapted from Balbo (2011: 387).

Although his suggestion is rooted in his own classroom practice, he has not empirically tested the effectiveness of the process he proposes, nor does he suggest a method of teaching it. Indeed, empirical research specifically focused on Latin education is scarce. Moreover, this research usually does not investigate the translation process, but examines reading and comprehending Latin as opposed to translating it. Therefore, the present section mostly discusses studies on the *reading* of Latin. These studies provide valuable insights into reading and comprehending the source text, which is where the translation process starts. However, we must bear in mind that *translating* is fundamentally different from *reading* and can best be considered as a type of source-based writing, as I discussed in Section 2.3.

I will start with an overview of the methodology and results of studies regarding the process of *reading* Latin. These studies are based on think-aloud protocols, observations of behaviour, questionnaires and even eye movements. Surprisingly, the study of eye movements and reading Latin were combined as early as the 1920s (Judd & Buswell, 1922; Buswell, 1928; Futch, 1935) when eye movements were investigated through photography. Judd and Buswell (1922) studied the eye movements of fourteen young students reading Latin prose. They concluded that the observed activities did not resemble reading: instead, students' behaviour resembled paraphrasing, they 'dwelled on words,' their eyes went back and forth. Students' behaviour generally showed confusion. Conclusions on the teaching of Latin in a manner that leads to this type of behaviour are scathing: "One can dismiss as absurd under existing conditions the idea that the ordinary pupil ever gets any literary thrills out of Latin. The whole subject has degenerated into an absolutely formal exercise in linguistic dissection." (Judd & Buswell 1922:156).

Buswell (1928) compared the reading behaviour of students in their first year of learning Latin through a direct method aimed at developing reading to that of the students of the first study (Judd & Buswell, 1922). The eye movements of the students who had learned Latin through the direct method showed a pattern resembling reading. The reading behaviour of these students was compared to the behaviour of beginning students of French reading French texts, and the patterns were generally the same. However, the reading of Latin texts was slower than that of the French texts.

Futch (1932) studied the reading behaviour of eight experienced Latin readers and twenty-seven beginning Latin readers through eye-movement photography. Easy Latin texts were presented to both groups. The experienced readers also read more complex texts, such as texts by Cicero. Analyzing fixation and regression patterns, Futch suggests four phases in learning to read a foreign language, developing from few fixations, short fixations and short regressions to patterns of variability in fixation length and finally to fixations

reduced both in length and in number and reduced regressions. The findings of these early studies on the behaviour of students when they are reading difficult, in this case Latin, texts are supported and refined by more modern eye-tracking research on reading complex texts in other languages: in beginners it is slow and can be characterised as deciphering more than as reading (Just & Carpenter, 1980; Göpferich, Jakobsen, & Mees, 2009).

Van de Walle and Van Houdt (2004) and Van Houdt (2008) studied the reading behaviour of university students in Latin through think-aloud protocols. They found that switching between top-down[40] and bottom-up[41] strategies, similar to Göpferich's switching between sub-competences, is crucial in a successful reading strategy. Although their research focused on reading and consequently did not cover target text production, it provides some insight in the reading behaviour of students at both secondary school and university level. The problems the students encountered when reading were mostly linguistic, which suggests that problems in translating may start in that same area.

In addition to these studies concerning *reading* Latin, some research into the *translation process* of secondary school students exists. Van Krieken (1981, 1982) reports on an explorative study of the translation process of Dutch students through think-aloud protocol analysis and observation. He concluded that students' activities are not as related to source language knowledge or text comprehension as expected, and observed that students translate mainly on word level. He characterised the observed translation activities as "playing Scrabble with words"[42] (1981:573).

Recently, Florian (2015) dived into the research gap of student translation and conducted an empirical study on the translation behaviour of German high school students in Latin. Through video and audio recordings, she qualitatively studied the translating behaviour of twelve students working in pairs, using think aloud protocols and questionnaires. Florian found that proficient students can successfully solve translation problems presented to them by relying on text comprehension, while their morphologic knowledge is supporting at best (Florian, 2015).

Research methods in the empirical process-studies discussed here are think-aloud protocol analysis (Van Krieken, 1981, 1982), observations (Van de Walle &Van Houdt, 2004; Van Houdt, 2008; Florian, 2015), questionnaires (Florian, 2015) and eye-movement analysis (Judd & Buswell, 1922; Buswell, 1928; Futch, 1935). Some objections to these methods are formulated by Krings (2005) and Olive (2010). They argue that students who think aloud

---

[40] *Zooming out, activating knowledge of context, genre characteristics, background.*
[41] *Zooming in on word or sentence level. Activating knowledge of morphology, syntactic structure.*
[42] *Translation by SL.*

while translating may experience cognitive overload, which interferes with the translating process itself and slows it down. Indeed, it has been proven that thinking aloud slows down processes in writing texts (Janssen, Van Waes, & Van den Bergh, 1996). Additionally, it can be difficult to verbalise automated processes, which suggests that thinking aloud studies miss out on automated translation activities. The method of observation reports observable behaviour only, so that thoughts and decision-making in participants remain unaccounted for. Questionnaires are always retrospective, as students answer them after reading or translating. An objection to this method is that answers tend to be more socially desirable than sincere. Eye-movement analysis is the least compromised method to research reading and translating processes, as it provides data on spontaneous, unfiltered behaviour, measured in real time. The thoughts that underlie the eye-movements have been subject to research and still are (Göpferich, Jakobsen & Mees, 2009; Hvelplund, 2014). Modern eye-tracking research investigates the translating process of either professional translators or university translation-students in modern languages as opposed to adolescents translating Latin. Therefore, based on this research, no firm conclusions can be drawn on the translation process of high school students translating Latin.

The study of the translation of Latin has not yet been done through modern eye-tracking techniques, neither in high school or university students nor in professional translators. Chapter three of this dissertation describes a study I conducted on translation activities performed by proficient students through eye-tracking and stimulated recall. I will discuss the additional methodology I chose for that study in Chapter three.

4   INTRODUCTION TO EDUCATIONAL SCIENCES

Educational sciences are a broad research field. In general terms, it covers the search for answers to three basic questions: 1) how people learn, 2) how people teach, and 3) what people learn. The answers to these questions vary over time and are influenced by developments in related fields, such as pedagogy, cognitive psychology and neuropsychology as well as the changing demands of society. This section does not aim at giving a full overview of these various answers, it mainly aims at providing a framework for this dissertation.

I established in Section 2 that a translation assignment can be considered a special type of writing assignment and that revision is a part of the translation process. Therefore, the present section focusses on research concerning the learning and teaching of writing and on clarifying some relevant basic concepts: learning aims, learning activities, principles of teaching (Merrill, 2002), teaching to write (Flower & Hayes, 1981; Graham & Perin, 2007),

observational learning (Case, Harris & Graham, 1992; Couzijn, 1999; Braaksma, 2002) and strategy instruction (Harris & Graham, 1996, 2002).

4.1 Learning and Instruction

4.1.1 Learning Aims and Learning Activities

Learning aims and learning activities are two basic concepts in education. A learning aim is the expected outcome of learning, and learning activities refer to the activities students must perform to reach the learning aim. In translating Latin, the learning aim seems to equal the translation assignment: as an outcome of their learning in Latin translation-class, students produce a written target text that should be closely related to the Latin source text in structure, phrasing, and meaning. Additionally, the target text should be readable and comprehensible independently from the source text. It has not yet been established what learning activities lead to attaining this aim. This gap in the instruction of Latin translation presents a serious problem, which teachers seem to have addressed by relying on tradition and personal experience. Common practice, as I discussed in Sections 3.1 and 3.2, suggests three leading ideas about learning activities in translation-teaching: 1) careful *linguistic analysis* of the source text leads to improved target texts on sentence level, 2) a systematic (PSOLMO/linear/positional) approach to the translation assignment on sentence level leads to improved target texts, 3) practice makes perfect. These ideas are based on tradition and beliefs rather than on empirical research. As discussed above, teaching Latin through a grammatical analytical approach may not lead to fluency in reading the source text (Buswell, 1928; Futch, 1932) or to producing a coherent target text (Kroon & Sluiter, 2010; Luger, 2015).

4.1.2 Instruction

Merrill (2002) has reviewed several instructional design theories with the aim of identifying common prescriptive principles for teaching. He found that all learning is based on five principles:
1. Learning is promoted when learners are engaged in solving real-world problems.
2. Learning is promoted when existing knowledge[43] is activated as a foundation for new knowledge.
3. Learning is promoted when new knowledge is demonstrated to the learner.

---

[43] *Merrill uses "the word knowledge in its broadest connotation to include both knowledge and skill, and to represent the knowledge and skill to be taught as well as the knowledge and skill acquired by the learner" (2002:45).*

4. Learning is promoted when new knowledge is applied by the learner.
5. Learning is promoted when new knowledge is integrated into the learner's world. (Merrill 2002:44-45).

Therefore, effective teaching must reflect these five principles. The aim of the current dissertation is to develop and test lessons in an experimental setting to teach students to produce more coherent target texts. In doing so, I will rely on Merrill's principles for the design of the lessons (Chapter 4) and apply them to teaching Latin translation. In the present section these principles are linked with teaching Latin translation in general.

The practice of Latin-teaching I described earlier seems strikingly similar to the type of instruction Merrill calls topic-centred: "Problem-centred instruction is contrasted with topic-centred instruction where components of the task are taught in isolation (e.g., "You won't understand this now but later it will really be important to you") before introducing the real world task to the students."(2002:45) This type of instruction is contrary to the first principle of teaching: "engaging learners in solving real-world problems." Admittedly, it may seem a stretch to call translating a Latin text a real-world problem, but the fact that the person giving the translation assignment and the intended reader are the same (i.e. the teacher) is certainly not helping. Examples of adapting the translation assignment into a real-world-problem assignment exist, e.g. by asking students to create a text for an explicit audience. In a translation project, high-school students translated texts on Alexander the Great that had never been translated into Dutch before. The target text was intended for publication.[44] In another project, students translated texts to create a book that was presented to their parents in a festive gathering.[45] The effects of these projects were not systematically researched, but students' and teachers' enthusiasm is attested.

The second principle, "activating existing knowledge as a foundation for new knowledge," is not systematically applied to the translation assignments in advanced Latin class, as the literary texts presented to students vary in genre, vocabulary and style, thus leaving little existing knowledge for activation besides grammar. It must be noted that activation of existing knowledge of the text is part of classroom practice when one specific text is discussed in a series of lessons. However, the activated existing knowledge then concerns students' knowledge of the text's narrative and cultural content rather than translation skills. A different approach to the translation assignment could provide a solution. For example, one could approach it as an assignment concerning meta-cognition and activating process- knowledge by focusing on the

---

[44] H. Stouthart (personal communication, 2012)
[45] M. Simons (personal communication, February 18, 2019)

question how to translate texts in general as opposed to the production of a translation as proof of linguistic knowledge (Section 3. 3).

Merrill's third principle, "new knowledge is demonstrated to the learner," is hardly any part of upper secondary Latin education in translation. Firstly, because *translation* is not considered to be 'new knowledge.' Secondly, Latin teachers rarely model the translating process of a *complete text.* Demonstrations are limited to micro-strategies such as 'how to translate a subjunctive in a subordinate clause,' or 'how to translate an absolute ablative construction.' Merrill labels instruction on new knowledge as "merely information followed by a few remember-what-you-were-told questions" (2002:48) as ineffective. However, this type of instruction is common in Latin translation class.

Merrill's fourth principle, "learning is promoted when new knowledge is applied by the learner" is captured in the popular 'practice makes perfect' belief: teachers believe that the more Latin texts students translate, the better they will become at translating Latin texts and they seem to assume that translating texts equals applying new knowledge. Merrill adds that learning through practice is promoted when "learners are guided in their problem solving by appropriate feedback and coaching, including error detection and correction, and when this coaching is gradually withdrawn" (2002:49). This type of guidance in the procedure of translating an entire text is not commonly observed in Dutch classroom practice.

The fifth principle, integrating new knowledge into the learner's world, is explained as follows: "learning is promoted when learners are encouraged to integrate (transfer) the new knowledge or skill into their everyday life" (2002: 50). This principle could be applied to translation when students apply principles of translation learned in Latin translation class in translating in their everyday lives. However, in classroom practice, the fifth principle is limited to translating gradually more complex Latin texts, which students will be able to tackle with the knowledge previously acquired.

The application of Merrill's five principles to teaching Latin translation shows that the practice of teaching translation rarely reaches the meta-cognitive level of teaching translation as a discipline. It more often concerns struggling with the source text's syntax and semantics.

4.2 Learning and Instruction of Writing

In Section 2.3, I followed Schrijver (2014) in arguing that translating can be considered a type of writing. Research on learning and instruction of *writing* is quite prolific, as opposed to learning and instruction of *Latin*. Rijlaarsdam, Van den Bergh, Couzijn, Janssen, Braaksma, Tillema, … Raedts (2011) present an overview of research on learning and instruction of writing. In the

early 1980s, Flower and Hayes developed an influential writing-process model that they improved in 1996. It is described by Rijlaarsdam et al. as follows:

The model contains three components: (a) the task environment, all materials external to the writer with the task at hand and the text produced so far as important elements; (b) the writer's long-term memory, with knowledge about the topic, audience, genres, and task approaches; and (c) a set of cognitive operations, globally distinguished in three categories. These three categories are planning activities (with goal setting, generating of ideas, and structuring of ideas); formulating and transcribing activities, resulting in some materialised language (text produced so far); and revising activities, containing reading already written text, evaluating, revising (intentional activity), and editing (automatically) (2011: 193).

The writing-process model of Flower and Hayes (1981) and Göpferich's translation-process model (2008) are similar in featuring *environment* and *cognitive operations*/*sub-competences* as main components. The writing and translating competence are very similar (Schrijver, Van Vaerenbergh, Leijten, & Van Waes, 2019). The similarity of both models supports the hypothesised transfer from learning and instruction of *writing* to that of *translation*. Writing, like translation, involves planning, formulating and revising. The important difference is that translating always involves a source text, from which the target text cannot diverge.

Graham and Perin (2007) performed a meta-analysis of writing-intervention research.[46] They classified the interventions into four groups: 1) process approach to writing instruction, 2) explicit teaching 3) scaffolding students' writing, and 4) more writing.[47] Based on their meta-analysis, they formulated ten recommendations concerning the instruction of writing, of which the following four are relevant for teaching the writing that is relevant in translation assignments:

1) Teach adolescents strategies for planning, revising, and editing their compositions (strategy instruction, mean weighted effect size[48] = 0.82; Grades 4–10).

(…)

4) Set clear and specific goals for what adolescents are to accomplish with their writing product. This includes identifying the purpose of the assignment (e.g., to persuade) as well as characteristics of the final

---

[46] *123 documents that yielded 154 effect sizes for quality of writing.*
[47] *This reminds us of the 'practice makes perfect' belief in Latin teaching.*
[48] *'Effect size' is a way of quantifying the size of the difference between two groups. An effect size of 0.10 is considered to be small; 0.30 is considered to be medium; 0.50 is considered to be large.*

product (e.g., addresses both sides of an argument; for setting product goals, mean weighted effect size = 0.70; Grades 4–8).

(…)

9) Engage adolescents in activities that help them gather and organise ideas for their compositions before they write a first draft. This includes activities such as gathering possible information for a paper through reading or developing a visual representation of their ideas before writing (for prewriting, mean weighted effect size = 0.32; Grades 4–9).[49]

(…)

10) Provide adolescents with good models for each type of writing that is the focus of instruction. These examples should be analyzed, and students should be encouraged to imitate the critical elements embodied in the models (for models, mean weighted effect size = 0.25; Grades 4–12).

(Graham & Perin, 2007:466-467)

I used these four discussed recommendations for the design of the lessons teaching translating Latin into coherent Dutch (Chapters 4 and 5).

Graham and Perin[50] researched the effects of explicit teaching of strategies and concluded that writing quality improved especially when the Self-Regulated Strategy Development (SRSD)[51] model was used. This model (Harris & Graham, 1996, 2002) features elements of Merrill's five principles of learning and consists of six basic stages of instruction:
1. develop and activate background knowledge
2. discuss the strategy
3. model the strategy
4. memorise the strategy
5. support the strategy
6. independent performance.

SRSD's first stage of instruction equals Merrill's second principle for teaching: activate and develop existing knowledge as a foundation for new knowledge. SRSD's second and third basic stages of instruction seem to reflect Merrill's third principle: learning is promoted when new knowledge is demonstrated to the learner. New knowledge, i.e. the strategy that is being taught, is discussed in SRSD's second stage, and modelled in the third basic stage. The modelling can be done by the teacher or by peers. The effectiveness of observational learning through peer modelling is supported by other

---

[49] *Graham and Perin note for recommendations 9) and 10) that variation in control conditions for the prewriting interventions should lead to caution in interpreting the effect-size.*
[50] *Graham and Perin, 2007:463.*
[51] *Harris and Graham, 1996, 2002.*

research (Schunk, 1995; Braaksma, 2002). In SRSD's stages four and five, the focus moves from instructor to student: the student memorises the strategy and uses the strategy with some form of support, e.g. a cheat sheet. These stages are related to Merrill's fourth principle: learning is promoted when new knowledge is applied by the learner. Graham and Harris promote using an acronym for memorising the various steps of the strategy, just as 'PSOLMO' is used for remembering the Latin translation strategy. They advocate emphasis on repeating the steps of the strategy (e.g. by using the acronym) as a result of which students will gradually remove support with growing confidence in using the strategy (scaffolding) and start to implement the strategy independently, reaching stage six: independent performance. The strategy has become integrated into the learners world, and learners will be able to use the learned skill or knowledge independently, inside and outside the classroom. This way, new knowledge is integrated into the learner's world (Merrill's fifth principle).

I used elements of the Self-Regulated Strategy Development framework as a foundation for the design of the lessons of the experimental study (Chapters 4 and 5). The Self-Regulated Strategy Development model is used and tested to teach strategies. As argued, the available translation strategies (PSOLMO strategy/ linear/ positional approach) do not meet the required focus on target-text coherence, being sentence level or reading strategies. A strategy for high school students to translate a Latin source text into a coherent Dutch target text does not yet exist. I hypothesise that a translation process using planning, formulating and revising (Flower & Hayes, 1981) will result in more coherent target texts. Therefore I will have to design a process-oriented strategy for translation focusing on target-text coherence (Chapter 4). This strategy focuses on the translation *process* instead of the translation *product.* The design of the lessons will reflect aspects of Self-Regulated Strategy Development as this is a tested and effective model to teach strategies.

5   CONCLUSIONS

The present chapter aimed at defining the theoretical framework for this dissertation as well as analysing the underlying problems students of Latin face in the production of coherent target texts and teachers face in teaching translation. It defined key concepts derived from translation studies and from learning and instruction of writing, relating those to the Latin classroom. The chapter also outlined directions for the design of an intervention that could offer some solutions.

In this chapter I have identified three main problems in the teaching practice of Latin translation which I have to address in the design of the lessons for my intervention:

1. The translation assignment itself. The *skopos* of the translation is unclear and a real-world translation assignment for Latin is hard to find.
2. Learning activities for producing a coherent as well as an equivalent target text are not defined. Knowledge of the translation process in high school students that leads to coherent target texts is lacking. Therefore, teachers do not really know what to teach their students to make them translate Latin into a coherent target text.
3. No evidence-based instruction method for teaching Latin translation is used in Latin education practice. As a base of shared knowledge is missing, teachers do not really know how to teach their students to translate Latin into a coherent target text.

In this dissertation I use the following definitions: *translating* is defined as a special type of target text production (Schrijver, 2014) and therefore is similar to writing in many ways. The *translation assignment* given to students will be 'to produce a coherent target text,' where *target-text coherence* is defined as a coherent text by itself (intratextual coherence). The target text is comprehensible to the reader without prior knowledge of the source text and conveys the main message of the source. The *translation competence of students* had not been clearly defined before this dissertation. Therefore, I will use Göpferich's model as a basis and I will investigate, through eye-tracking and stimulated recall, to what extent this model is applicable to proficient upper secondary school translators (Chapter 3). For the definition of *the translation process* I will use the four stages of the translation process in time: 1) orientation, 2) writing first draft, 3) writing second draft, 4) target text revision (Breedveld, 2002) and connect these stages with the writing-process elements *planning*, *formulating* and *revising* (Flower & Hayes, 1981).

In the design of the intervention (Chapter 4) I use Merrill's five principles for teaching (Merrill, 2002) and elements of Self-Regulated Strategy Development (Graham & Harris, 1996, 2002) for instruction of the strategy I will formulate based on the findings of chapter three.

CHAPTER 2
ASSESSING TARGET-TEXT COHERENCE

1   INTRODUCTION: TARGET-TEXT QUALITY ASSESSMENT

Chapter one described and analysed the problem of poor coherence in students' target texts, identifying three major underlying issues: 1) the translation assignment presented to students results in lack of *skopos*, 2) learning activities and a translation strategy for producing coherent target texts are not yet defined, and 3) no evidence-based instruction method for teaching coherent target text production is available. The present chapter identifies and discusses a fourth problem, which is the absence of a reliable tool to assess target-text coherence. This constitutes an additional problem as the main objective of this dissertation is to design lessons to improve target-text coherence and test the effect of these lessons. The first step in this design is to find an instrument to reliably measure target-text coherence before and after the lessons.

In current practice, the translation assignment is used to test students' linguistic skills, though teachers implicitly ask students to produce a target text that is "completely understandable and acceptable as Dutch."[52] Thus, when teachers ask students to translate a Latin source text into Dutch, the desired outcome is usually not entirely clear. The suggestion that the understandability and acceptability for a Dutch reader is part of the assignment is not reflected in the assessment, which is predominantly performed linguistically. Thus, the acceptability of the target text as Dutch is not systematically measured. This indeed would prove to be difficult, as a reliable instrument to assess the coherence[53] of the target texts that students produce has not yet been developed.

In the present chapter I discuss two types of quality assessment: analytic and holistic. *Analytic quality assessment* is currently used to grade Latin translation assignments ('colon-rating'), which is explained in section two. *Holistic quality assessment* is a method used in rating students' writing products. I explored two procedures for holistic assessment: 1) text scale rating with one or more anchor texts as frame of reference and 2) comparative assessment in pairs of target texts. Having performed exploratory studies on both types of assessment, I report on the testing of these models in sections three (scale rating with anchor texts) and four (comparative assessment in pairs). Holistic assessment is new in assessing Latin translation products. Determining the

---

[52] *Sicking (1968).*
[53] *I define target-text coherence as intratextual coherence, see Chapter 1.*

usefulness of both procedures for assessing target-text coherence is the main focus of this chapter.

Holistic assessment of target-text quality is difficult, as text quality is multi-dimensional and its assessment may very well be influenced by personal and therefore subjective criteria applied by the rater. Assessing linguistic equivalence of source text and target text seems objective, measurable and therefore safe, while assessing target-text coherence seems subjective and therefore questionable. Is it even possible to reliably assess target-text coherence? Leaving the relative safety of viewing the translation assignment as a test of linguistic equivalence, translating becomes the type of open-ended task of which Messick (1994) observes: "Evaluations of performance on such open-ended tasks usually rely on the professional judgment of the assessor, and some proponents view such subjectivity of scoring to be the hallmark of performance assessment" (Messick, 1994:5). However, in an intervention where the translation assignment will be 'to produce a coherent target text,' I will need a reliable instrument to assess target-text coherence.

As holistic text-quality assessment of high-school translations of Latin has not yet been studied, I have again turned to writing research, as I did for writing instruction in the previous chapter. In assessing writing assignments, both analytic and holistic models are studied and used (Wesdorp, 1981). Lesterhuis, Van Daal, Van Gasse, Coertjens, Donche, and De Maeyer (2018) studied the rating behaviour of teachers comparatively rating argumentative texts and find that:

based on this study we can assume that comparative methods in particular are a valid method for the assessment of complex skills, as the comparative methods enable the teacher to obtain reliable scores for complex skills more easily than using analytic models and teachers focus on these higher order skills while assessing the texts. (2018:15)

Research shows that the rating behaviour of the assessor is influenced by the scoring method used, whether this is analytic or holistic (Barkaoui, 2010; Lesterhuis et al., 2018). In the analytic model for Latin translation assessment the rating scale applies to one specific text, as a result of which the rater is primarily focused on applying the rating scale on elements of the text. Thus, the rating behaviour is influenced by the scoring method. Holistic assessment is less task-specific, as the standards the product should answer to are formulated in a more general sense. As a result the rating behaviour is primarily focused on the quality of the whole text that is rated (Lesterhuis et al., 2018). Holistic rating is less task-specific and therefore more generalisable than analytic rating and raters are more focused on the complete text than its elements. Therefore, holistic assessment could well be applicable to assessing coherence

in target texts in Latin translation as well and help raters to assess texts as a whole.

2   COLON RATING

The common approach to assessing Latin translation assignments in the Netherlands is to apply an analytic model: the colon-rating model. For this type of assessment the source text is divided into *cola* and a correct translation of each *colon* is rewarded with a number of score points, on a scale from one to three, that is determined in advance. The amount of score points depends on the difficulty of the *colon.* If several small errors can be made, the score for a correct colon is higher, usually with a maximum of three. Raters are instructed to deduct one score-point from the maximum score per colon for small errors (e.g. translation errors in tense or singular/plural) and to deduct all score-points for each colon that is not rendered correctly.

A colon is a Latin sentence or part of a sentence that conveys meaning. For instance, the sentence "*Huic et divus Augustus dedit secreta mandata, cum illum praeponeret Thraciae, quam perdomuit, et Tiberius proficiscens in Campaniam, cum multa in urbe et suspecta relinqueret et invisa*[54]" (Seneca, Epistula 83.14) is divided into five cola as:

Table 3: Example of Colon Scoring Model[55]

Colon	Maximum score
Huic et divus Augustus dedit secreta mandata, To him (the) deified Augustus gave secret assignments	2
cum illum praeponeret Thraciae, when he put him in charge of Thrace	2
quam perdomuit, that he controlled / had controlled (fully)	1
et Tiberius proficiscens in Campaniam, and Tiberius (did the same[56]), leaving for Campania	1
cum multa in urbe et suspecta relinqueret et invisa, when/though he left (a situation filled with suspicion and hate[57]) in the city/Rome	2

---

[54] *Dutch National Exam, Latin 2018.*
[55] *Dutch National Exam, Latin 2018.*
[56] *Translation of this element was provided in the annotations.*
[57] *Translation of this element was provided in the annotations.*

Assessment through this analytic colon-rating model scores very high on interrater reliability if performed by expert raters.[58] The colon-rating model is used to rate the translation task in the Dutch final central exams, as well as pre-exam translations. It is customary to rate pre-exam translation assignments by analogy to the practice in rating the final exams, i.e. to deduct one score-point from the maximum score per colon for small errors and all score-points in a colon that is not rendered correctly.

The colon-rating method tests the skill to render each *colon* of source text linguistically equivalent in the target text. Thus, students may very well translate one or more *cola* correctly – i.e. in a grammatically equivalent way – while the meaning of the sentence, section or text is lost on them (Kroon & Sluiter, 2010:28). Consequently, the coherence of the target text suffers, and a target-text *reader* can wonder what the text is about, while the target-text *rater* may not even notice the lack of coherence, as the model invites one to assess text quality merely on a colon level. In other words, a (grammatically) equivalent rendition of a source text (target text with a high colon-score) may yield an incoherent and consequently incomprehensible target text (Kroon & Sluiter, 2010).

The CITO[59] considers the index of items of the translation in the Dutch national exam very reliable (.88 in Latin final exams 2018) and considers one experienced rater, the teacher,[60] sufficient to grade it. The skill that is reliably measured by this model, however, is the ability to translate *parts* of the source text in *equivalent parts* of the target text, not providing a reliable measurement for target-text coherence.

To assess target-text coherence I must look into more holistic models for text-quality assessment, as the colon-rating model can be applied disregarding (in)coherence in a target text (Kroon & Sluiter, 2010). At the same time, a comparison between colon rating and a more holistic model could provide interesting insight into the relation between ST-TT equivalence and target-text coherence. A certain amount of linguistic equivalence between source text and target text needs to be observed for a text to qualify as a translation.

[58] *Opgaven, uitwerkingen en meer voor centrale examens 2018 vwo. (n.d.). Retrieved from https://www.cito.nl/onderwijs/voortgezet-onderwijs/centrale-examens-voortgezet-onderwijs/examenmateriaal-om-te-oefenen/vwo-2018/vwo-2018-tv1*
[59] *The Dutch organization for educational measurement.*
[60] *Bias is controlled by a system of appointing a second rater to all teachers to assess impartial rating.*

3   SCALE RATING WITH ANCHOR TEXTS

Research in Dutch writing education[61] suggests a holistic assessment with anchor texts as a reliable model for quality assessment. An anchor text is a text produced by a student and used as exemplary text (anchor) for a certain quality on the dimension of text quality. Trained raters compare texts to-be-rated with a set of anchor texts and decide on its quality in relation to these anchor texts. Thus a ranking order of texts is formed, from low to high quality.

Pollmann, Prenger, and De Glopper (2012) describe the practice of scale rating with anchor texts in Dutch writing education as follows: first, the scale developer ranks one hundred student texts intuitively in five quality levels: 1) very weak, 2) weak, 3) average, 4) strong, 5) very strong. Then, the scale developer chooses two texts representing each level and presents these ten texts to a team of seven expert assessors. The experts unanimously establish five anchor texts, each representing a quality level. The scale developer arbitrarily attributes a hundred points to the 'average' text, respectively 70 and 85 points to the two levels one standard deviation below average, and the levels above average respectively 115 and 130 points. To assess the quality of the other texts, each text was rated by three different assessors out of a team of student assessors trained in scale rating and using the anchor texts as points of reference. They provided individual scores for each text on a scale of 50 (quality below the 'very weak' anchor text) to 150 (quality above the 'very strong' anchor text) (Pollmann et al., 2012:19).

Text scale rating with anchor texts as frame of reference is used as a reliable instrument in holistic text-quality assessment for writing assignments[62] and it may offer possibilities for adaptation to assess target-text coherence in translations. The main adjustment in this type of assessment is that in rating translations the relation to the source text must always be part of the equation, as a target text that bears no relation to the source-text cannot be considered a translation, even if it is a coherent text in itself.

3.1 Exploring Scale Rating for Assessing Target-Text Coherence.

In a sub-study, I explored whether this type of comparative judgement could be applied to students' translations and provided a reliable score to indicate target-text coherence (Luger, 2016). At the same time, I wanted to learn what criteria determine the coherence quality of the target text. Additionally, I explored the correlation between colon rating and scale rating, hypothesising

---

[61] *Wesdorp (1981); Blok (1986); Pollmann, Prenger, and De Glopper (2012).*
[62] *Average correlation between six raters .65, average Cronbach's α 0.88 (Pollmann et al., 2012: 22)*

that some students produce 'technically' acceptable translations that are inco-
herent target texts (high colon rating, low scale rating), while others produce
a (rather) coherent target text, but make technical translation errors (low colon
rating, high scale rating). The challenge was to keep the relation between
source text and target text in view, without losing sight of the aim of perform-
ing a holistic assessment.

I conducted a small-scale study in 2014-2015 in my exam class,[63] working
in close collaboration with my co-supervisor Suzanne Adema.[64] The students
translated a Latin source text. The target texts were scored by two raters with
colon-scoring. Subsequently we worked in three phases to establish a scale:
1) we each chose an 'average' text as anchor, 2) we discussed our choices and
the criteria we used to determine which text we would use as anchor, 3) we
both rated the other target texts comparing them to our anchor text. After the
rating we explored correlations between scale rating and colon rating.

3.1.1   Task and Procedure

The translation task was scheduled early on a Monday morning the day after
autumn-break. The pressure to perform well was low, as the translation task
was not an official test and the grade did not count. The task was introduced
as practice for the translation assignment for the final exams 2015.[65] Students
translated a Latin text of 151 words into Dutch in ninety minutes. Students
had no reason to pay more attention to target-text coherence than usual, as
they worked under the assumption the target text would only be rated using
the regular colon-score model. Students were allowed to use the Latin–Dutch
dictionary with morphologic appendix (Pinkster, 2018). The source text (see
Table 4) was a letter by Cicero, *ad Familiares*, 14,4 (1-3).[66]

3.1.2   Establishing an Anchor Text

To establish an anchor text for the target-text quality assessment we had to
take an extra step compared to Pollmann et al. (2012): we had to ensure the
relation between target text and source text. Therefore, we started by text-lin-
guistically analysing and discussing the source text to describe its intratextual
coherence. In his letters Cicero shows a tendency to skip from one subject to
the other, as a result of which the coherence of sections can appear random.

---

[63] *The class consisted of twenty-four students (nine male, seventeen female) aged 17-18. As
two male students were absent, we had 22 available texts.*
[64] *Therefore, I mostly use the plural 'we' in the description of this exploration of anchor texts
for assessing target-text coherence.*
[65] *Subject of the Latin Central Exam 2015: Letters by Cicero, Pliny and Seneca.*
[66] *The source text was offered for translation practice in Jansen, Struyk and Hunink (2014).*

Other prose texts commonly show more explicit intratextual coherence, e.g. by using signal words.[67] Keeping that in mind, we summarised each section of the source text, formulating its main thought, describing tone and relation with other sections. We both took notes and used our notes as an outline of intratextual text coherence, assuming that discussing and agreeing on these topics would provide a shared frame of reference.[68] Having agreed upon an outline of the source text's intratextual coherence, we considered target texts that follow the source text's main thought, tone, and relation between sections to have the best possible intratextual coherence.

To establish a central anchor text, we followed the procedure for selecting anchor texts as described by Pollmann et al. (2012) intuitively splitting the target texts in three piles, representing texts of below-average quality, average quality and above-average quality all relative to this sample. The middle pile was expected to be the largest. By dividing the middle pile into three new piles, Pollmann et al. (2012) further refine the results, until they have one text left in the middle: the text anchoring the average coherence quality of the group, which is the empirical average. Following this procedure, we encountered a problem splitting our target texts in three piles using the outline of coherence as a guideline, as the pile of *below average* texts was largest (rater 1: $n = 12$ and rater 2: $n = 17$) while the *average* pile was much smaller (rater 1: $n = 2$ and rater 2: $n = 4$). Apparently, we had been assessing the coherence of the target text in relation to the source text instead of in relation to this sample, as a result of which our three piles did not represent the empirical average, but more traditional 'grades': *insufficient, sufficient, good.* After discussing this phenomenon, we re-evaluated and found that the empirical average in this case appeared to be an incoherent rendition of the source text. Re-evaluating, we found differences in how much we adhered to source-text equivalence: incomplete or incorrect rendition of the source text while the rest of the target text remained coherent resulted in ranking as *below average* in some cases and in *above average* in others. The discussion of the target texts that one rater considered to be below average, and the other above average, proved fruitful for refining the instrument. We formulated three dimensions in addition to the outline of intratextual-text coherence:
1. coherence and structure of the target text in itself,
2. target-text idiom, i.e. was the target text 'typical Dutch,'
3. extent to which source-text structure is followed, in content and in order.
We re-assessed the source text on a section level, distinguishing four main sections (*Table 4*).

---

[67] *Kroon (2007: Chapter 10); Adema and Van Gils (2015).*
[68] *The omission of using a written outline of the intratextual coherence of this text led to different interpretations of its outline of intratextual coherence.*

Table 4: Four Sections in Source Text: Translation and Main Thought

Section	Latin	Translation[69]	Main thought
1	Tullius Terentiae suae, Tulliolae suae, Ciceroni suo salutem dicit, Ego minus saepe do ad vos litteras, quam possum, propterea quod cum omnia mihi tempora sunt misera, tum vero, cum aut scribo ad vos aut vestras lego, conficior lacrimis sic, ut ferre non possim.	TO TERENTIA, TULLIOLA, AND YOUNG CICERO Yes, I do write to you less often than I might, because, though I am always wretched, yet when I write to you or read a letter from you, I am in such floods of tears that I cannot endure it.	Cicero greets his family The writer apologises for writing infrequently, stating two reasons: 1. General misery, 2. Confrontation with loss (by reading and writing letters) is unbearable.
2	Quod utinam minus vitae cupidi fuissemus! Certe nihil aut non multum in vita mali vidissemus. Quod si nos ad aliquam alicuius commodi aliquando recuperandi spem fortuna reservavit, minus est erratum a nobis;	Oh, that I had clung less to life! I should at least never have known real sorrow, or not much of it, in my life. Yet if fortune has reserved for me any hope of recovering at any time any position again, I was not utterly wrong to do so:	The writer utters regret that he did not commit suicide earlier. Argument: he would have suffered less.
3	si haec mala fixa sunt, ego vero te quam primum, mea vita, cupio videre et in tuo complexu emori, quoniam neque di, quos tu castissime coluisti, neque homines, quibus ego semper servivi, nobis gratiam rettulerunt.	if these miseries are to be permanent, I only wish, my dear, to see you as soon as possible and to die in your arms, since neither gods, whom you have worshipped with such pure devotion, nor men, whom I have ever served, have made us any return.	The writer expresses hope that things could take a turn for the better, depending on fate. If not, the writer expresses as his only desire to be with the addressee and die in her embrace. The reason: Gods nor men have shown gratitude.

| 4 | Nos Brundisii apud M. Laenium Flaccum dies XIII fuimus, virum optimum, qui periculum fortunarum et capitis sui prae mea salute neglexit neque legis improbissimae poena deductus est, quo minus hospitii et amicitiae ius officiumque praestaret: huic utinam aliquando gratiam referre possimus! | I have been thirteen days at Brundisium in the house of M. Laenius Flaccus, a very excellent man, who has despised the risk to his fortunes and civil existence in comparison to keeping me safe, nor has been induced by the penalty of a most iniquitous law to refuse me the rights and good offices of hospitality and friendship. May I some time have the opportunity of repaying him! | The writer provides the addressee with an update on his current situation. He has been staying with Laenius Flaccus for thirteen days. He characterises Flaccus as an excellent man who risks a lot by providing him shelter. The writer expresses the (feasible) wish to be able to repay him. |

The main thought of section one and three was rendered adequately by a considerable group of students, while section two and four were not rendered adequately. Our discussion resulted in the description of an empirical mean as: a target text that follows the pattern of alternately comprehensible (one and three) and incomprehensible (two and four) sections. Target text three follows this pattern, which initially led to a below average assessment in one rater and above average in the other. This text was chosen as the central anchor text, as it was showing skill with room for improvement, thus representing the 'middle pile.' This refined procedure reduces the holistic score to a semi-holistic five-item model, as it now represents coherence of four *sections* and of the text as a whole.

3.1.3   Rating the Texts

Having established the central anchor text and having refined criteria for assessing text coherence, both raters assigned 100 score points to the anchor text and then related the other texts to it, attributing 200 points to a text that was twice as coherent as the anchor text and 50 to a text that was half as coherent.

At this point both raters were very familiar with the source text and the comparative assessment with the anchor text took no more than about three minutes per text. It must be said that both raters struggled with the concept of holistic quality assessment. We felt that assessing target-text quality without looking too much into the details of the source text went against the grain of the classical philologist and we experienced increasing insecurity as to the reliability of this method.

3.1.4   Correlation and Significance

The insecurity we experienced was removed when we compared our assessments. *Figure 4* shows the correlation between the score points both raters attributed to the texts: the y-axis for rater one, the x-axis for rater two. The figure show five plateaus: rater one used an ordinal scale with five levels (50 -75 -100- 125 -150 points) shown as the plateaus in *Figure 4*, while rater two used a ratio scale, attributing a separate score (varying between 15 and 150 points) for each text. The difference in approach is not problematic, as this exercise was intended to explore possibilities to apply this method, focusing primarily on correlation of assessment in relation to the anchor text, and raters agree on most texts as to its position to the anchor text. However, five texts remain problematic: texts 4, 6 and 26 (rater one high, rater two low), and texts 10 and 22 (rater one low, rater two high). The discussion of the criteria for assessing these texts further refined the implicit norms both raters applied: the

'penalty' for clumsy Dutch wording and partially incoherent (or missing) sections varied. The ommission of the *ingratitude of gods and men* (section three) was rated differently, as was the penalty for the wrong translation of *salutis* (*genetive, singlar, salus*) as "greeting." Students apparently confused *salus* (safety, salvation, welfare) with *salutare* (to greet) or *salutatio* (greeting). As a result of which one rater deemed the whole section incoherent, while the other did not.

Figure 4: Scatterplot: Coherence Scores by Two Independent Raters.
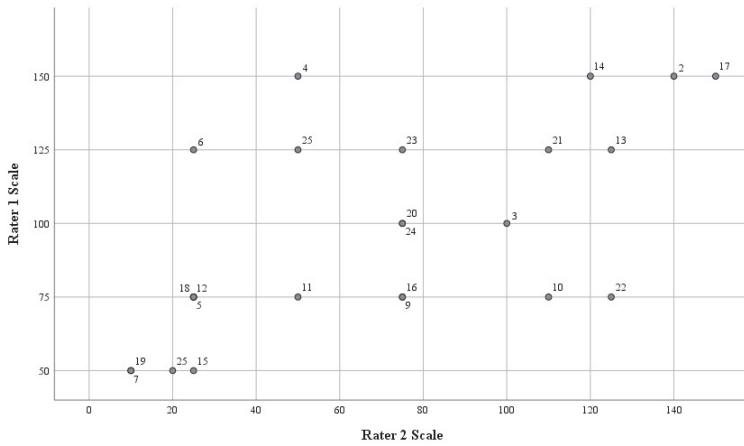
Figure 5: Scatterplot: Equivalence (Colon) and Coherence (Scale).
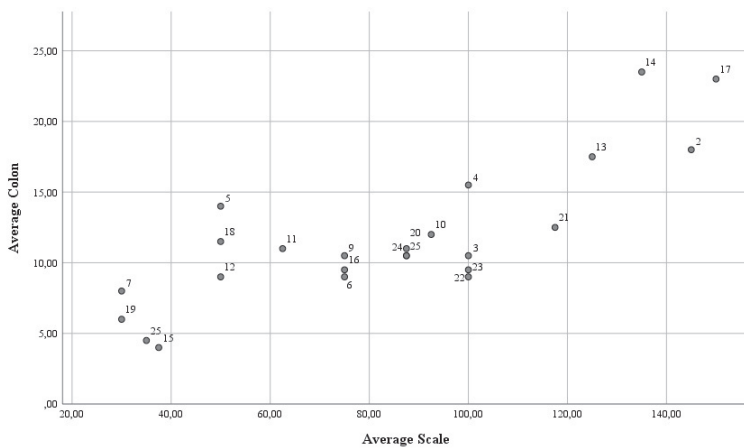
If the correlation is >.70 in assessment by three trained assessors familiarised with the rating scale, the assessment procedure is considered to be reliable, i.e. scores are not (too) dependent on a random team of assessors. In our exploratory scale assessment the correlation between the *two* assessors was $r = .62$ ($p <.001$), which makes it already a fairly reliable instrument. However, reliability would improve with a third assessor. Additionally, more specific and detailed discussion beforehand will lead to improved agreement in assessment. Reliability could be further improved if the assessors were more familiar with the procedure and with more than one anchor text.

In the colon-scoring model the maximum score was established at twenty-nine points on twenty items. Correlation between the two assessors is very high ($r = .95$, $p < .001$), as is to be expected with two expert assessors performing this type of assessment. With this high a correlation between raters in colon-scores, one experienced rater will suffice when we use colon-rating in assessing linguistic equivalence in the translations further on in this dissertation. The combined score total of both assessors is rather low (M 11,4; $SD =$ 5,2). This is an indication that the equivalence of source text and target texts, which is measured by this type of rating, on average, is wanting.

Having assessed the rater-reliability of both models, this exploration of scale rating provides an opportunity to compare students' scale and colon scores and investigate correlation. It is to be expected that correlation is not very strong, as different aspects of the target text are rated by scale rating (text coherence) and colon rating (linguistic equivalence). On the other hand, some correlation of the scores is likely, as linguistic skills are bound to influence the production of a coherent target text to some extent.

The correlation between colon score and scale score is significant in both assessors (rater one: $r = .77$; rater two: $r = .67$) which means that the types of scores have common grounds (about 60 to 45% respectively). In other words, the holistic score partly represents – as expected – linguistic accuracy. This means that knowledge of Latin plays a role in target-text coherence, but does certainly not tell the whole story of coherence quality.

Is the discrepancy between colon score and target-text coherence we identified in section two[70] reflected in this exploratory comparison? We compared the average scores of both raters for each type of assessment ($r = .81$). The scatterplot (*Figure 5*) of the average colon scores of both raters (y-axis) and average scale scores of both raters (x-axis) shows the scores form a (more or less) diagonal line, thus indicating fair agreement. However, when we take a closer look at *Figure 5*, we notice the diagonal is mainly formed by texts 7, 15, 19 and 25, scoring low in both colon and scale rating, and texts 2, 13, 14

---

[70] *A target text with a high colon score may lack coherence, while a coherent target text may yield a low colon score due to too many linguistic inadequacies (Kroon & Sluiter, 2010:28).*

and 17, scoring high in both models. When we leave these eight texts out, the impression of a diagonal disappears and a cloud of scores without any tendency of a line can be discerned just below the centre of the figure: average on one axis, not average on the other. This supports the claim that in some texts colon and scale rating differs.

To give an impression of a target text with a relatively high colon score[71] (14 out of 29 score-points) and low coherence score (25 out of 200 score-points)[72] I have provided an example in *Table 5,* left panel. An example of the opposite, a target text[73] with a relatively low colon score (9 out of 29 score-points) and high coherence score (125 out of max. 200 score-points) is shown below in *Table 5,* right panel.(Dutch followed by English).

These examples illustrate that different dimensions of the target text are measured by colon-rating (ST-TT equivalence) and by scale rating with anchor texts (intratextual coherence).

Table 5: Student's Translations of Cicero, ad Familiares, 14,4 (1-3)[74]

Section	Example 1: High equivalence, low coherence. Colon: 14/29 score-points; holistic score: 25/200)	Example 2: Low equivalence, high coherence. Colon: 9/29 score-points; holistic score: 125/200)
1	*(geen aanhef, SL)* Ik geef minder vaak brieven aan jullie dan ik kan, bovendien omdat weliswaar alle tijden miserabel voor mij zijn, maar vooral als ik óf schrijf aan jullie of als ik jullie brieven lees, ik op deze manier overmand word door tranen zodat ik niet kan verdragen.	Marcus Tullius groet zijn Terentia en Tullia en zijn kleine Cicero, Ik kan niet vaak genoeg een brief schrijven aan jullie omdat weliswaar al mijn uren toch miserabel zijn maar vooral (omdat) wanneer ik of schrijf aan jullie of jullie brieven lees, ik zo overmand word door verdriet, dat ik het niet kan verdragen.
2	O dat ik zo weinig vol van levenslust ben geweest! In het leven heb ik zeker niet veel of geen kwaad gezien. Maar als het lot van een of ander iemand ooit tamelijk veel hoop heeft behouden om gelukken terug te winnen is de vergissing minder gemaakt door ons;	Hoe verlang ik toch niet in leven te zijn. In dit leven zie ik niks meer dan zeker veel slechte dingen. Maar als er ons nog enige hoop rest en met een beetje voorspoed ons geluk terugwinnen is er door ons een vergissing gemaakt;

---

[71] *Rater 2.*
[72] *Text 5 in the scatterplots.*
[73] *Text 22 in the scatterplots.*
[74] *English follows Dutch.*

3   als deze ongelukkigheden blijvend zijn, ik spreek tot jou voor het eerst de waarheid, mijn liefste, wil ik je zien en sterven in je omhelzing, aangezien en niet de goden, die jij altijd gewetensvol hebt vereerd, en niet de mensen, aan wie ik altijd heb gediend, aan ons erkentelijkheid terug hebben gegeven.

   als deze omstandigheden blijvend slecht zijn wil ik werkelijk eerst dit van jou, mijn liefste, ik verlang jou nog één keer te zien en in jouw armen te sterven, aangezien noch de goden, waar jij gewetensvol voor gezorgd hebt, noch de mannen die ik altijd gediend heb, kunnen mijn dank aan jou teruggeven.

4   Ik ben 13 dagen bij M. Laenius Flaccus in Brundisium geweest, de beste man, die gevaar van zijn lot en zijn hoofd minder belangrijk heeft geacht dan mijn heil, en niet door de meest erge straffen van wie vriend en vijand er vanaf is gebracht de geldende verplichtingen ter beschikking te stellen. Zouden wij hem toch ooit grote dankbaarheid terug kunnen geven!

   Ik was dertien dagen bij M. Laenius Flaccus in Brundisium aansterkend; wie de kans op gevaar en onze hoofden minder belangrijk acht dan een groet en heeft vijanden van de wettelijke zeer strenge straffen af gebracht en staat vriendelijk in voor die geldige verplichtingen. Hoe kan ik deze man ooit bedanken.

---

1   *(No address SL)*
I give letters to you less often than I can, moreover because indeed all times are miserable for me, but especially when either I write to you or when I read your letters, I am in this way overwhelmed by tears, so that I cannot bear;

   Marcus Tullius greets his Terentia and Tullia and his little Cicero,
I cannot write a letter to you often enough, because indeed al my hours are miserable anyway but especially (because) when I either write you or read your letters, I am so overwhelmed by grief, that I cannot bear it.

2   O that I have been so little full of lust for life! In life I have certainly seen not much or no evil. But if some person's fate ever kept rather much hope to regain joys the error made by us is less

   How I long to not be alive at all. In this life I see nothing more than certainly many bad things. But if any hope is left for us and with a bit of success to regain our happiness a mistake is made by us;

3   if these unhappinesses are permanent, I speak to you the truth for the first time, my dearest, I want to see you and die in your embrace, for neither the gods, who you have honoured faithfully always, nor the people, to whom I have always served, have returned gratefulness to us.

   if these circumstances are permanently bad I truly want this from you first, my dearest, I long to see you one more time and die in your arms, as neither the gods, for whom you have faithfully catered, nor the men I have always served, can return my gratitude to you.

| 4 | I have been with M. Laenius Flaccus in Brindisium for 13 days, the excellent man, who considered the danger of his fate and his head less important than my well-being, and not by the most terrible punishments by which friend and foe have been dissuaded to make the current obligations available. Would we ever be able to return great gratitude to him. | I was thirteen days with M. Laenius Flaccus in Brindisium growing stronger; who considers the possibility of danger and our heads less important than a greeting, and has dissuaded enemies of the lawfully very strict penalties and guarantees kindly for these current obligations. How can I ever thank this man. |

3.2 Conclusions on Scale Rating with Anchor Texts

Scale rating with anchor texts as a model for assessing target-text coherence differs from its original use of rating writing assignments in that the relation between source text and target text remains crucial.[75] A reliable procedure requires an intratextual target-text coherence outline to maintain the relation with the source text, five anchor texts, clear definitions of the dimensions of assessment, agreement between raters on the approach of wording: from clumsy Dutch to a free yet coherent rendering of the meaning of the source text. In addition, a team of (at least) three assessors is needed for a reliable outcome and the team needs some training using this model, to overcome insecurity and detail-orientation.

Based on my experience in this exploration I suggest the following procedure to apply this type of assessment successfully:

1. Formulate[76] a written intratextual target-text coherence outline,
2. Determine the target text representing the empirical mean and use it as central anchor text,
3. Determine two anchor texts below the empirical mean and two above the empirical mean to refine the assessment,
4. Rate all target texts related to the anchor texts.

In assessment, clear wording and using the appropriate language register are positively weighed, whereas lexical errors leading to nonsensical meaning (e.g. *salutis:* greeting) weighs negatively. The target text is rated based on the outline of coherence and the anchor texts, while the assessor disregards the Latin source text to avoid focus on details of the source text. A slightly faulty rendition of the outline of coherence is judged mildly, while a serious aberration is judged severely.

---

[75] *Writing tasks based on sources are related to translation task inasmuch as a relation with the sources must be maintained. However, the close representation of the source text that is required in a translation task is much more strict.*
[76] *Before presenting the source text to students.*

In conclusion, this type of assessment can be used to reliably assess target-text coherence. However, it has some practical disadvantages. It requires some amount of preparation to set up the assessment: formulating a written intra-textual target-text coherence outline, determining and discussing anchor texts, training at least three assessors,[77] and finally organising the assessment.

4   COMPARATIVE ASSESSMENT IN PAIRS

It is fairly easy to intuitively determine quality difference in pairs: A is better than B, or B is better than A. The University of Antwerp developed an online tool for comparative judgement in pairs: D-PAC. The products that are to be rated (texts, drawings, film clips, et cetera) are uploaded in D-PAC. The tool offers a randomly composed set of two products next to each other on the computer screen and the assessor chooses which product is better. Through the programme assessors can be provided with a specific instruction. The software takes over the elaborate task of ranking the products that are to be rated, while the (human) assessor is merely asked to compare two products at a time. It is also possible to use the tool for peer feedback, as a function is provided to give feedback on the rated products and explain the choices the assessor makes. When all assessments are made, the programme ranks the texts in an order from least to most meeting the task. At the time of the pilot I performed (see 2.4.1), a reliability of .70 to .80 was calculated using the formula $(X*10)/Y=Z$, with X being number of texts, Y number of assessors and Z number of comparisons each assessor has to perform. A lower number of comparisons results in a lowered reliability: $(X*7,5)/Y=Z$, reliability .60 -.70.

Figure 6 provides a screenshot of an example of the assessment page, taken from the information provided by Maarten Goossens (D-PAC team) for the pilot I performed in 2017. Over the course of my project D-PAC has developed into a well-researched reliable tool for comparative assessment (Van Daal, Lesterhuis, Coertjens, Donche, & De Maeyer, 2016; Lesterhuis, Verhavert, Coertjens, Donche & De Maeyer, 2017; Verhavert, De Maeyer, Donche, & Coertjens, 2018).

---

[77] *Probably many more, depending on the number of texts to be assessed.*

Figure 6: Screenshot of D-PAC.[78; 79]

4.1 Exploring Assessment with D-PAC

I used D-PAC in a pilot aimed at exploring the ease and efficiency of this method for assessment. In this pilot, I decided to involve a class of students in my research, by making them both text producer and rater. Each participating student (n=28) was provided with a login through email. In class they translated a Latin text into Dutch and saved their translation as a PDF file. They uploaded the PDF file in the tool and in the following class students logged into their D-PAC account to assess translations. In the D-PAC account pairs of target texts were presented to them with an assignment, instructing them to select from each pair the most coherent, most fluent Dutch text and provide specific feedback as to the quality of the texts.

In the pilot, students made twenty comparisons each, in seventy minutes, a little over three minutes per comparison. Reliability of the assessment was calculated at .64. This was lower than expected, but the analysis of the assessments showed a somewhat erratic pattern, indicating little familiarity with the method or little experience with assessing texts. After removing one outlier (text) reliability increased to .67.

---

[78] *Source: M. Goossens, D-PAC stappenplan.pdf (personal communication, May, 2017).*
[79] *The text in Figure 6 reads: (Title) "Step 4: Judging". (Text in textbox in the lower-right) "Here you can give feedback. To generate the best learning results, do not only mark WHAT is good or less good, but also HOW this could be improved in the next draft." (Translation by S.L.).*

4.2 Conclusions on D-PAC as a Tool for Comparative Assessment in Pairs

The reliability in the pilot of .64 is promising and could be further improved by increasing the number of comparisons and by including more experienced raters or including some training. The main advantages of D-PAC over off-line forms of comparative judgement is that a lot of the hassle is taken over by the software, which provides ranking of least- to most appreciated text. The short time it takes to make the intuitive comparison is an additional advantage. Before I decided whether to use D-PAC to assess target-text coherence in my experimental study (Chapter 6), I further explored its use with a larger number of texts, a larger number of comparisons, and a team of teacher-assessors for assessing target-text coherence in translations, the results of which will be discussed in Chapter 5.2.

5   CONCLUSIONS

In section three, I argued that colon-scoring models provide insufficient information on the coherence of targets texts. Therefore, I decided to use a holistic model to asses target-text coherence in the texts students produce in the experiment for this dissertation (Chapter 6). Based on reliability and ease of use the choice between scale rating with anchor texts and comparative judgment with D-PAC is not difficult. The reliability of both assessment tools is high. Scale rating with anchor texts is a form of comparative assessment as it applies the basic principle of comparing in pairs more elaborately, one part of the pair being the anchor text. However, the process of establishing anchor texts was complicated. The use of D-PAC proved to be less time-consuming than scale rating with anchor texts. The interrater reliability of comparative judgment in general is good, provided that enough comparisons are made by a varied team of experienced assessors (Van Daal et al., 2016; Lesterhuis et al., 2017; Verhavert et al., 2018). D-PAC proved to be an intuitive and user-friendly tool to perform this type of assessment.

As linguistic accuracy remains an important feature of translations, all target texts in the experimental study (Chapter 6) have also been colon rated, to enable further observations of discrepancies in both models of assessment.

CHAPTER 3
HOW DUTCH ADOLESCENTS TRANS-LATE LATIN INTO COHERENT DUTCH[80]

1   INTRODUCTION

The problem of the incoherent target texts students in upper secondary education produce when they translate Latin into Dutch has been described in chapter one and analysed in connection with translation studies, teaching practice in the Netherlands, and educational sciences. However, we have also seen (Chapter 1, *Table 1*) that some students can produce a target text that is more or less coherent and that can be understood without the aid of the source text.

 This chapter addresses two questions: 1) what are the translation activities of successful student translators of Latin and 2) what successful strategies do these students use to produce a coherent target text? In this dissertation, a coherent target text is defined as a target text that is comprehensible to the reader without prior knowledge of the source text, and that conveys the main message of the source text.[81] The theoretical foundation of the current chapter lies in translation studies as argued in Chapter 1.2. Research has shown that translation competence involves several sub-competences. According to Göpferich (2008, 2009) strategic competence, which determines when the translator moves from one sub-competence to another, has a leading role in this process. As discussed in chapter one, research specifically focusing on the translation of Latin is scarce, and most research examines reading and comprehending Latin as opposed to translating it. Florian (2015) studied high-school students' translation behaviour and showed that students can successfully solve translation problems presented to them by relying on text comprehension, while their morphologic knowledge plays a supporting role at best.

The current chapter concerns a study of the translation process of eighteen students who are successful translators of Latin, i.e. who generally translate Latin into a coherent target text. Their behaviour was studied through eye-tracking and stimulated recall to avoid the problems of think-aloud studies.[82]

[81] *Intratextual coherence, see Chapter 1.2.3 and Chapter 1.5.*
[82] *Chapter 1.2.2.*

The data collected support the findings of Florian as to the role of text comprehension. Furthermore, the data suggest that a wide variety of problem-solving strategies, the ability to switch strategies, and the ability to use metalanguage to verbalise the chosen strategy is distinctive for a successful translation process in students. Since research shows that target text revision is an important feature of the translation process of professional translators[83] (Breedveld, 2002; Breedveld & Van den Bergh, 2003), this chapter also discusses the position of revision in students' translation processes.

2   METHOD

The participants performed two tasks on a computer. They translated a Latin fable and they edited a Dutch translation of another Latin fable while their activities were monitored by eye-tracker and screencast. Simultaneously, their keystrokes were recorded by Inputlog. Immediately after the tasks the participants were invited to verbalise their activities, stimulated by viewing the eye-tracking film (stimulated recall). They filled in an evaluation form about their perception of the tasks after the completion of all activities.[84] Afterwards the quality of the translation was scored through colon-scoring to provide insight into the ST-TT equivalence of the texts. Target-text coherence was not assessed, as the coherence of students' target texts was a selection criterion for participation (see 2.1).

2.1 Participants

Although the eye-tracking study aimed at analysing the translation activities of relatively successful student translators of Latin and at understanding what successful strategies they used to produce a coherent target text, some variety in the quality of the translators was needed to be able to distinguish more successful and less successful strategies. Therefore, I asked teachers to select students based on the students' Latin grade[85] and their variety of experience[86] in Latin, thus suggesting a variety in quality of the target texts they would produce. I invited thirty students[87] who had been suggested to me by their Latin teachers to participate in this study. Of these thirty students eighteen responded positively (male - female: 50% – 50%): one with 3.5 years, twelve with 4.5 years and five with 5.5 years of experience. In compliance with the

---

[83] *Chapter 1.*

[84] *In this dissertation the results of Inputlog and the written evaluation will not be taken into consideration. See below, Chapter 3.6.*

[85] *Ranging from above average to excellent.*

[86] *3.5 to 5.5 years of experience in translating Latin of 2.25–2.8 instruction hours per year.*

[87] *All students from Het 4e Gymnasium in Amsterdam, 15 to 18 years old.*

rules of the ethics committee of the University of Amsterdam, I provided those who responded with the details about the project and the consent form they were to present to their parents.

2.2 Task

2.2.1   Source text selection

The Latin source text for the eye-tracking study had to accommodate three criteria:
- It had to be coherent to facilitate the production of a coherent target text. This excluded a text consisting of a series of unrelated sentences or a fragment of a long narrative as a suitable source text.
- It had to be sufficiently complex to provide the translator with translation problems that elicited different types of problem-solving strategies, e.g. using the dictionary, activating different domains of knowledge, such as knowledge of antiquity, the source language or the target language (Göpferich, 2008).
- It had to be short. It was important that participants did not feel pressured and the time the task took was limited to prevent loss of concentration. Time pressure and fatigue influence translation competence (Göpferich, 2008) and had to be controlled in this experiment.

Fables by Phaedrus meet these three requirements: they are short, complex and coherent. Understanding the meaning of the fable and the genre characteristics contributes to a coherent target text production. There was one caveat, though, which is that the participating translators do not necessarily have the required knowledge of the genre and must therefore be provided with some type of introduction before performing their task. When such an introduction is provided, fables present suitable texts.

2.2.2   Task and Task Construction

To avoid that the outcome could be attributed to the specifics of one fable, I used two different yet similar fables: the fable of The Fox and the Grapes[88] and of the fable of The Fox and the Tragedy Mask.[89] Each fable consisted of four sentences. Each sentence contained one or two translation problems, as shown in Tables 6 and 7, in bold.

---

[88] *Phaed. 4.3. Abbreviated in this chapter as GRAPE.*
[89] *Phaed. 1.7. Abbreviated in this chapter as MASK.*

Table 6: The Fox and the Grapes (GRAPE), Translation Problems in Bold

	Sentence	Type of problem
1	**Fame coacta** vulpes alta in vinea uvam **adpetebat**, summis saliens viribus. [A fox, forced by hunger, tried to reach a bunch of grapes high in a vine, jumping with all its might.]	**Fame coacta:** construction: *it must be determined which noun is congruent with coacta (ablative singular with 'fame' or nominative singular with 'vulpes'?* **adpetebat:** coherence: *the imperfect is used conatively.*
2	**Quam** tangere ut non potuit, discedens ait [When he could not reach it, he said as he went away.]	**quam:** construction: *the relative pronoun cannot be used in Dutch is this way.*
3	Nondum matura es; nolo **acerbam** sumere. [You are not ripe yet; I don't want to eat sour (grapes).]	**acerbam:** coherence: *ellips, 'uvam' is missing.*
4	**Qui, facere quae** non possunt, verbis elevant, adscribere hoc debebunt exemplum sibi. [Those who mitigate with words what they cannot achieve, will have to consider this example for themselves.]	**Qui**: coherence: *this is the transition to the moral of the story.* **facere quae:** coherence: (as the meaning of quae was given) the difficulty is the inversion of the translation of 'quae' and 'facere' to create coherence.

Table 7: The Fox and the Tragedy Mask, Translation Problems in Bold

	Sentence	Type of problem
1	**Personam tragicam** forte **vulpes viderat**; [A fox once had seen a tragedy mask];	**persona tragica**: semantic/cultural: *tragedy mask* **personam vulpes viderat**: construction: *object, subject, predicate*
2	**Quam** postquam huc illuc semel atque iterum verterat, [After he had turned it over once and back again],	**quam:** construction: *the relative pronoun cannot be used in Dutch in this way.*
3	'O **quanta** species!' inquit '**cerebrum** non habet!'	**quanta**: construction/semantic *exclamation* **cerebrum**: coherence: *implicit antithesis*

[He said 'O, such beauty! But it
has no brain!']

4	**Hoc** illis **dictum est** quibus ho- norem et gloriam fortuna **tri- buit**, sensum communem **abs- tulit**. [This is said to those to whom fortune gave honour and glory, but from whom it took common sense away.]	**Hoc...dictum est**: coherence: *this is the transition to the moral of the story.* **Tribuit … abstulit:** coherence: implicit antithesis

In both fables, the story develops in sentences one to three and the moral of the story is conveyed in sentence four. The assignment started with a written introduction on the genre, which covered the typical animal characters of a fable and the text structure with moral and provided a fable in a Dutch translation as an example. After the introduction the translation task and the editing task were presented.

The Latin text was more extensively annotated than usual,[90] while the translation problems in each fable were intentionally left out of the annotations to avoid influencing which strategy the participants applied by suggesting a certain meaning. The translation assignment was similar to what students are used to. It read: "*Translate this fable into comprehensible and well-written Dutch.*"

The introduction and the tasks were tested in a group of proficient translators of the same age group at a different school, where it took students approximately thirty minutes to perform both tasks.

2.3 Research Design

The sequence of the fables and tasks was counterbalanced; ten participants translated fable GRAPE and edited fable MASK and eight participants translated fable MASK and edited fable GRAPE. Participants were randomly assigned to a group: participants in group A and B were presented the translation assignment first, the editing assignment second; participants in group C and D were presented the assignments in reversed order, as shown in *Table 8*.

---

[90] *Extensive annotations with complete matches lift the cognitive load (Göpferich, 2009).*

Table 8: Distribution of Editing and Translating Tasks over Groups

Group	Translate GRAPE	Translate MASK	Edit GRAPE	Edit MASK
A (n=4)		Task 1	Task 2	
B (n=4)	Task 1			Task 2
C (n=4)		Task 2	Task 1	
D (n=6)	Task 2			Task 1

2.4 Technical Equipment

I used a TOBII tx300 eye-tracker and a connected laptop. The eye-tracking software measured the fixation count, fixation duration and the visit count of the translator's eye movements in three areas of interest: source text, target text and dictionary. The software additionally created a screencast of the translation action combined with the visual representation of the eye movements in red dots and lines and filmed the translator's face by webcam. The participants worked with a keyboard and mouse connected to the laptop. The keylogging programme Inputlog was installed on the laptop.

Figure 7: Computer Screen Layout.

Area 1 SOURCE TEXT + ANNOTATIONS (BELOW TEXT)	Area 2 TARGET TEXT
	Area 3 LINK TO ONLINE DICTIONARY
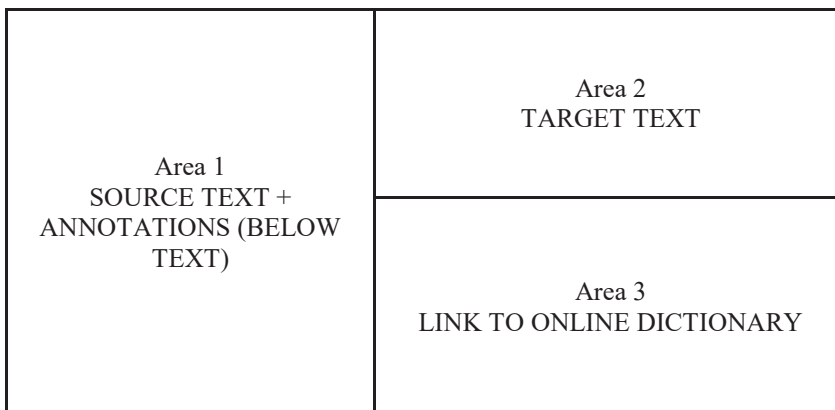
The introduction, task 1, and task 2 were presented as webpages on a website constructed in WordPress. Since the tasks were presented in this online web

environment and the participants used an online dictionary,[91] the laptop was connected to the internet. The computer screen was divided in three areas as shown in *Figure* 7.

2.5 Measurements

The six tools and methods that were used to collect data are summarised in *Table 9*. Inputlog records all keystrokes (including space bar and deleted text) in a timeline, provides a record of the process of text production, and produces a sequence in time of each key pressed, thus giving a visual representation of the writing process.

The stimulated recall interview was partly a spontaneous reflection of the participant on the eye-tracking film and partly prompted by the researcher with questions such as:

"So *quam* was a problem, do you remember how you tackled it?"

"What made you reconsider your initial translation?"

"How did you feel about your solution?"

The screencast software registered all action on the computer screen including the representation of eye-movements by moving red dots and lines, as the screenshot (*Figure 8*) shows. The target text the participants wrote was saved as a Word document. The evaluation form was a paper form with eight statements the participants could agree with on a scale of 1 (strongly disagree) to 5 (strongly agree):
- The introduction was helpful for assignment 1 and 2.
- Reading the introduction took me a long time.
- The (translation) assignment was difficult.
- The number of annotations on the (translation) assignment was too extensive.
- The (editing) assignment was difficult.
- I had enough time.
- The assignments were fun.
- I worked hard.

---

[91]*Woordenboek Latijn-Nederlands, online edition [Latin-Dutch Dictionary]. (2018). Amsterdam, the Netherlands: Amsterdam University Press. Retrieved form: http://www.latijnnederlands.nl/.*

Table 9: Tools and Methods to Collect Data

Tool / Method	Variable	How	Data
Eye-tracker	Translation process	Measuring fixation duration and count of eye movements	*Quantitative* Fixation duration, fixation count and visit count in three areas of interest: target text, source text, and dictionary
Inputlog	Target-text production process	Keystroke logging	
Stimulated recall	Translation process	The interview was recorded as screencast with audio	*Qualitative* Thoughts on the translation process underlying the eye movements and translation decisions the participant made
Screencast software	Translation process	Screencast	*Quantitative* 10s interval timeline of the translation activities
Translation product	ST-TT equivalence	Text in Word	*Quantitative* Translation problem solved/not solved Rest of the sentence correct/incorrect
Evaluation form	Perception of the tasks		*Qualitative* Thoughts on the assignment/ performance/tasks

2.6 Procedure

Testing the technical equipment. The research was preceded by a technical testing period from 1 April – 12 April 2016. When testing the eye-tracking software, it became clear that the measurements in area 1 (
Figure 7) were recorded, whereas in areas 2 and 3 they were not, although the eye movements did appear in the screencast as red dots (fixations) and lines (movements) (*Figure 8*).

Figure 8: Screenshot Area of Interest (AoI) Source Text.

This malfunction seemed to be caused by the fact that the source type that is to be recorded can be selected in the TOBII software, but when one website, such as the WordPress site, is selected, other sources in the same computer screen such as target text and dictionary cannot be recorded. To solve this problem, a new website was created,[92] integrating the WordPress site containing the tasks, the link to the dictionary and an area to produce the target text. In the target text area, a Word document in Inputlog could be opened where the participants typed their target text.

Setup. The eye-tracker and laptop were situated in a research area without direct daylight, as daylight can influence the recording of the eye movements

---

[92] *Thanks to Marco Kragten, who so patiently assisted me in the technical set-up of this study and the analysis of the eye-tracking data.*

(*Figure 9*). The research area was located at the participants' school to provide a sense of approachability.

Figure 9: Setup of the Eye-Tracker.
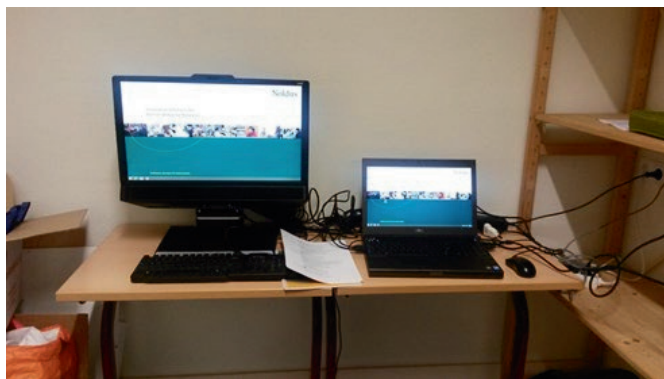
The research was conducted from 13 April to 1 May 2016, testing 1-5 participants per day. Each participant was instructed individually by the researcher before performing the tasks. The instruction involved technical instructions (eye-tracker calibration, opening the Word file in Inputlog) and some hands-on practice (the opportunity to browse the website briefly and to type some text in the Word file in Inputlog). The use of an online dictionary was not familiar to the participants. They were instructed to conduct a broad search by typing 'root*.'

All participants could take as long as they wanted for task 1 and 2. After these tasks, the participants filled in an evaluation form. All participants took the task seriously and worked hard. The participant who was scheduled late on Friday afternoon and two participants who were scheduled the morning after a school party reported they were less focused than usual.

Finally, the participants viewed the eye-tracking film and reflected on their translation behaviour, providing insight into the thoughts behind the eye movements. Their reflection was prompted both by the film of the eye-movements and by questions asked by the researcher concerning the translation problems and the translation process. The whole procedure took around sixty minutes per participant.

Data collection. Data were collected on both the translating and editing task. As the present chapter focusses on assessing translation activities rather than revision activities we use the data of the translation task only. Time limitations have prevented the analysis of the editing data within the scope of this dissertation. I will discuss the data collection following the tool or method of *Table*

9: eye-tracking, Inputlog, stimulated recall, screencast, translation product and evaluation. An overview of the data I collected successfully is presented in *Table 10*.

I experienced some difficulty in the data collection of the eye-tracking measurements. All eye-tracking data were missing from two participants' measurements. Additionally, the eye-tracking measurements in the area of interest (AoI) *target text* of some other participants were missing. Review of their screencasts however, showed that these participants had been looking at the target text, so apparently their gaze was not registered by the eye-tracker. In some other participants the measured eye-tracking values in the AoI *target text* were not consistent with the eye movements observed in the screencasts. Therefore, the target-text eye-tracking data of all participants were deemed contaminated and I decided to exclude these data from the analysis. A possible explanation for this could be that the opening of the Inputlogfile in the target text area on the website interfered with the eye-tracker's measuring. The quantitative data of the eye-tracking in the two remaining AoI's (*source text including annotations* and *the dictionary*) were analysed. The measured fixation count and the sum of measured fixation duration and were highly correlated: the more fixations, the larger the sum of the fixation duration ($r = .98$). The calculation of the mean fixation duration (fixation duration/fixation count) provided a useful new measurement: the average time a fixation lasts in an Area of Interest.

In the data collection by Inputlog a technical problem occurred that I was not able to solve. The timeline of the keystrokes showed the inexplicable appearance of chunks of text as if these chunks had been copied and pasted from another source. The screencasts however, showed a constant typing pattern, without any copy-pasting. Therefore, I could not use the Inputlog data to analyse the writing process. Fortunately, the translation process could be analysed by using the screencasts.

The recordings of the stimulated recall interviews of five participants were damaged and could not be played or analysed. Out of eighteen stimulated recall interviews, it was possible to use the recordings of thirteen interviews on the translation task for qualitative analysis.

The two participants whose eye-tracking films failed experienced some difficulty in the stimulated recall interview, as they had to remember what they had been doing based on a simple screencast without eye-tracking markings.

The screencasts of the translation process of all participants were saved, though –as mentioned before – in two screencasts the representation of the eye movements was not recorded.

The texts participants wrote as a translation were all successfully saved as Word documents. The evaluation form was filled in by seventeen participants.[93]

In spite of the difficulties in data-collection mentioned, sufficient data were collected for this study, as *Table 10* shows.

Table 10: Data Collected Successfully

Successfully collected data	Eye-tracker	Inputlog	Stimulated recall interview	Screencast of translation process	Target text	Evaluation form
Number of participants	16	0	13	18	18	17

2.6.1   Coding

Based on the translation activities reported by professional translators (Breedveld 2002) I coded the screencasts in 10s intervals to seven activities: 1) reading the Latin source text, 2) reading the annotations, 3) using the dictionary, 4) writing the target text, 5) improving the target text, 6) reading the target text and 7) "other". The activity of "formulating" was left out, as it is not visible behaviour in a screencast. This type of interval coding provides a timeline of the different activities involved in translating. Coding the screencast proved to be more difficult than expected, as the activities of the translating participants switch constantly even within 10-second intervals.

The solutions of the presented translation problems in the target text were scored as successful or not successful for each sentence of the target text (MASK 1–4 and GRAPE 1–4) by one coder. It was possible to code the first translation problem in each segment as solved (1/0), the rest of the sentence as translated correctly (1/0) and, in the segments where two translation problems occurred, it was possible to score the second problem as solved (1/0). Thus, each segment was provided with a maximum score of 2 or 3, as *Table 11* shows.

---

[93] *I forgot to present the evaluation to one participant.*

Table 11: Maximum Quality Scores per Segment

Fable	Segment	Score-point (max)
MASK	1	3
	2	2
	3	3
	4	3
GRAPE	1	3
	2	2
	3	2
	4	3

The data from the stimulated recall interview were typed verbatim and qualitatively coded by using semi-open coding via Nvivo.[94] The initial basis for the coding of the interviews was formed by the translation sub-competences of professional translators (Göpferich, 2008). I extended this basis with observations made while I was evaluating the protocols, as not all activities I observed could be captured by Göpferich's translation sub-competences. I rearranged Göpferich's sub-competences and added my observations. My *categories of knowledge* find an equivalent in the *domain (sub)competence* and the *(sub)competence in at least two languages* in Göpferich's model. My *categories of activity* are partly equivalent to Göpferich's *tool and research competence* (*use of the dictionary*) and her *translation routine activation competence*, in addition to which I have distinguished a metacognitive and a linguistic translation strategy. My *categories of mind* were added to Göpferich's model based on the interviews.

In *Table 12*, an example from the interviews is given for each category. We coded the interviews of 13 participants in Nvivo to these categories. The interviews were subsequently coded by two coders in the following categories:
1. Categories of knowledge: knowledge of antiquity and of the world, of Latin and of Dutch.
2. Categories of activity: use of dictionary, metacognitive translation strategy and the linguistic translation strategy.
3. Categories of mind: text comprehension, self-control, meta-commentary (e.g. "words starting with a letter 'q' always bother me.")

---

[94] *Nvivo is an intuitive tool used for coding and analysing data. https://www.nvivo.nl/.*

Table 12: Categories of Knowledge / Activity / Mind

Category of knowledge/activity/mind	Example from interview
Knowledge of antiquity	"Then I am thinking: mask! But I came to realise this quite late. It is not just a tragic mask, because, if, it becomes clear it is about actors."
Knowledge of the world	"[…] but it can't be about a person, it's about a mask. A mask doesn't have a back of the head."
Knowledge of Latin	"I wasn't sure about *viderat*, so I looked, what tense it was… But I think it is, errr, pluperfect."
Knowledge of Dutch	"And I add 'and', between 'this way AND that'."
Metacognitive translation strategy	"I just put it [meaning a word, SL] there for the time being. If I want to change it later on, I can delete it."
Linguistic translation strategy	"I saw *viderat*, which is the predicate."
Use of dictionary	"I saw *personam tragicam*, I thought 'well, a tragic person', so I wrote it down, but I immediately looked it up as well."
Text comprehension	"(…) but it can't be about a person, it's about a mask. A mask doesn't have a back of the head."
Self-control	Q: "when did you change the meaning?" A: "When I found out it is about a thing he picks up and then ... I don't remember what I wrote down eventually."
Meta-commentary	"If I could translate more freely, I would say: 'The moral of the story is…' but if it is for Latin class, I would stick to 'this is said'."

3 RESULTS: VARIOUS STRATEGIES IN PARTICIPANTS

The eye-tracking measurements of seven MASK and nine GRAPE translators were successful, with each participant translating four sentences, adding up to measurements for n = 28 in MASK and n = 36 in GRAPE. Analysing their eye-tracking data, I found considerable differences between participants in terms of how often they switched between the dictionary and the Latin source text (visit count) and how often they fixed their eyes there (fixation count). *Table 13* shows these large differences between participants. The *SD* in the measurements in visit count, fixation duration and especially fixation count is considerable, which means that participants varied considerably in the number

of times they fixated their gaze in the areas of interest *source text* and *diction-ary*. Considering these large differences between participants a quantitative analysis of the data would not likely yield reliable results. The small variance in the mean fixation duration shows that the duration of the fixations did not vary as much: even if the participants fixated more often, they did not, on average, fixate longer. This implies short looks and longer gazes are balanced in the participants.

Table 13: Behaviour in AoI's Source Text and Dictionary (M, sd)

	Source Text		Dictionary	
	MASK (n = 28)	GRAPE (n = 36)	MASK (n = 28)	GRAPE (n = 36)
Fixation count	271.29	358.72	115.64	232.25
	172.51	*243.09*	*118.77*	*244.38*
Fixation duration	61.52	88.00	29.16	60.47
	41.95	*67.89*	*31.31*	*63.52*
Mean fixation duration	.22	.23	.22	.22
	.036	*.044*	*.087*	*.091*
Visit count	45.29	65.58	15.96	36.14
	34.90	43.22	14.24	38.41

The observed differences between participants invite us to look at the data on a smaller scale. Therefore, I will start my discussion of the results by focusing on the translation process of one participant. I selected the participant for this case study because of his exemplary use of metalanguage, which led to an above-average successful solving of some, not all, of the presented translation problems. The first subsection (3.1) presents his results per method of data collection, the second subsection (3.2) returns to all participants and illustrates the use of metalanguage to solve two specific translation problems. The last subsection (3.3) discusses revision in the translation process of all partici-pants.

3.1 Case Study of John

I named the participant I decided to present in the case study *John*, by a ficti-tious name. His results will be discussed per method or tool of data collection, starting with the translation product which was used to assess the quality of his translation, through to the screencast to analyse his translation process, his eye-tracking data and finally his stimulated recall interview, thus painting a portrait of John as a translator. John is an eighteen-year-old boy, with 5.5 years

of experience in translating Latin and good Latin grades; he translated the fable of the tragedy mask (MASK). The fable in Latin is shown again[95] in *Table 14* with a translation[96] for better understanding.

Table 14: Fable of the Fox and the Tragedy Mask

Segment	Text
1	Personam tragicam forte vulpes viderat;
2	Quam postquam huc illuc semel atque iterum verterat,
3	'O quanta species!' inquit 'cerebrum non habet!'
4	Hoc illis dictum est quibus honorem et gloriam fortuna tribuit, sensum communem abstulit.
1	A fox once had seen a tragedy mask
2	After he had turned it over once and back again
3	He said 'O, such beauty! But it has no brain!'
4	This is said to those to whom fortune gave honour and glory, but from whom it took common sense away

John's quality scores. *Table 15* shows the quality scores of John's translation per sentence (MASK1–4). The second column shows the score for the first problem in each sentence, the third column scores the translation of the rest of the sentence and the fourth column shows the score for the second problem, if applicable.

Table 15: John's Quality Scores per Sentence

Sentence	Problem 1	Sentence	Problem 2	Total Score John	% Score	Max
MASK1	1	1	1	3	100	3
MASK2	1	1	X	2	100	2
MASK3	1	0	0	1	33	3
MASK4	1	0	0	1	33	3

Looking at the total scores of the quality of his solutions we see that John's translations in the first and second sentence are the maximum 3 and 2 points

---

[95] *See also* Table 7: The Fox and the Tragedy Mask
[96] *Translation by SL.*

(marked yellow). In the third and fourth sentence he solves the first problem correctly, misses the second problem in both cases and does not translate the rest of the sentence equivalently.

When we compare his scores tot the group average (*Table 16*) we see that his quality score is above average in the first two sentences, but below average in sentence 3 and 4. We will look into his translation process to discover the thoughts and actions that resulted in these scores.

Table 16: Quality Scores and John's Scores (Sentence Level)

Sentence	Max score	M	SD	John
MASK1	3	1.8	0.9	3
MASK2	2	0.8	0.7	2
MASK3	3	1.6	0.5	1
MASK4	3	2	1.1	1

John's screencast. Figure 10 shows the colour-coded timeline of John's translation. John translated for 15.18 minutes, showing the following pattern. At the beginning, we see long red chunks when John is looking up words in the dictionary or the annotations (eighteen times).

Figure 10: John's Translation Activities in 10s Intervals.

Legend:
Yellow = source text reading; Red = looking up (annotations/dictionary); Green = target text production; Blue = target text reading and improving; Grey = other.

He is reading the source text in short (yellow) intervals, while some writing (green) and revising (blue) occur. This is followed by a period of shorter activities, which means more shifting from one area to the other. Writing (green) and revising (blue) become more frequent, while source text reading (yellow) remains a frequent activity. Some intervals (grey) are coded as *off task*. Towards the end, reading of the source text remains frequent, whereas the frequency of looking up is decreasing and activities concerning the target text are almost as frequent as reading the source text. In the last two minutes his focus is almost entirely on the target text, indicating a revision phase concerning the text as a whole. In four instances John was *off task*, already checking out the

revision task outside the scope of this dissertation, before making his last revisions to the target text.

John's translation process shows a shift from orientation on the source text to writing and revising, which is in line with the translation process of professional translators as described by Breedveld (2002).

John's eye-tracking and stimulated recall. Table 17 and *Table 18* show John's eye-tracking measurements for each sentence. The measures of fixation count, fixation duration sum and visit count between sentences within this participant differ greatly, as the high standard deviations show. His fixation count in the source text ranges from 164 in MASK1 to 530 in MASK4. The variance is his fixation duration in source text is small (*Table 17*) whereas the dictionary can elicit slightly longer and shorter looks (*Table 18*).

Table 17: Source Text: John's Data for 4 Sentences

Source text	MASK1	MASK2	MASK3	MASK4	*Mean*	*SD*
Fixation count	164	357	233	530	321	160.59
Fixation duration sum	35.77	76.66	47.32	125.95	71.425	40.22
Fixation duration mean	0.22	0.21	0.2	0.24	0.2175	0.02
Visit count	28	49	35	25	34.25	10.69

Table 18: Dictionary: John's Data for 4 Sentences

Dictionary	MASK1	MASK2	MASK3	MASK4	*Mean*	*SD*
Fixation count	176	277	354	3	202.5	151.67
Fixation duration sum	47.26	67.11	95.69	0.57	52.66	40.01
Fixation duration mean	0.27	0.24	0.27	0.19	0.24	0.04
Visit count	22	37	40	1	25	17.83

Looking at the visit count, I noticed that it is higher in the source text than in the dictionary in MASK 1, 2 and 4, while this is reversed in MASK3 with 40 visit counts in the dictionary and 35 in the source text. Considering that

MASK3 consists of only seven words,[97] the visit count of 40 in the dictionary with 354 fixation counts is impressive. To understand why this happened we have to combine these data with the stimulated recall interview, where John explains his thoughts on his translating activities, stimulated by looking back at his eye-tracking film.

I will take a closer look at John's interview and eye-tracking results, discussing each sentence separately. *Table 18* shows that in the first sentence (MASK1) the mean fixation duration in the dictionary is relatively high (f = 0.27), which may be explained by the choice between *person* and *mask* John had to make. John explained in the interview that he tackled the first translation problem (*persona tragica*) in various ways: first trying the most equivalent translation: *a tragic person*, then, after looking *persona* up in the dictionary, changing *person* into *mask* and eventually, by activating knowledge of antiquity, John changed his translation into *tragedy mask*. This last change took place as a part of his revision phase at the end of his translation and is not included in the eye-tracking measurements of this sentence.

The second sentence contains the translation problem of the relative connection *quam*. The eye-tracking data show that John's visit count (f = 49) in the source text (*Table 17*) is higher than in the other three sentences, and in the dictionary (f = 37) (*Table 18*) it is the second highest. John keeps going back and forth between source text and dictionary (and the target text a bit, as we know only from the screencast). John explained in the interview that he was uncertain how to solve this problem. He tried the dictionary, but that was of little use, then he solved it by looking at the context:

> I knew the fox was turning something, but what he was turning into was missing. And then I thought, *quam* can refer to the object, and I thought 'Oh, then it must refer to the mask!' But it took a while before I figured that out.[98]

In the first two sentences, the higher visit count in the source text is explained by the type of strategy John said he used here. John had to think about the translation of *persona tragica* and *quam*, and he could not find the solution in the dictionary. Therefore, his focus was predominantly on the source text.

John said he solved the first problem in the third sentence (*quanta*) by looking in the dictionary, which is consistent with the eye-tracking data: only in this sentence visit counts in the dictionary (40 visit counts, 0.27 mean fixation duration) outweigh those in the source text. John solved the first problem correctly but stumbled in the translation of the rest of the sentence, as *Table*

---

[97] *MASK3:* O quanta species!' inquit 'cerebrum non habet! *[He said 'O, such beauty! But it has no brain!']*
[98] *All participants' quotes were translated into English by SL.*

15 showed. The second problem in this sentence is the implicit antithesis, which I categorised as a coherence problem (*Table 7*). John did not solve this coherence problem correctly, but we must bear in mind that the explicit translation of an implicit antithesis is not central to intratextual coherence: although adding a word improves the coherence of the target text, it is still comprehensible without it. Nonetheless, John overlooked this implicit antithesis both in his translation product and in his stimulated recall interview.

Most of the words in the last sentence were annotated in the task, and both coherence problems, the transition from story to moral and the implicit antithesis, cannot be solved by looking them up in the dictionary. The single visit count in the dictionary shows us that John realised that the answer to the problem could not be found there. In the interview John said about the fourth sentence that the main problem seemed to him how to phrase *hoc illis dictum est* in Dutch:

> Well, it still sounds a bit 'Latin' so to speak. I wouldn't do that in Dutch.
> If I could translate more freely, I would say: 'The moral of the story is
> …' but for Latin class, I would stick to 'this is said'.

As a result of this reasoning, John solved the first problem in this sentence correctly. He missed the second problem, the implicit antithesis, again and also failed to translate the rest of the sentence correctly.

John the translator. By combining all data, I paint the portrait of John the translator. It becomes clear from the coding of the screencast that his translation process roughly resembles that of professional translator as discussed in chapter one, since we can discern three out of four phases: 1) orientation (frequent dictionary use and source text), 2) text production (with dictionary and some revision) and 3) revision (Breedveld 2002). John does improve his target text, but does not produce a complete second target text.

The eye-tracking data show us that John's visit counts and fixation counts vary greatly per sentence and the interview reveals that he shows quite a large range of categories of knowledge and sub-competences. He uses alternative strategies if he finds his first strategy faulty. The eye-tracking and interview show that he is able to switch between these strategies, on some occasions even consciously.

Furthermore, it must be noted that John was the only participant translating this fable who solved the *persona tragica* problem correctly by using domain knowledge of ancient theatre practice.

When looking at both his results and his translation strategies, John is one of the better translators compared to the other participants. However, I selected students with good Latin grades in general. As this chapter aims to

investigate translation processes of proficient students, it is relevant to examine how the translation processes and problem-solving strategies of the other participants relate to John's. The following section will discuss the group's results, focusing on two translation problems all participants had to solve.

3.2 Results of the Group: Metalanguage and Problem Solving

As discussed above, the stimulated recall interview was prompted by the eye-tracking film. Watching their eye movements and gaze patterns, participants were able to see in real time where they paused to think during their translation process. In the interview, participants could explain what their thoughts had been at that specific moment. In doing so, differences between participants in the ability to explain their thoughts using metalanguage could be observed. The interviews provided insight in the way the participants tackled the translation, revealing some interesting discrepancies, as some valid strategies resulted in faulty translations and some faulty strategies led to good translations.

Section 3.2 discusses the solving of two translation problems: 1) the relative connection *quam* (MASK2 and GRAPE2) and 2) the translation of the moral (MASK4 and GRAPE4), focusing on the use of metalanguage in the stimulated recall interviews combined with the quality scores. I selected these two problems as they require different problem-solving strategies. The first problem presents an issue that cannot be solved by using the dictionary, thus eliciting other problem-solving strategies. I selected the second problem because text coherence is most acutely demonstrated by the translation of the transition of story to moral. The same problems occur in both fables, which is why the results of all participants can be considered.

3.2.1   Solving 'quam.'

The hypothesis was that solving the translation problem of the relative connection would require both knowledge of Latin and knowledge of the target language. The participant had to identify the relative connection and activate a translation routine, using knowledge of Latin, and subsequently apply an inversion in the translation of the relative pronoun and the conjunction *ut* or *postquam* (*Table 19*) using knowledge of Dutch to produce a coherent sentence in the target text.

Looking up a relative pronoun in the accusative case in the dictionary is never a successful strategy. However, a number of participants kept looking for a solution to the problem of *quam* in the dictionary but could not find it. As participant A[99] (GRAPE2) explained:

---

[99] *Apart from John, participants in this Section are identified by letters A–H.*

So, after that, I continued looking for '*quam*,' but I didn't succeed. Finally, I thought it could be '*quam + ut*' or something like that, it could be some combination, but that didn't work out either. So here I am looking up '*ut*' as well.

Table 19: The Relative Connection in Fables GRAPE and MASK

	Sentence	Type of problem
GRAPE2	Quam tangere ut non potuit, discedens ait: [When he could not reach it, he said as he went away]	quam: construction: the relative pronoun cannot be used in Dutch is this way.
MASK2	Quam postquam huc illuc semel atque iterum verterat, [After he had turned it over once and back again]	quam: construction: the relative pronoun cannot be used in Dutch in this way.

In the following quote, the linguistic phenomenon of the *relative connection* is identified correctly by participant B in MASK2, but its translation is incorrect:

> Then I saw *quam* and I thought it could be a relative connection. But then it said in the annotations 'he said' for *inquit*, so my text was not right.
>
> Why was that?
>
> Well, I translated 'after she' and then *verterat*, but it said 'he said' in the annotations. So then I thought it couldn't be a relative connection, because 'he' is the subject.[100]

Apparently, this participant was aware of the concept of the relative connection but did not realise that this could involve all cases, including the accusative *quam*. We see the beginnings of linguistic reasoning, but it does not lead to a good solution, as the knowledge is incomplete.

Conversely, a good translation can be based on a wrong analysis of the syntax, as the following part of an interview with participant C concerning GRAPE2 shows:

> I did look it [*quam*] up, but I didn't know how to translate it. I thought about it for a long time. Eventually I really didn't know what to do.

---

[100] *This participant scored 0 for the translation problem, and 1 for the rest of the sentence in MASK2.*

So, how did you solve that?

'Because' seemed to be okay here, I thought that really fitted the sentence.

So, what word did you translate as 'because'?

Quam

You translated: 'because he couldn't touch it'? You added 'it' then?

Yeah.

Participant C was looking into the domain of knowledge of Dutch: "I thought that really fitted the sentence," and consequently adapted his interpretation of the Latin text into his understanding of it, thinking "because" is the translation of "*quam*". The resulting problem, the missing object, was solved by knowledge of Dutch and target-text comprehension, because "it" was inserted as object. This last choice was not made consciously and was only verbalised after prompting by the interviewer.

Participant A provided us with another example of this phenomenon, switching from the dictionary, The where the solution could not be found, to another strategy:

> I translated it [*quam*] with 'like that': 'like that he could…' and then I added 'them', 'he could not touch them' and then I continued 'going away he said'…

By consciously adding the object, participant A showed a more explicit strategy than participant C. Participants A and C show that students use text comprehension as a means to achieve a correct solution for a translation problem, even when the use of the dictionary and their knowledge of Latin syntax fall short. They seem to rely on some sort of implicit understanding of the linguistic structure, which can lead to a correct translation, even when the participant thinks he is adding an object instead of translating *quam*. This reasoning shows some resemblance to John's solution, but while participants A and C more or less accidentally stumble upon the correct translation, John is *consciously* changing his strategy as he couldn't find *quam* in the dictionary: "I knew the fox was turning something, but what he was turning was missing. And then I thought, *quam* can refer to the object."

3.2.2   *Translating the moral*

In the last sentence of each fable (GRAPE4 and MASK4) the moral of the fable is presented. Therefore, this sentence has a more abstract type of content than the first three sentences (*Table 20*).

The hypothesis was that solving the translation problem of the moral would be facilitated by knowledge of the genre conventions. This knowledge could be triggered by the marking of the moral by *hoc exemplum* (GRAPE4) and *hoc dictum est* (MASK4) and the use of the relative pronoun without antecedent *qui* (GRAPE4) and *quibus* (MASK4) with the general meaning "(all) those who."

Table 20: The Transition from Story to Moral in Two Fables

Fable	Sentence	Type of problem
GRAPE4	Qui, facere quae non possunt, verbis elevant, adscribere hoc debebunt exemplum sibi. [Those who mitigate with words what they cannot achieve, will have to consider this example for themselves.]	Qui: coherence: this is the transition to the moral of the story. facere quae: coherence: (as the meaning of quae was given) the difficulty was the inversion of the translation of "quae" and "facere" to create coherence.
MASK4	Hoc illis dictum est quibus honorem et gloriam fortuna tribuit, sensum communem abstulit. [This is said to those to whom fortune gave honour and glory, but from whom it took common sense away.]	Hoc ... dictum est: coherence: this is the transition to the moral of the story. Tribuit… abstulit: coherence: implicit antithesis.

The variance in the quality scores (*Table 21*) suggests, again, some differences between participants. The high variance in the last sentences (MASK4 and GRAPE4), respectively 1.1 and 1.2, may be explained by the characteristics of the moral: participants either 'get it' or they do not.

The interviews provide some insight into the struggle. Participant E recognised the difficulty, but did not take the opportunity to write down the understood meaning of the sentence, scoring 0/3 in MASK4:

> "I suppose I could have phrased it better, but I didn't want to, you know, change it all, because I did understand what it's supposed to mean. […] But I just couldn't phrase it right."

Participant F was more adventurous and translated *dictum est* ("is said") as "applies to": "I did think the translation is a bit free, but 'is said' just didn't

feel […] good." The fact that this strategy pays off is reflected in the maximum score of 3/3 in MASK4 for participant F.

Table 21: Variance of Quality Scores per Sentence

Sentence	*Mean*	*SD*	Maximum
MASK1	1.8	0.9	3
MASK2	0.8	0.7	2
MASK3	1.6	0.5	3
MASK4	2	1.1	3
GRAPE1	2.2	0.9	3
GRAPE2	1.3	0.7	2
GRAPE3	1.2	0.6	2
GRAPE4	1.6	1.2	3

Participant G showed knowledge of Latin by expecting a common combination of words and by his awareness that words can be left out in Latin. He initially thought *verbis* had to be supplied to *illis* in MASK4. Eventually *self-control* supported by linguistic reasoning led to a correct solution of all problems in MASK4 for participant G (score 3/3):

> Well ... *hoc illis dictum* ... I immediately thought, *illis* means *illis verbis* 'by these words,' but eventually I changed that. So *illis* should be: 'said to those' so 'This is said to those, to whom Fate gave honour and fame' and then *"sensum communem abstulit"* but it is '*but* took away'(…)

Adhering to a *linguistic approach* combined with constant reflection on the storyline results in solving translation problems correctly, as participant H (GRAPE4) shows. This participant translates 'they' as subject, as if the subject in this sentence is identical to that of the sentence before. Thus, it is clear that the transition from story to moral is missed. However, he translates the second problem correctly:

> *Quae*, I changed the word order: I thought '*quae non possunt facere*' as 'the things they cannot do, they soften with words.' In Dutch 'they who do not do the things, soften with words' doesn't make sense. I found the first part [*of the sentence SL*] to be difficult and in the second part it was basically… well I didn't know if *adscribere* goes with accusative or dative/ablative, so I didn't know if it meant 'apply to this example' or 'apply to himself'. So at first I translated: 'they will have to apply this to' and then I didn't think it made sense that they should apply themselves to the example. And *sibi* would have to be accusative and I thought it didn't look like that. And I was thinking about *hoc*, it said in the annotations 'this,' but in Dutch 'this' can be used as a substantive

and adjective. So I thought 'this they will have to apply to themselves,' and then there was 'example' (*laughs*). And then I figured *hoc* could also be accusative, so it goes with *exemplum*!

In conclusion, this shows that in addition to sufficient knowledge of the source language, other strategies can help to solve the translation problems of translating the moral: *text comprehension*, which led to the free translation by participant F, and the *self-control* to be able to *change strategy if necessary* in participant G and H.

3.2.3   *Revision and problem solving*

Table 22 shows Breedveld's four phases and the specific activities for each phase in accordance with frequency (Breedveld 2002).

Table 22: Frequency of Translation Activities per Phase

Phase	Most frequent activities	Less frequent activities
Orientation 'first run-through' (very short)	Formulating Reading ST	Close to absent: other activities
Text-production 'second run-through'	Formulating Writing Reading TT	Reading ST and Evaluating TT; Dictionary and some planning; Low frequency: other activities
Improving the text 'third run-trough'	Reading TT	Some formulating; Evaluating TT same as 2nd run-through Little writing Reading ST practically absent
Assessing the fluency of the text 'fourth run-through'	Reading TT	More writing than formulating

Note: Based on (Breedveld (2002)

Breedveld's model of the translation process in professional translators as described in chapter one consists of four phases: orientation (first run-through), text production (second run-through), text production (third-run-through), and revision to assess the fluency of the text (Breedveld, 2002).

    The four phases Breedveld discerns and the behaviour observed in the participants in my eye-tracking study appear to be somewhat different: in high-school students' behaviour the third and fourth run through are almost identical, whereas activities from the phases of text production (second run through) and improving the text (third run through) appear to be mostly

intertwined. Therefore, I propose a slight shift for my dissertation in naming of the phases for describing the high-school translation process: (1) Orientation, (2) Text production (first draft), (3) Text production (second draft), and (4) Revision.

As revision activities such as reading and evaluating TT are present in the translation process that professional translators show (Breedveld & Van den Bergh, 2002: 330), and become more frequent towards the end, I hypothesised that proficient students would also revise their target texts towards the end of their translation process. The analysis of the screencasts showed some variation in the position of revision in the observed translation processes. In most participants' screencasts a revision phase could be discerned after the translation was completed, while some participants more dominantly revised per sentence. However, all students take time to revise their written text after completing their target text, some longer than others. This suggests that *revision* is indeed an activity in the translation process of students who qualify as proficient translators. Thus the translation process again resembles the writing process, considering the part revision plays in both (Van den Bergh, Rijlaarsdam & Breetvelt 1994).

To take a closer look at the revision patterns I selected four participants who translated the fable MASK: two students who had quality scores *below* average (3/11 and 5/11) and two participants who scored *above* average: John (7/11) and another participant (8/11). I will compare the role of revision in the translation processes of these four participants. It can be hypothesised that the quality of the translation correlates to the amount of time spent on the production of the translation: the longer the process, the higher the quality.[101] Therefore, I will first have a look at the quality scores of their translations and the time they spent translating (*Table 23*). The maximum quality score for the fable was 11 score-points, the mean quality score of the group was 6.125, *SD 1.73*.

Examining the time participants spent on the translation task (*Table 23*), it is noticeable that participant 1, with the lowest score, spent the shortest time on the translation task. This seems to support the hypothesis of a correlation between translation time and quality score. However, no firm conclusion can be drawn as the quality score of John – who did not translate as long as participant 2 – is above average.

---

[101] *See Chapter 1.2.3 (Gerloff, 1988 as cited in Breedveld & Van den Bergh, 2002:330).*

Table 23: Quality Scores and Translation Time

Participant	Score total	Time
1	3	13.02
2	5	16.51
John	7	15.19
3	8	22.35

In all four timelines (*Figure 11*) it can be observed that revision activities such as reading and improving the target text (blue) become more frequent towards the end of the process.

Figure 11: Four Translation Processes in 10s Intervals.

Legend:
Yellow = source text reading; Red = looking up (annotations/dictionary); Green = target text production; Blue = target text reading and improving; Grey = other.

Looking closely at the actual changes participants made during the revision activities, it becomes clear that the participants with lower quality scores did change their target text, but did not always improve its quality, failing to correct a solution that was wrong to begin with. For example, a participant translated *fame* as if it read *fama*, and subsequently revised the target text by changing *rumour* to *story*: though the activity was scored as revision, it was no improvement either as a result of insufficient analysis of the morphology or as a result of imprecise source-text reading.

In conclusion, I observed that revision activities are present in all participants' translation processes. A further analysis of the content of the revision activity showed that it depends on factors such as knowledge of Latin and text comprehension whether the revision activities result in actual improvement of the target texts.

4   CONCLUSIONS AND DISCUSSION

I return to the questions this chapter aimed to answer 1) what are the translation activities of successful student translators of Latin and 2) what successful strategies do these students use to produce a coherent target text?

The eye-tracking study conducted with eighteen students whose selection was based on their proficiency as translators yielded four translation activities I expect to contribute to coherent target-text production in general: (1) Source-text comprehension,[102] (2) Target-text revision,[103] (3) The ability of students to reflect on the translation process,[104] and (4) The use of a metalanguage to talk about translation decisions when switching between types of knowledge and between strategies.[105]

I found that recognition of linguistic features in the source language does not necessarily lead to a correct translation in the target language. Labelling the linguistic phenomenon correctly did not necessarily lead to the correct translation of the relative connection (3.2.1, participant B). At the same time an implicit understanding of the linguistic structure, knowledge of Dutch or target-text comprehension could lead to a correct translation, even when participants were unable to verbally explain their reasoning (3.2.1, participants A and C). In the stimulated recall interviews on solving the translation problems of the relative connection and the moral, I noticed that a broad range of types of knowledge, such as genre knowledge, source-text comprehension and target-text comprehension, as well as a conscious switching between these *types* combined with a conscious switching between *strategies* and the ability to reflect on the translation process such as John displayed, are decisive in a successful translation process.

I would therefore advise that these findings be used in order to try to help students improve their translation skills. Text comprehension can lead to a correct translation even when morphology or syntax is not analysed correctly, so we need to encourage students to use that type of strategy in addition to the more traditional linguistic approach.

The selected fables met the requirements of short, complex and coherent source texts. The results for both fables do not suggest significant discrepancies in difficulty between fables. I have specifically observed that knowledge of genre characteristics, e.g. the moral in a fable (3.2.2), is helpful in understanding the source text and consequently in producing a coherent target text.

Furthermore, teachers need to encourage students to talk about their translation decisions using metalanguage as John does (3.1.3), as this is a tool for

---

[102] *Sections 3.1.1 and 3.2.2.*
[103] *Section 3.2.3.*
[104] *Sections 3.1 and 3.2.*
[105] *Sections 3.1.and 3.2.*

reflection on their translation behaviour. In addition to the development of this metalanguage, target-text revision as such must be encouraged in all students. When students practice and learn how to revise and to reason about their choices, the coherence of their target texts is expected to improve. Not all changes are improvements, as we have seen in section 3.2.3. The quality of the revision is related to the content of the activity. I expect that improved metacognition of the translation process is reflected in the use of metalanguage and will thus contribute to more successful revision activities.

The number of participants (n= 18) does not allow robust statistical results about the quantitative data. However, the approach as a case study and the analysis of the interviews do provide interesting insights into the translation strategies of the participants. A more even distribution over the experience levels of Latin would have been preferable, but the younger students I approached were more hesitant to participate. The participants were highly motivated and enjoyed working with high-tech equipment to perform an old-school task. Furthermore, the gift certificate they received after completing the tasks may have helped their motivation.

As described above, the laptop used had to be connected to the internet. The wifi signal in the room was weak at times, and in several instances the test was interrupted by loss of connection. This problem was solved by using a hotspot on the researcher's mobile phone. The test of the technical setup for the eye-tracker revealed problems that were partly solved by adjusting the web environment, as discussed in Testing the technical equipment). I assumed that this new web environment would not interfere with the use of Inputlog, and the integration of the Inputlogfile in this setup was not tested. When unforeseen problems in data collection occurred, an Inputlog expert I consulted suggested that these problems could be related to this integration. The collection of data in Inputlog may have been disturbed by the maximising and minimising of the Word file. Additionally, the contamination of some eye-tracking measurements could also have been related to the combined use of the eye-tracking software and Inputlog. For future eye-tracking studies, I recommend more extensive testing of the setup, as this could detect such problems at an earlier stage.

I chose to set up a research area at the participants' school in order to facilitate participation, and this approach proved to be successful. Participants were relaxed and did not report excessive stress or anxiety. However, following the assistance with the set-up by Marco Kragten, continuous expert technical support in this environment was missing, which eventually proved to be a disadvantage. In future studies of this type I would recommend a research area embedded in a technical support system, in addition to more extensive testing.

The website containing the introduction and the two tasks was not fixed; participants could scroll up and down. This prevented defining separate AoI's for text and annotations in the analysis of the eye-tracking data. In future use of eye-tracking to study the translation process in students I would recommend a fixed text, to be able to identify the AoI's more precisely.

Fortunately, the loss of Inputlogdata could be compensated by the 10s interval scoring of the screencasts of the translation process. The loss of eye-tracking measurements however, impeded full analysis of the eye-tracking and this was frustrating.

Still, the eye-tracking study provides a valuable contribution to the understanding of the translation process in proficient student translators. Chapter four will discuss how to teach these translation activities to students who are less proficient by nature than the eighteen participants in the eye-tracking study.

CHAPTER 4
DESIGNING LESSONS
TO IMPROVE TARGET-TEXT
COHERENCE

1   INTRODUCTION:
LESSONS TO IMPROVE TARGET-TEXT COHERENCE

The present chapter aims at offering a solution for two of the problems I identified in the first chapter of this dissertation: 1) teachers do not really know *what* to teach their students to make them translate Latin into a coherent target text, and 2) no evidence-based instruction method for teaching Latin translation is used in Latin education practice, so teachers do not really know *how* to teach their students to translate Latin into a coherent target text.[106] To offer a solution for the first problem I present a process-oriented translation strategy focusing on target-text coherence. I offer the design principles for lessons teaching this strategy provide a solution to the second problem. The chapter starts by establishing *what* the process-oriented translation strategy should look like (Section 2) and continues by establishing *how* this strategy should be taught, formulating design principles for the lessons (Section 3.1) as well as criteria for source-text selection (Section 3.2) and a detailed description of the lessons (Section 3.3). The overall aim of this dissertation is to test the effect of the designed lessons on target-text coherence. Therefore, this chapter also covers the lessons I designed for a control condition (Section 4) and the experimental design I set up to test the lessons in both conditions (Section 5).

2   DEVELOPING A PROCESS-ORIENTED TRANSLATION
STRATEGY

In chapter three, I concluded that *source-text comprehension* can lead to a correct translation even when morphology or syntax is not analysed correctly. I also hypothesised that the ability to *reflect on the translation process itself* and the ability to verbalise these reflections using *metalanguage* contributes to coherent target-text production. Furthermore, analysis of the stimulated recall interviews suggested that a *conscious switching between types of knowledge* and *between strategies* while writing the target text are distinctive for a successful translation process. In addition to the findings from chapter three, we

---

[106] *Chapter 1.5.*

must keep in mind that target-text coherence in professional translators can be improved by *instruction on writing* (Schrijvers, 2014) and *on revision* (Breedveld, 2002; Breedveld & Van den Bergh, 2002; Robert & Van Waes, 2014)[107] as argued in chapter one. The importance of revision for coherent target-text production concurs with the observation that all participants in the eye-tracking study revised their target texts. I have operationalised these findings as five skills. The development of these skills in students is crucial for the production of coherent target texts:
-   the skill to reflect on the translation process,
-   the skill to use metalanguage[108] for reflection on translation decisions to *consciously* switch between types of knowledge and strategies,
-   source-text comprehension skills,
-   target-text production skills,
-   target-text revision skills.
The development of the skill to reflect on the translation process presupposes knowledge of the existence of a translation process. However, this knowledge may very well be rudimental or even absent in students. Mere awareness in translators of the existence of a revision phase may influence revision behaviour and therefore may favour the production of a coherent target text. I combined the development of these five skills to formulate my process-oriented translation strategy in four components presented in *Figure 12*: (1) Reflection on the translation process, (2) Source-text comprehension, (3) Target-text production and (4) Target-text revision. The importance of knowledge of the translation process as well as the skill to use metalanguage for reflection and conscious decision-making permeates the whole process-oriented strategy, as is reflected in *Figure 12* by the vertical position of "reflection on the translation process." I connected the other three components, source-text comprehension, target-text production and target-text revision, with the appropriate phase of the translation process in the middle section of *Figure 12* (Breedveld, 2002). After source-text comprehension, the focus of the process-oriented translation strategy moves away from ST towards TT. Each phase in the middle section is characterised by activities that are most frequent in this phase, presented in the right section of *Figure 12*. These activities are teachable translation activities (compare *Table 22:* Frequency of Translation Activities per Phase (Breedveld, 2002)).

The development of the skills underlying each component should be the content of the lessons aimed at improving the coherence of students' target

---

[107] *Explorative tests show that students who write a coherent target text take time to revise the text (Chapter 3.3.2); Robert and Van Waes (2014). In addition to this Schrijver (2014) shows that improving writing skills in translators improves the translation.*
[108] *Chapter 3.1.*

texts. Students should be offered knowledge of the translation process, develop their skills to reflect on it (left section), get used to the order of the phases of the translation process (middle section) and be trained in the activities most frequent in those phases (right column).

Figure 12: Process-Oriented Translation Strategy[109].
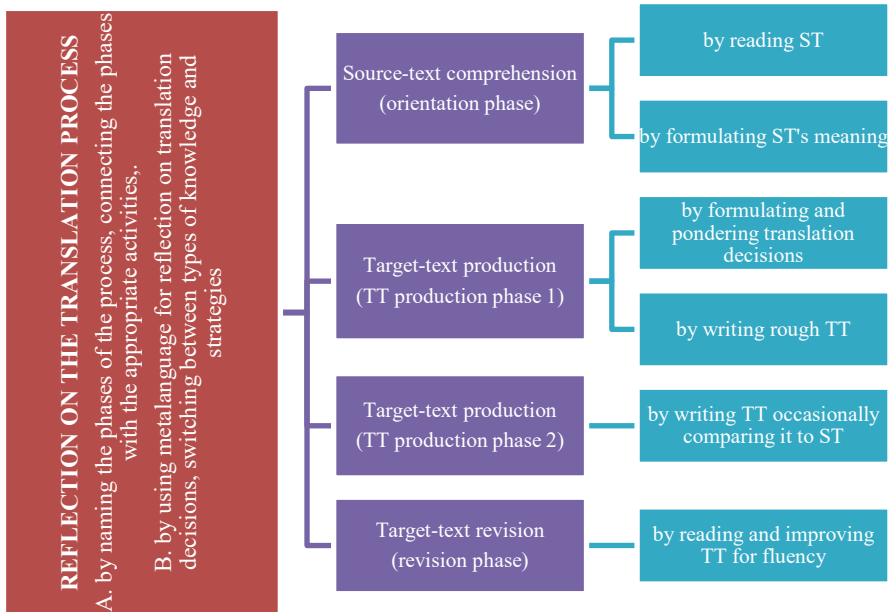
3   DESIGNING LESSONS TO IMPROVE TARGET-TEXT COHERENCE

Now that the content of the lessons has been (roughly) formulated, the next step is to establish how to teach students this process-oriented translation strategy. It seems viable to rely on research concerning learning and instruction of writing. In chapter one I argued that writing and translating competences are very similar (Schrijver, Van Vaerenbergh, Leijten, & Van Waes, 2019) as both writing and translating involve planning, formulating and revising. The similarity of the models of writing and translation competences supports a transfer from learning and instruction of writing to that of translation.

---

[109] *Based on Breedveld (2002) and the eye-tracking study (Chapter 3).*

The design of lessons to instruct students to use the process-oriented translation strategy requires: (1) design principles, and (2) criteria for the selection of appropriate source texts.

3.1 Design Principles

Translating a text is a complex process,[110] and I have formulated a process-oriented translation strategy, as reflected in *Figure 12*, to guide students to the production of a coherent target text. For the design of the lessons teaching strategy, I formulated three design principles, of which the first and second focus on the *instruction method* of the strategy, and the third focuses on the two competences the lessons aim to develop, 1) strategic competence concerning the *process* of translating and 2) communicative competence concerning *meaning and coherence* of source text and target text. These design principles guided and controlled the design process and served as evaluation criteria for the assessment of the design.

Principle 1: Self-Regulated Strategy Development (Graham & Harris, 1996; 2002) is a tested instruction form to teach a (new) strategy. As argued,[111] I used SRSD as a framework for the design of the lessons instructing the process-oriented translation strategy I developed.

Principle 2: Effective learning and instruction of writing is achieved with lessons following the four evidence-based recommendations by Graham and Perin (2007:467).[112] I used these recommendations in the design of the lessons teaching coherent target-text production.

Principle 3: *Self- Regulation* is a key aspect of SRSD. Students should be able to perform the strategy independently, being aware what they are doing (translation process) and what they are producing (a coherent target text). They need to be able to reflect on process and product. To develop 1) strategic competence concerning the *process* of translating and 2) communicative competence concerning *meaning and coherence* of source text and target text, I aimed at raising *awareness* of the translation process and at the development of metalanguage for reflection and conscious decision-making. I developed learning activities to raise this awareness.

The way I used each of these principles is explained in the text below. In addition to these three specific principles for the lessons, the design is founded on Merrill's general principles for teaching (Merrill, 2002:44-45, see Chapter 1.4.1.2).

---

[110] *Balbo (2011), Breedveld (2002), Göpferich (2009).*
[111] *Chapter 1.4.2.*
[112] *Chapter 1.4.2.*

3.1.1 First design principle: Self-regulated strategy development.

The close relation between writing and translating suggests that tested methods for writing instruction are well suited for teaching students to produce coherent target texts as well.[113] As Graham and Harris originally intended Self-Regulated Strategy Development for writing instruction in primary education, I have adapted it in this section both in terms of content and audience, to make it applicable for the instruction of the process-oriented translation strategy in upper secondary education. The adapted six stages form the backbone of the lessons and provide a rough, chronological sketch of the training. Table 24 shows the application of the stages of strategy-instruction to translation.

Table 24: Strategy Instruction Applied to Process-Oriented Translation[114]

Strategy Instruction Stage	Strategy instruction applied to the process-oriented translation strategy.
1. Develop and activate background knowledge	1. Develop and activate knowledge of (a) the translation process and (b) text coherence
2. Discuss the strategy	2. Analyse the translation process and discuss the four phases of the process, which develops awareness of the translation process.
3. Model the strategy	3. Model key activities for the text production and revision phase and use of metalanguage.
4. Memorise the strategy	4. Memorise the strategy by explicitly formulating and discussing translation process and translation decisions, developing metalanguage.
5. Support the whole strategy	5. Support the process-oriented translation strategy by assignments (scaffolding) for each phase. Students translate a Latin source text into a coherent target text, following the strategy.
6. Independent performance	6. Translate a Latin source text into a coherent target text (strategy not supported by guiding assignments for each phase).

Note: strategy instruction based on Harris and Graham

---

[113] *Chapter1.4.2.*
[114] *Harris & Graham (2002)*

Stage 1: Develop and Activate Background Knowledge

The concepts of *text coherence* and *translation process* must be activated or introduced in the first stage of the instruction, as these concepts are key to the experiment as a whole. I want to start by making students aware of their own translation process (stage 1a) and the concept of text coherence (stage 1b), as a successful translation process leads to a coherent target text. All upper-secondary students of Latin have experience in translating and have formed translation habits they are possibly unaware of. As it is not a common feature of Latin class to pay attention to the process of translating, it is to be expected that students are unaware that their habits form a process. I expect that raising awareness of their individual translation process is a prerequisite to improve these habits.

Stage 2: Discuss the Strategy

Adapted to the process-oriented translation strategy, *discuss the strategy* means that students discuss the four phases of the translation process based on an analysis of their translation process. Discussing the strategy promotes awareness and awareness is the first step for students in improving the translation process. Analysis of their own translation process and comparing it to peers' and professional translators' processes are the next steps. In this SRSD-stage each phase of the translation process must be discussed. The discussion needs to highlight that source-text comprehension starts in the orientation phase and that understanding the source text is crucial throughout the translation process, as we have seen in chapter three: source text comprehension is the *conditio sine qua non* for coherent target-text production. The discussion must also cover that target-text production is a skill in itself, and that target-text revision is important to achieve target-text coherence.

Stage 3: Model the Strategy

Students need to see how key activities within the strategy are performed. Target-text production and target-text revision are key activities in the coherent target-text production process of translators and modelling is the way to demonstrate these activities. In this stage we will model key activities for the text production and revision phase. I will use screencasts of my eye-tracking study[115] as a model for the text revision and text production activities performed by students who produce coherent target texts. Students watch a set of two screencasts.

---

[115] *Chapter 3.*

Stage 4: Memorise the Strategy
Students consolidate the awareness of the phases of the translation process and the activities that play a role in each phase by discussing and formulating improvements of the individual translation process. To stimulate the development of metalanguage, students analyse the activities they watched in the screencasts and compare their observations in pairs. They assess what they consider to be good or weak text production and text revision activities. In my adaptation this stage is not characterised by memorising the strategy, for example, by using an acronym.

Stage 5: Support the Whole Strategy
In this stage students perform the translation strategy supported by scaffolding assignments for each phase. I want students to consciously go through all four phases of the translation process, by offering assignments that stimulate appropriate activities for each phase. This means that this is the first time in the strategy instruction students experience the complete translation process and actually translate a text independently, trying to apply the strategy.

Stage 6: Independent Performance
In the last stage, students work towards independently going through all four phases of the translation process to independently translate a Latin source text into a coherent target text. A reminder summarising the four phases and the dominant translation activities for each phase can still be used if required, until the strategy is sufficiently internalised.

3.1.2   Second design principle: learning and instruction aimed at coherent target-text production.

In addition to strategy instruction (SRSD), I used the four recommendations for teaching writing[116] by Graham and Perin (2007) for the design of lessons aimed at coherent target-text production. Two recommendations overlap with SRSD, as both strategy instruction and modelling are recommended for writing instruction by Graham and Perin as well. Additionally, Graham and Perin recommend setting "clear and specific goals for what adolescents are to accomplish with their writing product" (2007:467). In the teaching of coherent target-text production, this would mean that the concept of target-text coherence must be part of the lessons as well as a translation assignment that refers to the production of a coherent target text.[117] The recommendation that

---

[116] *Discussed in Chapter 1.*
[117] *Chapter 1.3.2.*

students "engage (..) in activities that help them gather and organise ideas for their compositions before they write a first draft" (2007:467) can be applied to lessons teaching coherent target-text production by emphasising the importance of the orientation phase and offering activities that help students to understand the source text before they start target-text production. *Table 25* summarises Graham and Perin's four recommendations and my adaptation to the instruction of the process-oriented translation strategy.

Table 25: Application from Meta-analysis of Writing to Translation[118]

Recommendation for writing instruction (Graham & Perin, 2007)	Recommendation applied to instruction of the process-oriented translation strategy
1. Teach strategies for planning, revising and editing.	1. Covered by design principle 1.
2. Set clear and specific goals for the writing product.	2. Provide a clear translation assignment.
3. Stimulate students to gather and organise ideas before writing a first draft.	3. Stimulate students to understand the source text before writing a first draft of the target text.
4. Provide good models.	4. Covered by design principle 1 (stage 3).

3.1.3   Third design principle: learning activities aimed at raising awareness of (elements of) the translation process.

Strategic competence concerning the *process* of translating and communicative competence concerning *meaning and coherence* of source text and target text are developed by learning activities aimed at raising awareness of the translation process. Arguably, this awareness of the translation process leads to self-regulation in students, therefore *awareness* is closely related to the element of Self-Regulation in the SRSD of Graham and Harris. Learning activities concerning the third design principle raise 1) awareness of the strategy, 2) awareness of target-text coherence, 3) awareness of the translation activities belonging to the translation phases, and 4) awareness of the meaning of the source text.

Awareness of the strategy: The lessons must make students aware of the translation process strategy, its phases and the activities that that are linked to those

---

[118] *Graham and Perin (2007)*

phases. Awareness of the translation process in itself may improve the process strategy. Awareness of the translation process is stimulated by the following learning activities: identifying, categorising and comparing translation activities.

Awareness of target-text coherence as a characteristic of quality may improve the coherence of the text students produce themselves. Awareness of text coherence is stimulated by assessing coherence in target texts written by others and ranking these texts (compare and contrast). To improve their revision skills students are stimulated by the following learning activities: revision of an incoherent text into a coherent text, comparison of the revised texts, and revision of their own target text.

Awareness of the translation activities: Throughout the lessons students are developing metalanguage on translating to enable them to a) communicate about their translation process, b) compare their own *modus operandi* to that of others, and c) *consciously* switch between different categories of knowledge that translation requires. The development of metalanguage is stimulated by the following learning activities: verbalising translation activities, formulating strategies, giving and receiving feedback on translation activities and classroom discussion. Therefore, the lessons should feature group work and working in pairs, in addition to individual assignments.

Awareness of the meaning of the source text: to stimulate focus on the meaning of the source text, which may improve text coherence in the target text, the lessons concern the phases of *writing* and of *revising* the target text in the translation process. In the phases of writing and revising, thinking about the meaning of the text is stimulated through the learning activity of visualization by drawing.

These three design principles have led the design of the lessons. The following section discusses the selection criteria for the Latin source texts meeting the needs of lessons to improve target-text coherence.

3.2 Source Texts: Selection Criteria and Selection

The selection of source texts for the lessons to improve target-text coherence was guided by the same criteria I used for selecting source texts for the eye-tracking study: I needed short texts with a clear and coherent storyline, offering a certain degree of structural complexity. For the eye-tracking study I had selected fables by the Roman author Phaedrus as source texts, as they present short, complex and coherent stories, where the fables proved to be sufficiently

complex to elicit different problem-solving strategies in the participants.[119] This suggested that the fables would also be well suited for the lessons on coherent target-text production. The strict structure of fables, with three building blocks: situation, action, and moral, at either the beginning or the ending of the fable, provides a useful structural format for the translator. Additionally, fables are well-suited for the experiment as each fable tells a story with a moral and we aimed to make students aware that source-text comprehension is crucial for translation: the comprehension of the moral helps the translation of the story and vice versa. Consequently, students had to be provided with an introduction on the genre of the fable in the lessons.

I selected thirteen short fables with a clear storyline as source texts,[120] some of which were presented in Latin, some in a Dutch translation and some in Latin with a Dutch translation printed next to it. *Table 26* shows the selected source texts and their characteristics. It also indicates whether the fable is presented in Dutch translation or in Latin.

Table 26: Features of Selected Source Texts

Title of fable	Words	Moral[1]	Latin	Dutch
1. Ovis, cervus et lupus (I,16)	46	B	X	
2. De vulpe et uva (IV,3)	47	E	X	
3. Vulpes ad personam tragicam (I,7)	43	E	X	
4. Vacca, capella, ovis et leo (I, 5)	75	B	X	X
5. Asinus ad senem pastorem (I,15)	61	B		X
6. Mustela et homo (I,22)	75	E		X
7. Asinus et leo venantes (I,11)	86	B		X
8. Lupus et agnus (I,1)	94	E	X	
9. Duo calvi (V,6)	45	E	X	
10. Mulier parturiens (I,18)	41	B	X	
11. Canes famelici (I,XX)	37	B	X	
12. Canes et corcodilli (I,XXV)	51	B	X	
13. Muli duo et latrones (II,7)	80	E	X	X

[1]B = at the beginning / E = at the ending of the fable.

The source texts that were presented only in Latin were provided with extensive annotations to minimise the use of the dictionary. Research suggests that the complexity of choosing the right semantic value for a word while translating may lead to cognitive overload in the translator (O'Brien 2006, 2008). I

---

[119] *Chapter 3.2.2.*
[120] *Fable 2 and 3 were also used before, in the eye-tracking study.*

hypothesised that looking for the right meaning in a dictionary, which is in itself a complex activity for students (Bartelds, 2018), may distract from the production of a coherent target text. In addition to the annotations, a hard copy Latin-Dutch dictionary (Pinkster, 2018) was available, as students are used to using the morphologic help the dictionary offers in an appendix.

3.3 Lessons to Improve Target-Text Coherence

I introduced the design principles for the lessons to instruct the process-oriented translation strategy in Section 3.1 and the source-texts selection for these lessons in Section 3.2. The present section presents the lessons I designed to teach students how to produce coherent target texts when they translate Latin into Dutch.

Table 27 presents the five strategy components, the three design principles, and the work format that I expected to be best fitted for each learning activity. The colours used in Table 27 match the colours of the three sections (left, middle and right) of the process-oriented translation strategy in Figure 12. In addition to the specific design principles for the lessons, the lessons are based on Merrill's principles for teaching, as all well-designed lessons are (see Chapter 1.4.1.2). I would like to add to the information in *Table 27* that all lessons are building on the elements taught in the lesson(s) before and that participants are encouraged to apply the skills they learn to their regular Latin classes. I present the manner in which I distributed the learning activities over the actual lessons in the text following Table 27.

The first lesson aims at development of knowledge of the translation process by making students aware of their translation process (SRSD stage 1A). I use an interval scoring assignment to do so. In this type of assignment, students are asked to score their activity every 30s (marked by a beep-signal) on a form while performing a task.[121] Students perform a translation task and score translation activities as identified by Breedveld (2002): reading ST, formulating, using a dictionary, writing TT, reading TT, and improving TT. I use this assignment of scoring of translation activities as the starting point for the whole experiment: 'Notice your translation process'.

---

[121] *For this method see Olive, Kellogg and Piolat (2002).*

CHAPTER 4

Table 27: Strategy Components, Design Principles and Work Format

Four components of the Process-oriented translation strategy (I-IV)	Principle 1: Six stages of SRSD applied to the process-oriented translation strategy (1-6)	Principle 2: Four recommendations by Graham and Perin (2007) (a-d)	Principle 3: Learning activities aimed at self-regulation	Work Format
I. Reflection on the translation process	1A. Develop and activate knowledge of the translation process.		Notice your translation process.	I
A. Development of knowledge of the translation process	1B. Develop and activate knowledge of text coherence.		Discuss your decisions on coherence	P/W
			Assess and argue different types of translation.	I/P
B. Development of metalanguage for reflection.	2. Analyse the translation process and discuss the four phases of the process, which develops awareness of the translation process.	b. Set clear and specific goals for the writing product.	Compare with others and contrast (the phases of) your translation process.	P
			Discuss your translation process.	W
			Discuss the models' text production/ revision activities.	P/W

II Development of source text comprehension skills.	1B.	c. Stimulate students to gather and organise ideas before writing a first draft.	Read the source text to understand the story.	I
			Draw the story and compare drawings.	I/P
			Re-tell the story to someone unfamiliar with it.	P
IB, III Development of target text production & IV Development of target text revision skills	3. Model activities for the text production and revision phase and use of metalanguage.	d. Provide good models.	Observe and categorise the models' text production activities.	I
IB, II & IV	4. Memorise the strategy by explicitly formulating and discussing the translation process and translation decisions, developing metalanguage.	d. Provide good models.	Compare and contrast observed activities in models with your own activities.	I/P
IA, II, III	5. Support the strategy by assignments (scaffolding).	a. Teach strategies	Write chunks of the target text, focusing on meaning.	I
II, III	4 and 5.	a. and b.	Write a (more) fluent target text, focusing on meaning.	I
IV	4 and 5.	a. and d.	Observe and categorise the models' revision activities.	I
I, III & IV	4 and 5.	a. and d.	Compare and contrast with your own activities.	I/P

IA & IV	4 and 5.	a. and b.	Revise on text coherence by marking each sentence + (coherent) / – (not coherent) - without source text knowledge. - with source text knowledge (other's target text). - with source text knowledge (own target text).	I
		a. and c.	Read the source text focused on the story. Write chunks of the target text, focusing on meaning. Write a (more) fluent target text, focusing on meaning.	I I I
I-IV	6. Students translate a Latin source text into a coherent target text, following the strategy.	a. and b.	Revise on text coherence by marking each sentence + (coherent) / – (not coherent) - with source text knowledge (own target text).	I

Note Work Format: I: Individual, P: Pair work, W: Whole class

 I had experimented with this method before in regular class to assess whether the beeps interfered with the translation process and whether this was a useful tool for students to reflect on their translation process. After the translation task, I scored the activities each student had noted in excel, colour-coding each activity. I printed these coloured visualizations of each student's translation process and cut them into slabs of paper. I used these coloured slabs the following lesson as a starting point for reflection on the translation process. I found that, after some initial annoyance, most students were not disturbed by the beeps and they found the coloured slabs insightful. This interval assignment is very easy to perform and highly effective as an instrument for reflection on the translation process.

The second lesson aims at 1) development of knowledge of the translation process (SRSD stage 1A) and 2) development of metalanguage for reflection on the translation process (SRSD stages 1B and 2). Students analyse, compare and contrast their translation process, discussing different types of translation and defining text coherence. Students are provided with a visual representation of their translation process by colour-coded slabs of paper. In order to create the awareness that different translation processes exist, but that revision at some point is always part of a complete translation process, students analyse their process and compare and contrast it in pairs and with the translation process as described by Breedveld (2002). Students will be made aware that different types of translation[122] exist, to convey that translation is not an activity with only one correct outcome. To highlight that, students perform an assignment to label three translations of one fable as equivalent (literal), communicative or literary. To keep the lessons clear, I decided to limit the scope to these three types of translation. Students will be made aware that Latin translation classroom practice aims at the production of an equivalent translation, while the target text must always be a coherent Dutch text as well.

The third lesson aims at 1) development of source-text comprehension skills by exploring the genre of the fable (SRSD stage 1B) and 2) development of target-text revision skills (SRSD stage 3). The lesson starts by repeating the phases of the translation process, the types of translation and the definition of a coherent text. Students individually mark the building blocks of a fable in Dutch translation and visualise the story by drawing the fable in three pictures: situation, action, moral. Then, they compare their drawings in groups of three, each member having studied a different fable, and formulate genre characteristics and expectations on sequence of the building blocks. Students watch screencasts of two student translators from the eye-tracking study performing revision activities (modelling) and contrast and discuss the revision activities they observe in pairs. Having defined appropriate revision activities in

---

[122] *Chapter 1.*

classroom discussion, they formulate a personal revision strategy (development of metalanguage). They practice their revision skills, starting by individually assessing coherence of four target texts[123] by marking each sentence as + (coherent) or – (not coherent) without source text and revising the least coherent target text into a more coherent text. The lesson ends with an assignment to improve the personal revision strategy.

The fourth lesson aims at 1) development of text production skills (SRSD stage 4) and 2) performing the full translation process guided by assignments (scaffolding, SRSD stage 5). The lesson starts by repeating the phases of the translation process and the announcement that today's goal is to improve target-text production skills. Students watch text production activities of two student translators as models (screencast) and contrast and discuss the text-production activities they observed. Having defined appropriate text-production activities in classroom discussion, they formulate a personal target-text production strategy (development of metalanguage).

Then, students form groups of four, and each student translates one out of four different fables into a coherent target text, guided by six assignments:

1. Mark the building blocks (source-text comprehension, translation phase 1: orientation).
2. Summarise each building block focusing on meaning, draw the story in three pictures in the order: situation, action, moral (source-text comprehension),
3. Write a translation of each building block in the order of the source text (translation phase 2: TT production first draft).
4. Exchange your first draft with a peer and draw his/her building blocks, discuss differences between drawing and first draft with a peer (source-text comprehension and development of metalanguage).
5. Write a more fluent target text (translation phase 4: TT production second draft).
6. Exchange your second draft with (another) peer, revise your peers text on text coherence by marking each sentence as + (coherent) / – (not coherent).

The lesson ends with two questions encouraging reflection[124] and the transfer of the skills they learned to their regular Latin classes.[125]

The fifth lesson is aimed at 1) developing source-text comprehension skills (SRSD stage 1B) and 2) going through the complete translation process without the guidance of assignments (i.e. without scaffolding, SRSD stage 6). The

---

[123] *Translations of 'the Fable of the Mask' that were produced by eye-tracking study participants. The students translated this fable themselves in lesson 1 for the interval scoring.*
[124] *"This is what I learned in the past lessons about the translation process:…"*
[125] *"This is how I'm going to use this knowledge in the future when I translate Latin texts:…"*

lesson starts by repeating the phases of the translation process and the genre characteristics. Students apply their knowledge of the genre characteristics to a new fable (presented in Latin and Dutch) and explore two linguistic features 1) the use of tenses in situation-action-moral and 2) the use of direct speech in situation-action-moral. Then they translate an unknown, extensively annotated fable into coherent Dutch. For this independent performance of the translation strategy, a written reminder of the phases and appropriate translation activities for each phase will be provided and can be consulted if necessary.

4   SETTING UP AN EXPERIMENTAL STUDY: CONTROL CONDITION

An experimental study aimed at improving the text coherence of students' translations needs a control condition to establish its effect. Therefore, I designed lessons for a control condition as well. These lessons are based on the current 'business as usual' in teaching Latin translation, which is guided by the belief that a morphologic and syntactic understanding of the Latin source text leads to a correct translation.[126] Therefore, lessons in a control condition focus on linguistically analysing Latin source texts on sentence level.

In the design of the lessons for the control condition I applied Merrill's principles, as the intention was to offer students good lessons. The recommendations of Graham and Perin apply specifically to writing instruction, as a result of which they should not be applied to the 'business as usual' lessons in the control condition.

The lessons in the control condition are equal to the lessons in the experimental condition in 1) number, 2) work forms, and 3) texts and annotations, to be certain that differences in outcome can be attributed to the lessons. Obviously, the approach to the source texts in both conditions differs fundamentally. In the lessons in the experimental condition source-text comprehension is merely the starting point for target-text production and revision, while in the control condition the understanding and analysis of the source text's linguistic structure is leading in the lessons ('business as usual'). Error analysis of translated texts is a method which is occasionally used in regular classes to help students reflect on the types of errors they make while translating (Bekker & Van Oeveren, 2008). Error analysis focuses on linguistic errors on sentence level. Students categorise their errors as semantic errors (wrong meaning of word), morphologic errors (e.g. plural instead of singular) or syntactic errors (wrong analysis of sentence structure).

Error analysis is applied after completing a translation task: students identify what type of translation error they made to be able to avoid that type of

---

[126] *Chapter 1.3.*

error in the next translation task. The exercise is aimed at improved knowledge of the source language and builds on the belief that translation is primarily a linguistic task, concerned with the morphologic and semantic understanding of the text.

It differs fundamentally from revision in that error analysis is not aimed at improving the current target text as a part of the translation process, but is aimed at improved translation skills in the next translation task. I chose error analysis as an element of the lessons in control condition to provide students with a sense of systematic procedure to improve their translation skills and to create a certain uniformity in the rather vague 'business as usual.' I made sure that the same source texts and work formats were used in both conditions.

5   AN EXPERIMENTAL STUDY
TO IMPROVE TARGET-TEXT COHERENCE

I aimed at testing whether the lessons in the experimental condition would have effects as to the target-text coherence of the target texts students produced. Therefore I had to use a pre-test to assess initial text coherence, and a post-test to assess an effect for target-text coherence. I needed the same measurements for a control condition, to be able to assess whether effects were related to the lessons.

5.1 Research Question and Hypothesis

The process-oriented translation strategy underlies the main research question: will instruction of a strategy that focusses on the translation *process* and that is aimed at the development of metalanguage on translation, coherent target text production and revision skills lead to the production of more coherent target texts than a traditional instruction, focusing on the translation *product* and aimed at developing knowledge and linguistic understanding of Latin?

I hypothesised that *if I use methods such as SRSD, modelling and observational learning*, which have been tested and found effective in writing education*, to instruct upper secondary Latin students on the translation-process strategy*, focusing on target-text production and target-text coherence, with explicit attention to target-text revision, these students will develop their knowledge of the translation process and use of metalanguage, as a result of which *the coherence in the target texts will improve.*

5.2 Measurement Design

I used fable 1 *(Ovis, Cervus et Lupus*) and fable 2 (*De Vulpe et Uva*) for pre-test and post-test. I previously used the latter of the two fables for the eye-

tracking study. The experimental study accommodates four groups (A and B experimental, C and D control). Thus, I could check differences in difficulty of the source texts by crossing the source texts I used for pre-test and post-test in both conditions. *Table 28* shows the design of the experimental study.

Table 28: Design of the Experimental Study

	Session 1	Session 2-5	Session 6
Experimental A B	Pre-test: Text 1 Text 2	Intervention	Post-test + learner report Text 2 Text 1
Control C D	Pre-test: Text 1 Text 2	Business as usual	Post-test + learner report Text 2 Text 1

As I wanted the experiment to be as compact as possible, I used the first lesson partly for the pre-test and the last lesson, the independent performance, as a post-test. The common and specific features of the lessons in experimental and control condition are shown in *Table 29*.

Table 29: Common and Specific Features of Two Conditions[127]

Session	Experimental	Control
1 Pre-test Common	Translation of fable 1 or 2 on computer (I) Evaluation (I)	
Specific	1. Scoring the translation-activities of fable 3 in 30s intervals (I)	Translation of fable 3 (I)
2 Common	Assignment on types of translation: equivalent, literary, communicative (fable 4) (P) Peer-comparison (P) Feedback (P) Working in pairs / individually (P/I)	
Specific	1. Analysis of interval scores (I) 2. Comparing translation process: own with translation professional translators (I)	1. Checking translation fable 3 (I) 2. Error analysis of translation of fable 3 (I)

---

[127] *(I) = Individual ; (P) = Pair work; (C) = Whole class; (G) = Group work*

	3. Comparing interval scores, giving and receiving feedback (P) 4. Participating in classroom discussion on translation process (C) 5.1 Listening to introduction of the concept 'text coherence' (C) 5.2. Assessing text coherence of 4 translations of fable 3 (P) 5.3 Ranking a translation on coherence (G) 5.4 Discussion of the ranking (C)	3. Comparing error analysis, giving and receiving feedback (P) 4. Participating in classroom discussion on commonly made errors, discussion of relevant grammar (C) 5. Comparing results of assignment on types of translation (G)
3	Reflection on translation habits (I) Working individually (I) Carrousel assignment (G/I)	
Common		
Specific	1. Text comprehension Carrousel (G) 1.1 Listening to introduction on fable genre: building blocks (moral, situation, action) (C) 1.2 Marking moral, situation, action in Dutch translation of fable A (5), B (6) or C (7) (I) . 1.3 Drawing a cartoon in three pictures of the translation of the same fable as 1.2 (I) 1.4 Peer comparison of marking and drawing (G) 1.5 Formulating expectations of text coherence (G) 2.1 Observation of text revision (screencast) (I) 2. Formulating feedback on revision activities in screencast (P) 3. Participating in classroom discussion on text revision (C) 4.1 Revision of least coherent translation (P) 4.2 Comparing revised texts (P) 4.3 Peer feedback on revised text (P) 5. Participating in discussion on revision (C) 6. Online evaluation in Socrative (I)	1.Translation Carrousel (G) 1.1 Translation of fable A (5), B (6) or C (7) (I) 1.2 Checking translation (B, C or A) of a peer with a model (being not familiar with source text) and making error analysis of that translation (I) 1.3 Drawing a cartoon in three pictures of the translation of fable C, A or B (I) 1.4 Discussion of the errors that were made in the three fables' translations: formulation of tips and tricks (G)
4	Reflection on translation habits (I) Classroom discussion (C) Working individually (I)	
Common		
Specific	1. Observation of text production (screencast) (I) 2. Formulating feedback on text production activities in screencast (P)	1.1 Translation of fable 8 (I) 1.2 Reviewing the translation and discussing moral and relevant grammar (C) 1.3 Making an error analysis (I)

	3. Participating in classroom discussion on text production (C) 4. Applying knowledge of the full translation process (guided: marking building blocks, drawing, comparing, writing chunks, writing text, revising) translating one fable (fable 9, 10, 11 or 12) (I and P) 5. Participating in classroom discussion of both fables (C) 6. Listening to explanation of linguistic characteristics in Latin text related to the three building blocks (C)	2.1 Translation of fable 9 (I) 2.2 Checking own translation with a model (I) 2.3 Making error analysis (I)
5 Common Specific	Reflection on translation habits (I) Working individually and in pairs (I/P)	
	1.1 Marking moral, situation, action in fable 13 (with translation given) (I) 1.2 Marking linguistic characteristics in Latin text (I). 2. Applying the knowledge of the full translation process (with a reminder) translating fable 8 (I) 3. Using a (projected) slide as a reminder of phases in translation process (if needed) (I) 4. Peer feedback on revised text (P)	1. Translation of fable 10, 11 and 12 (P) 1.1 Participating in classroom discussion reviewing the translation and discussing moral and relevant grammar (C) 1.2 Making an error analysis of the translation of one of the three translated and discussed fables (I)
6 Post-test Common	Translation of fable 2 /1 on computer (I) Learner report	

6   CONCLUSIONS

Previous research suggested that to achieve target-text coherence four components were important: reflection on the translation process through metalanguage, text comprehension, text production and revision. I operationalised these elements and used them as components for my process-oriented translation strategy: (1) the skill to reflect on the translation process, (2) the skill to use metalanguage for reflection on translation decisions to consciously switch between types of knowledge and strategies, (3) source-text comprehension skills, (4) target-text production skills, and (5) target-text revision skills. The instruction of this process-oriented translation strategy should be based on both SRSD (Harris & Graham, 1996, 2002) and four recommendations for writing instruction (Graham & Perin, 2007).

I designed an experiment for two conditions, consisting of a pre-test, four lessons and a post-test. In the experimental condition, the process-oriented

translation strategy is taught with an emphasis on the development of text pro-duction and revision skills, which means that the focus in the translation pro-cess gradually shifts from source text to target text. Therefore, the lessons in this condition are primarily target-text oriented as opposed to source-text ori-ented, though comprehension of the source text remains the starting point. This target-text oriented approach to translating Latin would be novel in sec-ondary education, as the focus in the Dutch classroom[128] traditionally lies more on the source text and its morphologic and syntactic analysis than on target-text production. In the control condition students translated fables in a setting similar to 'business as usual', focusing on linguistic analysis of the Latin source text with the support of error analysis.

I set up a trial run for the experimental study to assess whether it could be performed as planned and whether students found the experimental lessons to be sufficiently clear and did what they were expected to do in both conditions. The trial run of the experiment and the improvements made as a result of it are extensively discussed in chapter five of this dissertation.

---

[128] *Chapter 1.3.2: The Translation Assignment in Latin Class.*

CHAPTER 5
EXPERIMENTAL STUDY -
TRIAL RUN AND IMPROVEMENTS

1   INTRODUCTION TO THE TRIAL RUN

The previous chapter described the development of a process-oriented translation strategy, design principles and design of lessons to teach that strategy. These lessons were made part of an experimental study to improve target-text coherence in students' target texts. The experiment consists of lessons focusing on the development of text production and revision skills as well as the development of metalanguage in an experimental condition, and lessons characterised as 'business as usual' in a control condition. The present chapter describes the trial run of this experimental study in sections two to six, as well as its redesign based on the findings in trial run (Section 7).

We[129] formulated ten questions the testing had to answer concerning fidelity of implementation (O'Donnell, 2008),[130] results as to the intended outcome of the experiment, and procedures. High fidelity increases reliability that effects can be attributed to the lessons. The first set of questions related to fidelity of implementation:

1. Did the source texts that were selected for the lessons present sufficient starting points for discussing target-text coherence?
2. Did students in both conditions perform the tasks as instructed?
3. Did students find the experimental lessons to be sufficiently clear?
4. Did teachers in both conditions find the instructions to be sufficiently clear?
5. Was the duration of the lessons as planned?

The answers to these questions can be summarised by the conclusion that in the experimental study fidelity was low. It seemed that insufficient clarity and the bad timing of the experiment at the end of the schoolyear, as well as classroom dynamics have led to low time on task and low fidelity. Especially in the control condition, the teacher instruction turned out to be insufficient and led to differences between the two groups in the control condition. Measuring duration was an important point of interest for the testing. When designing the experiment, we could only guess at the time it would take students to perform

---

[129] *The research discussed in the present chapter as well as in Chapter 6 was performed in close collaboration with my co-supervisor Suzanne Adema and my supervisor Gert Rijlaarsdam. Therefore, I will mostly use the plural 'we' in these chapters.*
[130] *I discuss O'Donnell more elaborately in Chapter 5.6.1.*

the tasks. Therefore, the testing of the experiment was also aimed at measuring how much time students needed to perform the tasks. We analyse the problems with fidelity in more detail in section six. The redesigned experiment explicitly addresses these problems. The source-text selection, however, was successful: the testing confirmed that translating fables offers a good opportunity to address and discuss text coherence.

The second set of questions related to the intended outcome:
1. Were the source texts used in pre-test and post-test equally difficult?
2. Was an effect of condition measured as to the target-text coherence in the post-test?
3. Was an effect of condition measured as to the translation process in the post-test?

It is difficult to find Latin source texts that do not differ in difficulty. We aimed at establishing whether differences between texts were significant by crossing the fables in pre-test and post-test. Both questions about effects of condition cannot be answered due to fidelity issues as a result of which effect measurements could not be reliably established for this pilot.

The third set of questions related to procedure:
1. Was the technical set-up (hardware and software) in pre-test and post-test adequate?
2. Was the length of the source texts used in pre-test and post-test adequate?

The procedural aspects of the experiment are discussed in sections 7.5 and 7.6 and have led to procedural improvements for the redesigned experiment.

We used classroom observation, personal communication with teachers and students, registration of attendance and lesson books, a questionnaire, and a learner report to collect data for the testing. These instruments are more elaborately discussed in section four.

The three design principles for the experimental condition as formulated[131] were asserted. The trial run led to four major improvements:
1. a new research design;
2. for both conditions, adaptation of:
   a. the duration,
   b. the grouping of students,
   c. teacher instructions;
3. a newly formulated design principle (control condition);
4. a thorough redesign of the lessons in the control condition based on the new design principle.

The experiment was redesigned as an experimental study with two competing conditions (Section 7).

---

[131] *Chapter 4.3.1.*

At the end of chapter five we add section nine to report on the lessons in the experimental study of chapter six. This section aims at being an inspiration for teachers as well as an anecdotal approach to the experimental study that would be inappropriate in chapter six.

2   PARTICIPANTS

We planned to test the lessons, pre-test and post-test[132] in a school setting during the last weeks[133] of the academic year 2016-2017 in the pre-exam level (grade 11) of the school where I work as a teacher. All four Latin teachers in grade 11 were willing to test the intervention in their regular classes. The experiment was tested in four groups by four different teachers. *Table 30* shows the distribution of participants over groups and conditions.

Table 30: Distribution of Participants over Conditions

Experimental (n= 52)	A (n=25)	B (n=27)
Control (n= 51)	C (n=28)	D (n=23)

Teachers A, B and C were regular and experienced teachers, with established working relationships with their classes. Teacher D was a replacement teacher, whose relationship with the class was still developing and occasionally problematic as to classroom discipline. All students were 16-18 years old.

3   MATERIALS AND MEASURES

For each session a lesson book was provided, containing explicatory texts and classroom assignments. Teachers were provided with a written lesson plan for each lesson by email. *Table 31* shows how and when we collected data to measure clarity, text coherence, fidelity and duration of the experiment.

We collected data on *fidelity* by:
- Registering *attendance*, as participants had to be present to be able to do what they had to do.
- Scoring the lesson books: to what extent did participants fill in the assignments?
- Assigning learner reports at the end of the experiment: what did participants report to have learned?

We collected data on *clarity of the lessons* by:

---

[132] *See Chapter 4, Table 28: Design of the Experimental Study.*
[133] *1-26 June 2017.*

- *Lesson observation*: I was able to observe and report teacher and student behaviour for most classes in group A, C and D in notes, if my schedule permitted. I taught group B myself.
- *Personal communication*: The teachers and I discussed clarity of the teachers' instructions before and after classes in an informal setting. The teachers described the lessons I did not observe in person afterwards.
- *Questionnaire:* At the end of session three in the experimental condition participants were asked to complete an online questionnaire checking whether they thought the experimental lessons were clear enough.

Table 31: Variables, Instruments and Measures

Variable	Instrument	Measurement occasion
Fidelity	Attendance registration	Class
	Lesson books registration	Class
	Learner report	Post-test (6)
Perceived clarity	Lesson observation	Class
	Personal Communication	Spontaneous communication with teachers
	Questionnaire	Experimental condition (3)
Duration	Time measuring	Class
Target-text coherence	Comparative judgement (D-PAC)	Pre- and post-test (1, 6)
Translation process	Screencast	Pre- and post-test (1, 6)
Technical set up	Observation	Pre- and post-test (1, 6)

Table note: numbers refer to session.

We measured *duration* by *registration:* We registered how many 70-minute classes were needed for each designed lesson unit.

We measured *target-text coherence* by *D-PAC*: the online tool for comparative judgement discussed in Chapter 2.2.4. The PDF-files with the translations were uploaded in D-PAC and the texts were rated by teachers of Dutch (4), English (1), History (2), Latin (10), and Philosophy (1).[134] Three teachers were removed because they made too few comparisons (1 to 6 comparisons). The others performed a total of 1027 comparisons, with the average number of comparisons being 68,5 (r = 76).

---

[134] *Appendix A, Table LV.*

We collected data on the *translation process* by *Screencast:* We wanted to record target-text production to observe participants' revision activities in the translation process using screencast software. We selected the online screencast programme *Screencast-O-Matic*, as it is user friendly and we received positive reviews of it from peer researchers. This programme had a free version with a limited recording capacity of 30 minutes and a payed version with unlimited recording time. The free version provided enough recording time for the translation of one fable.

We tested the *technical set-up* in pre-test and post-test by *observing* the functionality of:
- *Word* on the desktop computers in one of two computer rooms of the school.
- *Email,* as students emailed their text in PDF to the main researcher.
- *Screencast recording*, as students recorded their translation activities as screencast.

Students were provided with a Latin-Dutch dictionary (Pinkster, 2008) in hardcopy.

4   PROCEDURE

The experimental study consisted of two test sessions and four instruction sessions of 90 – 120 minutes each (Chapter 4, *Table 29*). All six sessions had to be tested within the regular schedule (lesson units of 70 minutes; 60 minutes on Tuesdays). Therefore, the experimental study had to be performed in more than six regular classes and sessions and classes overlap only partly.[135] The experimental sessions in group A were halfway finished when group B started, so some improvements in the material and instructions were made for group B. All tasks in experimental and control condition were designed to be performed in class. Participants were not required to do homework.

The testing was supervised in each group by the teacher. Fables for the tests could easily be translated in the allotted time. The teachers reported that the participants experienced the tests as rather short. Some issues with the annotations of the Latin source texts were reported, e.g. students in both conditions asked questions about the annotation for '(*advocare*) *sponsum*' in Text 1, as they were not familiar with the meaning of the Dutch 'borg staan voor' (*to warrant*). These issues were noted for improvement.

---

[135] *As shown in Appendix A, table LVI and LVII.*

5   DATA COLLECTION

The teachers collected the lesson books at the end of each class and I administrated whether all books were there and all assignments were made. All teachers registered attendance in their classes.

During the experiment, a total number of thirty-one classes was taught, of which I observed nineteen, making notes during class. Due to the school schedule, five classes were taught in my absence. These lessons were evaluated afterwards with the teachers. I taught seven classes myself, making notes afterwards.

The crossing of Text 1 and 2 in pre-test and post-test went according to plan in groups A, C and D. In the pre-test of group B, I accidently distributed the same tests as in class A, thus interfering with the crossed pre-test-post-test design for the experimental condition.

When we used *Screencast-O-Matic* we found that the firewall in the school internet environment prevented its use. Therefore no data were collected of the translation process. The learner reports were filled in by all participants who attended the post-test.

6   DISCUSSION OF RESULTS

As announced in section one we formulated ten specific questions to test the experiment, concerning fidelity, outcome and procedures. The first question was whether the source texts that were selected for the lessons presented sufficient starting points for discussing target-text coherence. The trial run confirmed that translating fables offers a good opportunity to address and discuss text coherence. As expected the clear structure of these texts offers handles to students for translation and interpretation, while the storyline and the moral are abstract enough to challenge the students' text comprehension skills. Lessons in both conditions offered genre knowledge and an introduction to different types of translation as common features and students responded favourably to these assignments. The present section discusses the remaining nine questions regarding fidelity, outcome and procedures, as well as solutions for the problems we encountered.

6.1 Discussion of Fidelity

According to O'Donnell (2008), five principles for fidelity of implementation are:

> (a) *adherence*, whether the components of the experiment are being delivered as designed; (b) *duration*, the number, length, or frequency of sessions implemented; (c) *quality of delivery*, the manner in which the

implementer delivers the program using the techniques, processes, or
methods prescribed; (d) *participant responsiveness,* the extent to which
participants are engaged by and involved in the activities and content
of the program; and (e) *program differentiation,* whether critical fea-
tures that distinguish the program from the comparison condition are
present or absent during implementation (2008:34).

The trial run showed that improvements were required on the tested aspects
of fidelity[136] that we numbered 2 to 5 in the introduction to this chapter.[137]
- Did participants in both conditions perform the tasks as instructed? (*ad-herence,* O'Donnell, 2008).
- Did participants find the experimental lessons to be sufficiently clear? (*participant responsiveness,* O'Donnell, 2008).
- Did teachers find the experimental lessons to be sufficiently clear? (*quality of delivery*, O'Donnell, 2008).
- Was the duration of the lessons as planned? (*duration,* O'Donnell, 2008).

The data showed that students in both conditions did not perform the tasks as
instructed. Not one lesson book was filled in completely, the variance in at-
tendance[138] was large, and we observed low *time on task*. Students reported
that they did not find the experimental lessons to be sufficiently clear. Class-
room observation showed *off task* activities and it was noted that students
asked questions about the purpose of the assignments.[139] Classroom-manage-
ment issues as well as students' social interaction were observed to lead to *off
task* activities.

Teacher A expressed insecurity about teaching the experimental lessons.
Teacher B, as the designer of the experiment, unsurprisingly reported no prob-
lems as to clarity. Although teachers C and D did not explicitly express that
they experienced problems, the observations showed differences in approach
that caused differences between lessons in the control condition. Apparently,
the lesson plans left room for different approaches. 'Business as usual' in
Latin classrooms has never been prescribed or described for the Dutch context
and may differ from teacher to teacher and classroom to classroom (Chapter
1.5).

The duration of the lessons varied between groups and conditions. It
seems acceptable to redesign the experiment for six sessions of 100 -120
minutes each (approximately 10.5 h. in total). The experimental condition in
its current form apparently needs less time: six sessions of 80 minutes each

---

[136] *We did not collect data on* program differentiation *(O'Donnell, 2008), as we were confi-
dent that lessons teaching the newly developed process-oriented translation strategy would
differ substantially from 'business as usual'.*
[137] *page 123.*
[138] *i.e. O'Donnell's* duration.
[139] *i.e. O'Donnell's* responsiveness.

should be sufficient (approximately 8 h. in total). However, to avoid differences between conditions, the lessons in in both conditions should be approximately the same length in the redesigned experiment.

Some solutions to improve fidelity in the redesigned lessons are similar for both conditions, while some regard each condition separately. Increasing clarity and reducing *off task* time is an important aim in the improvement of the material in both conditions, even though the reasons for *off task* behaviour differ per condition. I will first discuss the explanations for low fidelity in both conditions, followed by the condition-specific issues.

Causes for low fidelity in both conditions. We identified six causes that could explain the deviations of the planned implementation of the lessons in both conditions:
1. the duration of the project and the overflow of lessons in different classes,
2. the variance in attendance,
3. social interaction in the classroom,
4. the lack of clarity as to the purpose of assignments,
5. teacher-paced lessons, leading to 'waiting time' for students,
6. complex collaborative assignments, leading to 'waiting time' for students.
In a revised experiment these issues can be avoided by improving the setting (issues 1-2) and the lessons (issues 3-6). A setting of the actual experiment in classes that are not restricted in time by the school schedule would eliminate the first issue. The last weeks of the schoolyear were lacking in structure, which may have resulted in the variance of attendance (issue 2). It was therefore not the best time to test the lessons. Variance in attendance in the actual experiment can be avoided if students participate voluntarily and are being rewarded for their participation in an experimental study that takes place at the beginning of the school year. Using groups of participants that have no shared social classroom context and teachers that have no other relation to the participants than being a 'translation trainer' could increase a sense of professionality and result in more *on task* behaviour (issue 3). The redesigned lessons in both conditions have to provide clarity of purpose (issue 4). The lessons must reduce waiting time for students by offering participants more autonomy as to their pace and by offering less complex collaborative assignments (issues 5-6).

Causes for low fidelity specific to the experimental condition: Specifically in the experimental condition, o*ff task* activities also occurred when students felt insecure, either about what they had done, about what they were doing or about what they had to do next.

This problem can be addressed with a redesigned intervention which more explicitly offers room for discussion with the group as a whole.[140] The group discussions - without labelling contributions to the discussion as *correct* or *wrong* - provide a sense of 'closure' or 'clarity' for students, while maintaining the necessary insecurity. Moments of classroom reflection on *where we are and where we are going* would be a natural beginning for this discussion. The right moment for the discussion must be found depending on the progress the group as a whole is making.

The improvements in the material will not be sufficient if the teacher who is supervising the process does not have a clear idea of the process and aim of the intervention. Focus and overview of the students in the experimental condition was impaired by insufficient teacher instruction on reflection and process. The teacher must be able to explain the purpose of the assignments. Therefore, more detailed teacher background information must be included on the purpose and overall process of the experiment.

Causes for low fidelity specific to the control condition: In the control condition, *off task* activities also occurred when students were giving up on error analysis and when students were bored. This can be addressed by a redesign where error analysis is more explicitly taught and more variation is offered in addition to the translation-correction-error analysis routine.

Also, we should reduce the number of fables we offer, as it took students longer than expected to translate the fables in class. The lack of variation in assignments students experienced can be solved by adding some questions on the texts they have translated, which is quite similar to regular teaching material. It is important that lessons in this condition remain primarily focused on the analysis of the Latin text and ST-TT equivalence, as the different focus between conditions (process-oriented and product-oriented) must be maintained.

In the control condition, teachers' instructions on the purpose of the assignments and the overall process must be improved as well, to keep questions about the annotations or assignments from arising and to increase the similarity in the execution of the lessons. It could also be useful to provide Power-Point presentations of the texts that highlight the focus of text discussion in class.

---

[140] *After both observation tasks, but also after the + / - method and revision assignment.*

6.2 Discussion of Outcome

We wanted to test whether the source texts used in pre-test and post-test were equally difficult (Section 1, question 6), but cannot offer firm conclusions due to the failed crossed pre-test-post-test design. The results of group D, the only group where the test order of fables 2-1 was realised, showed lower average coherence scores in post-test than pre-test, which suggests that fable 1 was slightly more difficult to understand than fable 2. However, this suggestion is highly speculative for two reasons. First, statistical power was lost as a result of the failed execution of the crossed pre-test/post-test design of the tests.[141] Second, group D suffered a problematic teacher-student relationship, which means a teacher effect cannot be excluded. Therefore, no firm conclusions can be drawn. To improve the reliability of the tests a second fable will be added to the tests in the redesigned experiment (see also Section 6.3).

Given the problems we experienced with fidelity (Section 6.1) as well as with procedure (Section 6.3) no clear results on effects of conditions are available.

6.3 Discussion of Procedure

The technical set-up of hardware and software in pre-test and post-test (Section 1) were adequate as to the use of the computer rooms, Word, and the emailing of the target texts. The use of *Screencast-O-Matic* turned out to be problematic due to the firewall of the school environment.

The final question we wanted to answer by the trial run was whether the length of the source texts used in pre-test and post-test was adequate. It is impossible to find two fables that present exactly the same difficulties for the tests. The reliability of a test relies partly on its length. The fact that in the tests both fable 1 and fable 2 were perceived as short offers the possibility of adding to the power of the tests by extending them with the translation of a second fable. The translation of two fables instead of one in each test will make the tests more reliable.

7   REQUIRED IMPROVEMENTS OF THE EXPERIMENTAL STUDY

The trial run showed that the experimental study required improvement, as became clear in the previous section. In the following sections, these improvements as well as the maintained features of the lessons will be described.

First, however, we must again discuss an important issue in the curriculum of Latin. As discussed in Chapter 1.4.1, a base of shared knowledge how to

---

[141] n = *80 for order fable 1-2;* n = *23 for order fable 2-1*

teach Latin translation is missing. The trial run disclosed that this lacuna led to fundamental problem with the lessons of the control condition and this problem was not sufficiently controlled by the use of error analysis as a tool. The assumption that 'business as usual' streamlined by the application of error analysis would provide stable control condition proved to be too optimistic. In this trial run it became even more clear than before (see Chapter 1.3): business as usual is not 'to teach translation', business as usual is 'to translate.' A research design as in the trial run with a control condition consisting of two separate groups featuring 'business as usual' presented a problem, because the differences between both groups in the control condition would be too big. A prescribed and more detailed teacher instruction could remove those differences. However, *any* prescribed lesson plan on how to teach translation would be a deviation from 'business as usual' for the individual teacher. The research design had to change.

This presented a problem: if teaching to translate by teaching how to perform an error analysis became the objective in the control condition instead of 'business as usual', new design principles on teaching this skill had to be formulated. This design principle is that error analysis must be explicitly taught as a means to improve ST-TT equivalence (see Section 9.2). Thus, the experimental study testing the process-oriented translation strategy evolved into an experiment with two competing experimental conditions:

Process condition: an experiment to improve target-text coherence. The lessons in this condition focused on the *translation process.* The design principles for the lessons in the process condition remained the same as for the initial experimental condition: lessons focused on the *translation process* and aimed at improving the participants' target-text coherence through 1) knowledge of the phases in the translation process, 2) awareness of their individual translation process and 3) the development of revision skills, writing skills, and a metalanguage about translating.

Product condition: an experiment to improve ST-TT equivalence. The lessons in this condition focused on the *translation product* and were aimed at improving the equivalence of source text and target text through knowledge of Latin morphology and syntax. The resemblance to 'business as usual' remained only in the sense that teachers mainly discussed these aspects of the Latin source text: students translated a Latin ST, the TT is discussed and improved afterwards in class. The lessons were redesigned on a new design principle: error analysis must be explicitly taught as a means to improve ST-TT equivalence (see Section 9.2).

7.1 Maintained Features of the Lessons

The present section describes in short what features of the tested lessons were maintained in the redesign of the lessons. The subsections 7.2 to 7.7 describe more elaborately what improvements were made as to duration, lessons, teacher instruction, tests, procedure and measurements.

    Four common characteristics of the lessons in both conditions remained: source texts, genre knowledge, introduction to different types of translation, and applied work forms.

    Source texts: The fables by Phaedrus we used as Latin source texts for the tests and lessons were satisfactory, as they met the required characteristics of being short, coherent and sufficiently complex texts.[142]

    Genre knowledge: The translation process is characterised by an orientation phase before writing. An aspect of this phase is activating genre knowledge. Genre knowledge was part of the intervention in both conditions; in the experimental condition as part of the process-strategy for translation, in the control condition as useful knowledge when translating fables. This was maintained in the redesigned lessons for process and product conditions.

    Different types of translation: In both conditions, participants were taught about three types of translation equivalent, literary and communicative. In the redesigned experiment we will use the same assignments concerning genre knowledge and types of translation but remodel the distribution over the lessons to a more compact form to improve clarity.

    Work forms: Participants worked individually and in pairs in both conditions. Classroom discussion was a part of each lesson and each lesson started with reflection on what was learned. As these features facilitate learning, they will be maintained in the redesigned experiment.

The common and specific features of all redesigned sessions, including all tests are presented compared to the original design in separate tables in Appendix B. In addition to these Appendices, I discuss the lessons in both conditions in section nine.

7.2 Improved Fidelity of Implementation: Duration

To avoid that effects could be attributed to a difference between conditions in the numbers of lessons, we decided that the number of lessons for the redesigned experiment should again be equal in both conditions, although the time the lessons required differed between conditions. We established that a total of four lessons would suffice to cover the content of the experimental

---

[142] *See Chapter 4 and 5.7.*

condition and adjusted the content of the control condition accordingly. The tests were added before and after the lessons in separate sessions.

7.3 Improved Fidelity of Implementation: Participant Responsiveness

To improve participant responsiveness, we redesigned the lessons in both conditions based on the results of the trial run.

Process condition: To improve participant responsiveness we redesigned the lessons for the process condition regarding the complex collaborative assignments. We redesigned all group assignments as pair work, retaining discussion as a feature while reducing the complexity in the organization of the assignment. By reducing the group assignments and organising pair work according to individual pace, we aimed at more student-paced lessons, to improve *time on task* and a sense of autonomy*, while clearly marking moments for class discussion.

Product condition: The lessons for the product condition were redesigned on the new design principle: error analysis must be explicitly taught as a means to improve ST-TT equivalence. To improve participant responsiveness and *time on task,* we designed more varied lessons in the *product condition* by adding questions about the fables for participants to answer after translating. These questions also served as a starting point for classroom discussion. This type of questions reflects common classroom practice, but it was omitted in the original design. We also decreased the number of fables that had to be translated. In this condition we redesigned most group assignments as pair work as well, retaining discussion as a feature while reducing the complexity in the organization of the assignment. We redesigned the complex group assignment in lesson 3 to a more straightforward group assignment where participants compare, select and present tips and tricks on translation.

7.4 Improved Fidelity of Implementation: Quality of Delivery

In both conditions, we aimed at improving clarity and fidelity by providing teacher instructions with (1) clearly marked moments for classroom discussions to provide participants more security, and (2) explicit topics for classroom discussion and preferred outcomes.

7.5 Improvements in Tests for the Redesigned Experimental Study

In the redesign both pre-test and post-test consisted of two fables each. For this purpose, we moved some fables from the lessons to the tests. This was convenient as we wanted to add variation to the lessons in the *product condition* and offer a smaller number of fables for translation practice. We added a

delayed test to the redesigned experiment, to measure the persistence of effects.

For pre-test, post-test and delayed test we chose two fables with approximately the same word count in total (107- 114 words). In some fables the moral was placed at the beginning (B), and in some fables the moral was placed at the end of the story (E). We improved annotations for the fable *Ovis, cervus et lupus* based on the remarks students made in the testing. *Table 32* shows characteristics for each fable of the redesigned tests.

Table 32: Characteristics of Fables for Tests

Test	Fable	Wordcount	Moral[143]	Title
Pre	1	55	B	Canis per Fluvium Carnem Ferens
	2	54	B	Ovis, Cervus et Lupus
Post	3	44	E	Canes Famelici
	4	72	B	Musca et Mula
Delayed	5	60	B	Canes et Corcodilli
	6	47	E	De Vulpe et Uva

7.6 Improvements in Procedure for the Redesigned Experimental Study

The redesign of the experimental study had to address the problems with attendance. In a school setting, 100% attendance is never guaranteed, which suggested performing the experiment outside school hours. Additionally, participants could be induced to attend all lessons and tests by offering an incentive.

The redesigned *product condition* had become an experimental study on ST-TT equivalence. Therefore, data on ST-TT *equivalence* in the tests had to be collected, which meant target texts additionally were colon rated. The rating could be performed by one expert rater, after establishing significant interrater reliability.[144]

The problems with the screencast could be avoided by using a different type of measurement for the translation process, namely measuring by means of a questionnaire. This provided insight into the participants' perception of their translation process. The data collection of the translation process by a screencast programme could provide a backup if needed, if its accessibility

---

[143] *B = at the beginning, E = at the end.*
[144] *The reliability of colon rating with one expert rater remains sufficient, as argued in Chapter 2.*

would be improved by preparing the installation of such a programme with the IT staff.

The questionnaire about perceived clarity in the *process condition* (after lesson 3) was extended to the *product condition*, as the clarity of the error analysis instructions needed to be measured as well. In the redesign I used questionnaires on three more occasions for more detailed fidelity measurements: after completing the lessons, after post-test and after delayed post-test. The total number of evaluations in both conditions in the redesigned experiment is four.

For the redesign we colour-coded the lesson-books' covers and tests to avoid mistakes in distribution of the materials.

7.7 Adaptation of Measurements for the Redesigned Experimental Study

The measurements for the redesigned experimental study in two competing conditions were adapted as presented in *Table 33* according to the changes made in the material.

Table 33: Variables, Instruments and Measurements

Variable	Instrument	Measurement occasion
Fidelity	Attendance	Class
	Lesson books	Class
	Questionnaire	Post-test (6)
		Delayed post-test (7)
Perceived clarity	Questionnaire	After session 3
Target-text coherence	D-PAC	Pre-test (1)
		Post-test (6)
		Delayed post-test (7)
Target-text equivalence	Colon rating	Pre-test (1)
		Post-test (6)
		Delayed post-test (7)
Translation process	Screencast	Pre-test (1)
		Post-test (6)
		Delayed post-test (7)

Table note: numbers refer to session.

8   CONCLUSIONS ON THE TRIAL RUN

We designed an experimental study consisting of an experimental condition to test the effect of lessons aimed at improving target-text coherence by the

instruction of a newly developed process-oriented translation strategy and a
control condition characterised as 'business as usual' (Chapter 4). The trial
run of this experimental study (Chapter 5) brought issues to light regarding
the stability of the control condition, where 'business as usual' did not provide
sufficient direction for the lessons, and regarding fidelity of implementation.
To address the first issue, we decided to redesign the study into an experi-
mental study with two competing experimental conditions: a *process condi-
tion* aimed at improving target-text *coherence* and a *product condition* aimed
at improving source text-target text *equivalence*. The issues on fidelity of im-
plementation regarded 1) quality of delivery, 2) participant responsiveness
and 3) duration (O'Donnell, 2008). The trial run suggested changes to improve
fidelity of implementation: 1) *quality of delivery* could be improved by more
specific teacher instructions, 2) *participant responsiveness* by reduction of the
complexity of assignments in both conditions and variation in assignments in
the *product condition*, 3) *duration,* by timing the experiment outside the
school schedule, earlier in the schoolyear.

9   REDESIGNED LESSONS AND THEIR EXECUTION

Chapter six will discuss in depth the redesigned experimental study we per-
formed in the year 2017-2018 and its effects. In that chapter, we will present
the whole experiment, following the usual structure of reports on experimental
studies, which provides little room for anecdotes or examples. In preparation
for chapter six, we aimed at providing a narrative for the lessons as well as
some examples of the work the participants delivered. This is what this section
aims at doing: it discusses the execution of the redesigned lessons in both con-
ditions of the experimental study. With the present section we hope to inspire
teachers of Latin translation by describing the lessons of both conditions and
illustrating these descriptions with examples from students' material. Section
9.1 shows some overlap with Chapter 4.3.3, as many assignments of the orig-
inal lessons remained. A detailed overview of the design and redesign of the
lessons is provided in Appendix B.

### 9.1 Instruction Sessions in the Process Condition

The sessions in the *process condition* meet the design principles as presented
in Chapter 4.3.3.

*Instruction session 1: Reflection on the translation process and genre charac-
teristics*. After completing the pre-test in the first test session, the participants

had performed an additional translation task[145] while scoring their translation activities every 30 seconds.[146] The first part of this first instruction session was devoted to development of knowledge of the translation process and metalanguage, followed by an introduction to three different types of translations and the genre characteristics of fables. We had decided to move all aspects of genre characteristics to the first instruction session, to emphasise that this knowledge is relevant at the beginning of the translation process (source-text comprehension). Participants were introduced to the following concepts: source text, target text, metalanguage, literary translation, equivalent translation, communicative translation, target-text coherence, and genre characteristics of fables, e.g. building blocks and linguistic characteristics.

At the beginning of this first instruction session each participant received a colour-coded slab of their individual translation process. The time spent on the translation task as well as the reported activities varied among participants, as is illustrated by two examples in *Figure 13*.

Figure 13: Translation Process Colour-Coding.
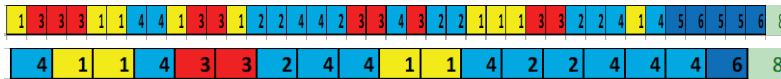
Legend[147]: Yellow = 1 reading /marking Latin; Red = 3 looking up a word; Light blue = 2 Formulate / 4 Write TT; Dark blue = 5 Improve TT / 6 Read TT; Pink = 7 other; Green = 8 ready.

This variation provided a starting point for reflection and discussion. Participants were asked to analyse the pattern of their activities, compare it with the translation process as described by Breedveld (2002), then compare it with a peer and give and receive written feedback on the process and its possible improvement before discussing it. To prevent the insecurity participants reported in the pilot, these tasks were followed by classroom discussion on the variations and qualities of translation processes and the concepts of source text, target text and metalanguage. Participants found it surprising to see the extent to which these processes could differ. Talking about the translation process and target-text quality without focus on knowledge of Latin seemed inspiring and new to many participants. The discussion of the translation

---

[145] *Translating the text of the fable of The Fox and the Tragedy Mask.*

[146] *For this method see Olive, Kellogg and Piolat (2002), discussed in Chapter 4.*

[147] *The categories are partly different compared to* Figure 10 *and* Figure 11 *as students in instruction session 1 could report also on not visible behaviour, e.g.* formulate. *I maintained yellow for* source-text reading *and red for* looking up words*, but I distinguished activities concerning source-text writing 1 and 2 (light blue) and revision (dark blue) as well as off task activities (pink) and having finished the assignment (green).*

process, without any evaluation of the source text or any form of analysis of ST-TT equivalence was specific for the process condition.

Students were introduced to the concepts of literary, equivalent and communicative translations by qualifying three translations of the same fable as such, followed by a classroom discussion introducing the notion that all translations must be at least a *coherent target text*. The concept that a target text should at least be understandable Dutch was an eye-opener for some participants.

The concepts of text coherence and genre characteristics in fables were introduced and the genre characteristics were explored through assignments. Participants learned the building blocks by analysing a translated fable, drawing a three-picture comic of the fable and comparing it with a drawing of a different fable by a peer. They struggled to draw the moral, realising that the moral is most difficult to draw as it is most abstract. They solved this problem in various ways, for instance by inserting the moral in a text balloon (*Figure 14*) or by having a text-only moral picture (*Figure 15*). In comparing the two fables they also noticed that the moral can occur either at the beginning or the end of a fable. After that, they identified linguistic characteristics for each building block by studying a fable in Latin with a Dutch translation alongside it. The search for linguistic characteristics was primed with questions. Thus, genre characteristics were argued and experienced by the participants instead of explained by the teacher. The lesson ended with a short written recap of the learned concepts.

Figure 14: Example 1: Fable 'Mustela et Homo' in Three Pictures.[148]

---

[148] *The Weasel and the Man.*

Figure 15: Example 2: Fable 'Mustela et Homo' in Three Pictures.[149]

Instruction session 2: Revision. The second instruction session was devoted to the last phase of the translation process: revision. The lesson consisted of two parts: 1) theory and observation of revision, 2) practicing revision to improve revision skills.

The lesson started with a recap of the two main concepts from the previous lesson (Merrill, 2002) 1) the four phases of the translation process and 2) target-text coherence. The recap was introduced by asking two questions,[150] followed by reflection on their own revision behaviour as analysed in the previous lesson and their expectations of the effects of revision for the improvement of the coherence of the target texts they write when they translate Latin. This reflection was prompted by questions from the lesson book.

Following this, participants focused on the revision phase of the translation process. After a short general introduction on the revision phase, participants individually viewed two eye-tracking screencasts of proficient students revising their translation to identify the activities these translators performed. They could watch the screencast as often as they liked via a link to YouTube, and many watched several times, as the eye-movements were very fast. They compared and contrasted their observations in pairs (Modelling: Graham & Perin 2007; Harris & Graham, 1996; 2002). This phase of the lesson was completed with a classroom discussion to prevent the insecurity participants reported in the pilot. Both screencasts were watched again with the entire group, while teacher and participants discussed what they had observed, leading up to the conclusion that text comprehension is leading for successful revision.

---

[149] *The Weasel and the Man.*
[150] *1) What do we mean by a 'coherent text'? 2) What are the four phases of the translation process?*

The second part of the lesson was aimed at practicing revision in two steps: first, participants individually ranked four target texts from more to less coherent using the +/- method, so marking coherence with + and incoherence with – (*Figure 16*).

Figure 16: Example. Assessing Coherence, Using the +/- Method.

Les 3 Tekstrevisie observatie   **Groep B**

1. **Pas de +/- methode toe (individueel):**

tekst 1
De vos over een masker van een tragedie

De vos had toevallig een masker van een tragedie gezien; nadat hij zich van alle kanten herhaaldelijk had gedraaid om deze, zei hij: "O, wat een verschijning! Het heeft geen schedel!"

Dit is gezegd van hen voor wie het lot eer en glorie toezwaaide, maar het gezonde verstand heeft afgenomen.

tekst 2
De vos en het tragische masker

Een vos had bij toeval een tragisch masker gezien. En nadat hij dat eenmaal hierheen en nogmaals daarheen had gedraaid, zei hij: "O, wat een grote schoonheid, maar hersens heeft het niet!" Dit is gezegd voor hen, aan wie het lot eer en roem toebedeeld heeft, maar gezond verstand heeft afgenomen.

tekst 3
De vos naar het tragische personage

De vos had toevallig een tragisch personage gezien;
die nadat hij eenmaal hierheen en opnieuw daarheen was gedraaid,
'O grote schoonheid, hij heeft geen hersenen!' zei.
Degene aan wie dit gezegd is heeft het lot eer en roem gegeven
en het heeft zijn gezond verstand weggenomen.

tekst 4
Een treurig persoon zag toevallig een vos;
Nadat hij eenmaal hierheen en opnieuw daarheen had gedraaid, zei hij:
'Oh hoeveel schoonheid! Hij heeft geen hersenen!'
Dit heeft gezegd dat aan wie het lot eer en roem heeft gegeven,
hij het gezond verstand heeft weggenomen.

2. **Dit is voor mij de volgorde van meest naar minst coherent:** 2 3 1 4

Second, they compared their ranking with a peer and decided which text was least coherent. They individually revised this text into a coherent text, then assessed its quality in pairs, again using the +/- method (*Figure 17*).

The lesson was concluded with an individual assignment to describe the revision activities they performed and a questionnaire. Some participants reported in the questionnaire that they found it difficult to analyse the screencast.

Figure 17: Example of TT Revision in Pairs, Using the +/- Method.
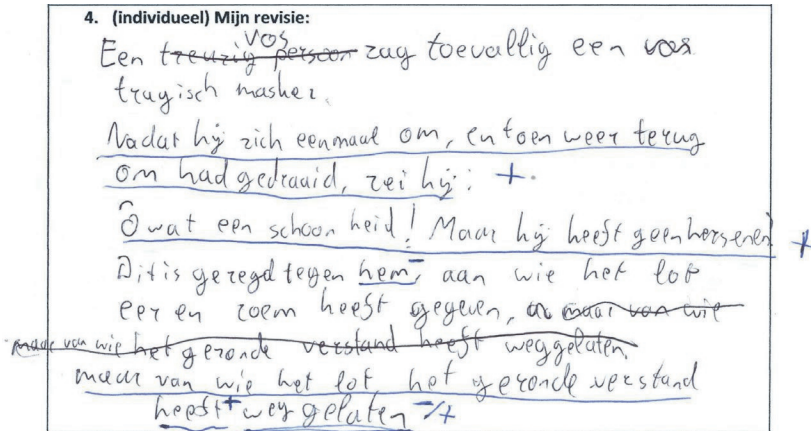
Instruction session 3: Target-text production. The third instruction session was devoted to the target-text production phase of the translation process. The lesson consisted of two parts: 1) theory and observation of target-text production, 2) practicing in order to improve target-text production skills.

The first part of the lesson followed the pattern of the second instruction session: recap of the translation phases, introduction of today's topic, reflection on participant's own target-text production behaviour, theory of coherent target-text production (first and second draft), observation of target-text production in screencasts, compare and contrast screencasts, classroom discussion.

The second part of the lesson involved practicing target-text production, which meant a fable had to be translated. Please note that before this moment hardly any Latin was discussed in the lessons in the *process condition*. To offer participants the opportunity of giving and receiving feedback on the target text untainted by prior knowledge of the story, two different fables were used for this assignment. Participants were guided through the phases of the translation process by performing smaller tasks, some individually, some in pairs:

1. Source-text comprehension (orientation):
-   Build understanding of the source text by reading it, by formulating its meaning and by marking the three building blocks.
-   Visualise the fable by drawing a cartoon of three pictures: situation, action and moral. (*Figure 18* and *Figure 19*)

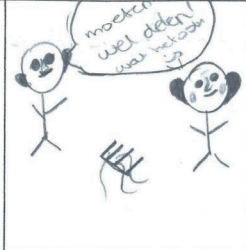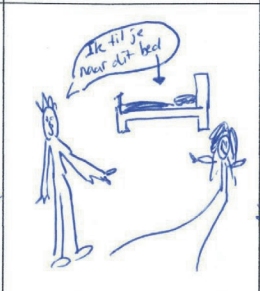*Figure 18: Example 1. Orientation by Visualisation: 'Duo Calvi'*[151]*.*

Figure 19: Example 2. Orientation by Visualisation: 'Mulier Parturiens'[152]*.*

2. Target-text production 1: Write a rough translation of each building block (situation, action and moral.)
3. Target-text production 2:
-   Find a peer who translated the same source text and compare. Build and formulate meaning.
-   Write the second draft of the target text.
4. Revision:
-   Find a peer who translated the other fable, read each other's second draft and assess coherence using the +/- method. (*Figure 20*)

---

[151] *Two bald men.*
[152] *A Woman Giving Birth.*

Figure 20: Example of Second Draft with +/- Peer Feedback.

-    Use the feedback on your second draft to revise your text into a more co-
     herent target text (*Figure* 21).

Figure 21: Example of Target Text of Figure 15 Revised.

To check whether all participants understood the story and moral of the fables
the assignment was followed by a classroom discussion of the moral of both
fables.

     This type of assignment offers scaffolding practice of the process-oriented
translation strategy. The experimental study featured this assignment only

once, as the intervention lasted only four lessons. To stimulate independent performance of the process-oriented translation strategy we would recommend integrating this type of assignment in regular Latin translation classes.

Instruction Session 4: The Complete Translation Process
The fourth and final instruction session aimed at going through the complete translation process without scaffolding, so without the guidance of assignments. In the lesson book, a written reminder of the four phases of the translation process and the appropriate translation activities for each phase was provided, which students could consult if necessary. The first part of the lesson followed the usual pattern: recapping the translation phases, introducing today's topic: going through the complete translation process independently or semi-independently and reflecting on one's own translation behaviour. During the performance of the translation task a slide was projected, that read "translate the fable into a coherent text," to remind all participants of the translation assignment. Students were instructed to use a pencil for phase one (Orientation, *Figure 22*) and two (Text Production 1), blue pen for phase three (Text Production 2) and red pen for phase four (Revision). *Figure 23* offers an example of phases 2 to 4.

Figure 22: Example of Orientation Activities (phase 1).

Figure 23: Example. Target-Text Productio[153]n

9.2 Product Condition

The lessons in the *product condition* are based on the teaching principles as formulated by Merrill (Chapter 1.4.1.2). In addition, the principle is met that error analysis must be explicitly taught as a means to improve ST-TT equivalence. Participants need to understand how to perform an error analysis and, more importantly, why error analysis is relevant for improving the performance of a translation task.

Instruction session 1: Introduction of error analysis and genre characteristics. After completing the pre-test in the first test session, the participants in the product condition had performed an additional translation task, translating the same text[154] as the participants in the process condition, but without the scoring assignment.

The first part of the first instruction session was devoted to the introduction of error analysis as a tool for improving translation skills, the second part to genre characteristics of fables and an introduction of different types of translations.

---

[153] *TT1: Blue ink; TT2: Pencil; Revision: Red ink.*
[154] *Translating the text of the fable of The Fox and the Tragedy Mask.*

After an introduction on the usefulness of error analysis to improve translation skills, participants received the target text they had produced in the additional translation task and reviewed their translation using a model translation and an error analysis form (*Figure 24*). In this form, they indicated errors concerning tense and number (morphology), constructions (syntax) and miscellaneous errors. Then they filled in an error analysis for a peer (*Figure 25*) and discussed the differences in pairs.

Figure 24: Example of an Error Analysis Form.

Zin	Tijd	Ev/Mv	Constructie	Anders, nl...
1				X (vertaling)
2				
3	X		X	
4				
5				
6				
7				
8				
totaal	1		1	(

Figure 25: Example of Error Analysis by a Peer[155]

Zin	Tijd	Ev/Mv	Constructie	Anders, nl...
1				x verkeerde vertaling 'musicam'
2				x tael…ig is herhaald van de 1º zin
3			x huc vertaald als onderwerp	
4				x verkeerde vertaling quanta
5			x quibus vertaald als onderwerp ipv 'tegen diegenen'	
6				
7				
8				
totaal				

---

[155] *Same ST, different TT as Figure 24.*

The comparison of *Figure 24* and *Figure 25* shows some variation in the error analyses. In these examples the translator (*Figure 24*) is less specific than the peer (*Figure 25*), who provides feedback in words instead of mere crosses and general terms. During the classroom discussion that followed, the participants discussed how to perform the error analysis as specifically as possible to improve their next translation. In the classroom discussion, the teacher was instructed to use the source text to explain the relative connection to the class. The discussion of the source text, including the evaluation of the ST-TT equivalence was specific for the product condition.

In the second part of the first instruction session, the three building blocks of the fable were introduced and applied to the fable of the Fox and the Tragic mask they had just corrected. This instruction session presented genre characteristics in a traditional way, by stating the fable's characteristics and importance and offering an assignment to apply the genre knowledge. It was not made explicit that there is a link between this type of genre knowledge and text comprehension.

Students were introduced to the concepts of literary, equivalent and communicative translations using the same task as in the process condition: qualifying three translations of the same fable. In a classroom discussion, the teacher stressed that translations in Latin class are aimed at *equivalence*. An important difference between conditions was that the concept of target-text *coherence* was not specifically addressed in the product condition.

The lesson concluded with a written statement, in which participants formulated their intention as to how to perform a translation task next time.

Instruction Session 2: error analysis and questions on source-text comprehension (I). The lesson started with a recap of the concepts introduced in the first instruction session: three types of translation (communicative, equivalent and literary), the fact that translations in Latin class are aimed at equivalence, and that fables consist of three building blocks: situation, action and moral.

The aim of the second lesson was to expand participants' experience translating fables by practicing translation and making error analyses. To add some variation to the lesson, source-text comprehension questions were added to the activities of instruction session 2.

The first part of the lesson consisted of the following translation task: each participant individually translated the fable "Asinus ad Senem Pastorem" (the Donkey to the Old Man) into "understandable and readable Dutch." Then they compared and contrasted their target text in pairs, discussed differences and, if necessary, improved their target text with red pen (*Figure 26*). Afterwards, participants answered source-text comprehension questions in pairs. The task was concluded by a classroom discussion of the source text, guided by a

PowerPoint presentation to achieve similarity in the discussion for the two groups in this condition. Additional improvements in the target text were to be marked with pencil.

Figure 26: Example of a Corrected Target Text.
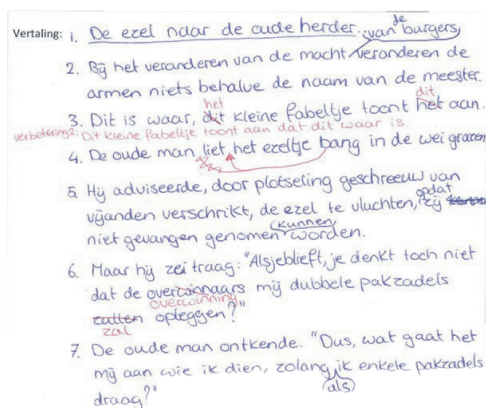
Figure 27: Example of Reflection on Errors and Plan to Avoid Those.
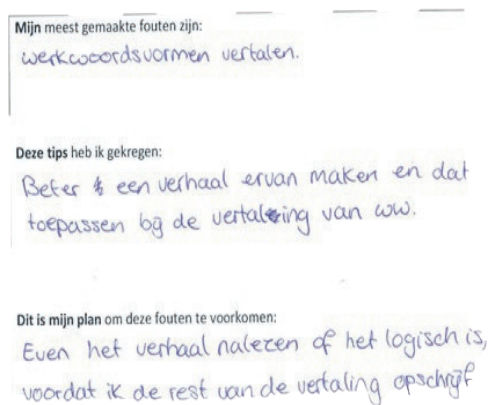
In the second part of the lesson, a second translation task was presented: this time participants translated "Mustela et homo" (The Weasel and the Man) into "understandable and readable Dutch," corrected their target text comparing it to the provided target-text model and made an error analysis of their target text. When they had finished, three or four peers formed a group, compared their errors and made a poster of the most common errors, adding tips to avoid those errors in the future. Each participant individually noted what they had learned in their lesson book (*Figure 27*), and each group presented their poster

to the rest of the class. The lesson was concluded with a questionnaire as part of the data collection (see Chapter 6.2.3).

Instruction session 3: Error analysis and question on source-text comprehension (II). The lesson started with reflection primed by two questions: 1) what have you learned in the previous lessons? and 2) what is your personal plan to improve your translations? Again, the teacher underpinned the usefulness of error analysis as a tool for reflection on translation errors and improving translation habits.

Participants performed two translation tasks during the third lesson. The first task was to translate "Duo Calvi" (Two Bald Men) into "understandable and readable Dutch." After translating this fable, participants corrected their target text by comparing it to the provided model and making an error analysis. Then, they were to find another participant to exchange error analyses with and evaluate the errors the peer had made and formulate feedback and tips. Each participant had to exchange with as many peers and collect as many tips as possible. *Figure 28* shows an example of three collected tips: note that using the PSOLMO-strategy is given as a tip.

Figure 28: Example of Collected Tips and Tricks.

After collecting tips and feedback, participants were asked to write the tips and feedback they found useful on a post-it and stick the post-it on the whiteboard in one of four categories: 1) knowledge of language, 2) translation strategy, 3) precision, and 4) other (*Figure 29*). This task was concluded by a classroom discussion aimed at improving participants' insight into the areas the tips and feedback targeted, so that they realised where room for improvement was to be found as well as what could lead to the improvement. The pictures

show that the quadrants 1) knowledge of language and 2) translation strategy contain the most post-its.

Figure 29: Post-Its with Categorised Tips.
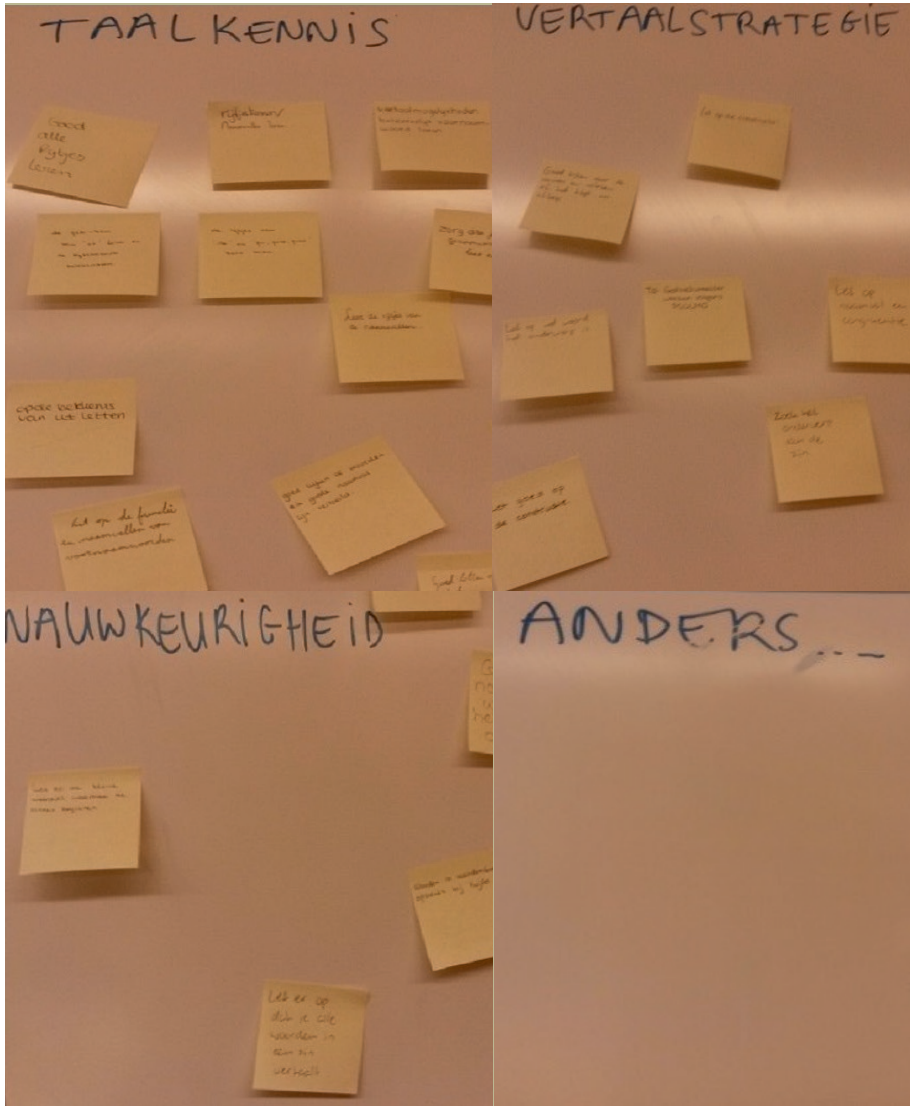
In the second part of lesson 3, participants translated "Mulier Parturiens" (A Woman giving Birth) into "understandable and readable Dutch" either individually or in pairs, as seen fit. This fable was corrected with the whole

classroom, covering at least the following grammatical elements from the source text: *absolute ablative (*e.g. *instante partu* and *actis mensibus*), the paradigm of the 4th declension (e.g. *gemitus*) and deponent verbs (e.g. *hortatus est*), as well as the meaning of the fable.

Participants individually answered written questions on source-text comprehension of both fables in their lesson books as a conclusion for the lesson.

Instruction session 4: Error analysis and source-text comprehension in questions (III). The fourth lesson started by looking back on what the error analyses had brought. Each participant wrote a statement in their lesson book and then translated "Lupus et Agnus" (The Wolf and the Lamb) into "understandable and readable Dutch."

 The translation task was followed by a class discussion of the source text covering grammar and syntax. Participants were instructed to improve their target texts using pencil (*Figure 30*).

Figure 30: Example of a Target Text Corrected with Pencil.
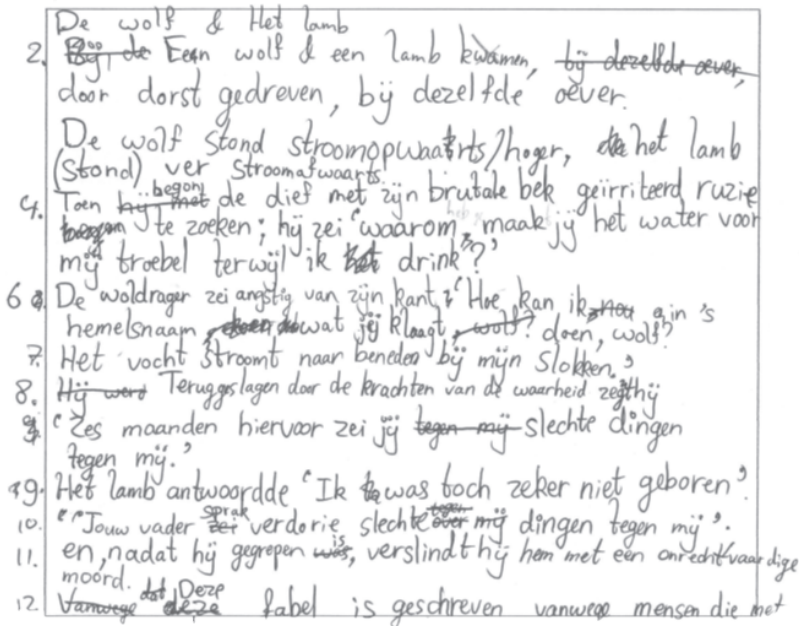
The discussion also covered the errors they had made, what errors were still frequent and what differences they had noticed, comparing the first and the

present lesson. After the classroom discussion of the fable, participants could either draw the fable in three pictures (*Figure 31* and *Figure 32*) or make an error analysis, as seen fit.

Figure 31: Example 1 of a Drawing of Fable "Lupus et Agnus".
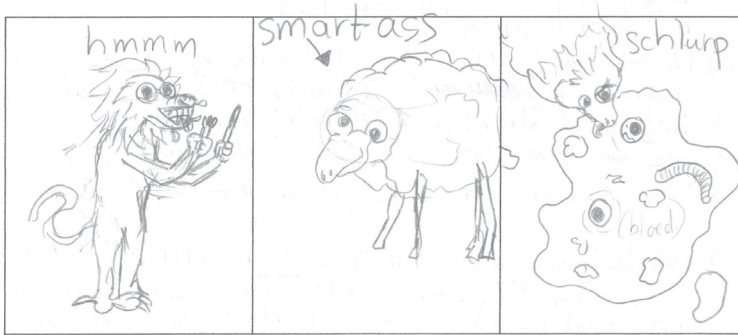
Figure 32: Example 2 of a Drawing of Fable "Lupus et Agnus".

Section nine presented an impression of the lessons and the students' products. The complete report on the experimental study and its effects is provided in chapter six.

CHAPTER 6
EFFECTS OF TWO INTERVENTIONS

Improving Target-Text Coherence and Source Text-Target Text Equivalence

1   INTRODUCTION TO THE EXPERIMENTAL STUDY

The object of this dissertation was to study whether the text coherence of target texts produced by students in upper secondary education translating Latin can be improved. The previous chapter described the trial run and redesign of an experimental study aimed at teaching a newly developed process-oriented translation strategy. The redesign resulted in an experiment consisting of two competing experimental conditions:

1. a *process condition* focusing on the development of text production and revision skills as well as the development of metalanguage, with instruction characterised as strategy instruction, and
2. a *product condition* focusing on error analysis and source text-target text equivalence, where the instruction method could still be characterised as *business as usual*.[156]

The present chapter reports on the experimental study and its results.

The internal validity of an experiment is crucial: otherwise no conclusions between cause and effect can be made. It implies that effects in post-test and delayed test can only be reliably attributed to the experiment if fidelity of implementation in both conditions is established, if differences in group composition prove not to be significant between conditions, and if scores in pre-test do not significantly differ.

---

[156] *The design of this experiment is argued in Chapter 4, its redesign in Chapter 5.*

Observing O'Donnell's principles for fidelity of implementation,[157] we[158] aimed at assessing that teachers taught what they were supposed to teach and that participants did what they had to do within conditions *(adherence* and *quality of delivery)*. We assumed that lessons with a clear purpose that are perceived as useful elicit high responsiveness. To measure *participants' responsiveness,* we administered a questionnaire at the end of session 3, session 5, post-test, and delayed test. We asked participants what they *thought* they had learned and what they *expected* of the use of what they had learned. We also asked if they thought *the lessons* to be *useful* (session 3) and *pleasant* (session 5). It is equally important to assess sufficient differences between conditions: *process* and *product condition* should differ on key elements *(programme differentiation)*. We asked participants if they had performed specific learned activities in the last lesson. We asked in both conditions whether participants performed orientation and revision activities. These elements are always implicitly part of the translation process, but we explicitly taught them as key elements in the translation process in the *process condition*.

We established a balanced group composition as to participants' self-reported grade, school type, and age in attributing participants to conditions (Section 2.2). Having established equality in group composition, fidelity of implementation (Section 2.6), and having established that scores in pre-test do not significantly differ between conditions (Section 2.7), the main questions we wanted to answer were:

1. Do participants in the *process condition* produce more coherent target texts in post-test and delayed test, compared to participants in the *product condition*?

2. Do participants in the *product condition* produce more equivalent target texts in post-test and delayed test, compared to participants in *process condition*?

Secondary, we wanted to explore the role of proficiency in Dutch and in Latin for both conditions. It can be argued that writing coherent texts for translations could be a result of writing skills in the target language (Dutch), and that high

---

[157] *'(a) adherence - whether the components of the experiment are being delivered as designed; (b) duration - the number, length, or frequency of sessions implemented; (c) quality of delivery - the manner in which the implementer delivers the program using the techniques, processes, or methods prescribed; (d) participant responsiveness - the extent to which participants are engaged by and involved in the activities and content of the program; and (e) program differentiation - whether critical features that distinguish the program from the comparison condition are present or absent during implementation' (O'Donnell, 2008:34)*

[158] *The research discussed in the present chapter was performed in close collaboration with my co-supervisor Suzanne Adema and my supervisor Gert Rijlaarsdam. For the statistical analysis of the data we gratefully relied on the advice provided by Huub Van den Bergh. Therefore, I will mostly use the plural 'we' in this chapter.*

equivalence scores in the post-test could be a result proficiency in Latin. When a condition relies strongly on the target language proficiency, one may expect an interaction between the learning condition and the proficiency in the target language, for instance. Therefore, we have also looked into the following:

3. To what extent do the effects of conditions depend on initial proficiency in Latin or Dutch (self-reported and scores at pre-test)?

We found that participants in both conditions significantly improved both their coherence and their equivalence scores in post-test. This outcome is nuanced for different types of participants in the results paragraph (Section 3.2).

The present chapter follows the usual structure of reports on experimental studies and consists of three sections: introduction, method and results. In the results section, it is common to present the outcomes of statistical calculations. The possible readability issues I mentioned earlier considering the diverse backgrounds of its intended readers (Chapter 1.1) are most acutely present in this chapter. Therefore, some modifications were made to accommodate the reader by moving the statistics to the footnotes as much as possible.

2   METHOD

2.1 Experimental Design

The experiment was performed, as mentioned, in two conditions that we labelled: *process* and *product*. The design consisted of three measurement occasions: a pre-test, a post-test and a delayed post-test (maintenance). All participants attended four extracurricular[159] Latin translation lessons of circa 120 minutes each. *Table 34* shows the design of the experiment.

Table 34: Experimental Design

	Pre-test	Experiment	Post-test	Maintenance post-test
	Translation task	4 sessions, 2 hrs each	Translation tasks	Translation tasks
Week	1	2-5	6	11
Process	O_1	$X_{process}$ and target text	O_2	O_3
Product	O_1	$X_{product}$ and source text	O_2	O_3

Note: O = observation (measurement); X = intervention. No contact between week 6-11: Instruction delay.

[159] *Thus addressing the problem with attendance experienced in the trial run, see Chapter 5.*

Lessons in the *process condition* focused on improving the translation process and the Dutch target text, offering no specific training in Latin, the source language. Lessons in the *product condition* focused on the translation product and the Latin source text, which is 'business as usual' in Dutch Latin classes; to concur with the *process condition*, the *product condition* was set up as a turbo translation training, focusing on error analysis.

2.2 Participants

Selection of Student Participants

We aimed at recruiting both weak and strong translators as participants for the experiment, with a variance in motivation for participation (intrinsic and extrinsic), from various school backgrounds, to assure that effects could not be attributed to participants' shared experiences with learning Latin translation. Student participants were recruited via their regular Latin class. Forty-five Latin teachers of eleven secondary schools in Amsterdam were approached by email with the request to have the researcher visit their regular Latin class and inform students about the experiment. The criterion for participation was that participants had finished the basic Latin curriculum.[160]

In the classes of the fourteen teachers who responded favourably, the researcher provided a short introduction to the experiment and interested students were invited to write down their names and email addresses. Within a week, these students received a Google-form to confirm their participation and to collect background information. This form was filled in by students of eight different schools. In response to the form, they received another email with more specific information about the project itself and the informed consent form for their parents or guardians.[161] Participation was voluntary, and participants were allowed to withdraw at any time. Participation was encouraged by offering a monetary reward upon completing the whole experiment.[162] Thus we recruited a diversely motivated group of participants, as is indicated by the reasons given to participate. These varied from "I need tutoring for Latin and might as well get paid for it" to "I want to help develop methods for teaching Latin translation." Furthermore, the mix of participants from different schools assured that effects could not be attributed to participants' shared experiences with learning Latin translation.

---

[160] *I.e. participants had to have finished a textbook based on the list of basic morphologic and syntactic knowledge of Latin, which is generally done in lower secondary education.*
[161] *All participants under 18 turned in signed informed consent forms.*
[162] *Thus addressing the problem with attendance experienced in the testing as described in Chapter 5.*

Distribution of Student Participants Over Conditions
Student participants were randomly assigned to the two conditions, taking into account self-reported Latin grade points, grade and school, to distribute those variables evenly across conditions as shown in Table 35. Participants in each condition were split into a Monday (A and C) and a Thursday (B and D) group, according to their availability. Participants were allowed to change groups within the condition if their school schedule changed during the weeks of the experiment. *Table 35* shows relevant demographic characteristics of the participants.

Table 35: Relevant Characteristics of Participants (Percentages)

	Process	Product	Total
Year of birth			
1998	2,4	0	1,2
1999	7,1	10,3	8,6
2000	38,1	51,2	44,4
2001	40,5	30,8	35,8
2002	11,9	7,7	9,9
Grade			
10th	7,2	5,1	7,4
11th	52,3	56,4	53,1
12th	40,5	38,5	39,5
Grade points Latin			
Fail	2,4	5,1	3,7
Insufficient	21,4	23,1	22,2
Sufficient	42,9	38,4	40,7
Satisfactory	23,8	23,1	23,5
Good	9,5	10,3	9,9
Grade points Dutch			
Fail	0	0	0
Insufficient	0	0	0
Sufficient	11,9	20,5	16
Satisfactory	54,8	48,7	51,9
Good	33,3	30,8	32,1
Schooltype			
Lyceum	50	59	54,3
Grammar school	50	41	45,7
Gender			
Female	66,6	71,8	67,9
Male	33,3	28,2	32,1

The experience with authentic Latin texts is indicated by grade: 10[163]th grade equals 0-2 months of experience, with each following grade adding twelve months of experience. Both year of birth and grade are taken as variables, as some participants have skipped a grade, while others failed a class. This means that, in itself, neither grade nor age equals experience.

Participants' grade points in Latin are a relevant characteristic as the distributions over conditions must not differ. The experiment is partly based on target-text writing, therefore grade points for Dutch are included as variables as well. Participants were asked to report their grade points for Latin and Dutch in one out of five categories on the intake form:
-    Fail (grade points 1 - 4 out of 10)
-    Insufficient (grade points 4 - 5.5)
-    Sufficient (grade points 5.5-6.5)
-    Satisfactory (grade points 6.5- 7.5)
-    Good (grade points 7.5 or higher).
School type is considered a variable as well, as Latin is taught as a compulsory subject in Dutch grammar schools,[164] whereas students at a Lyceum can choose to follow Latin classes. This difference might influence motivation or accomplishment in participants. The school type the participants attend is either Lyceum or Grammar School.

Selection of Instructors and Distribution over Conditions
Teaching experience and availability were leading in the selection of the teachers, as well as familiarity with the school the experiment was conducted at. Four teachers (A-D) were selected to teach the lessons. *Table 36* shows relevant characteristics.

Table 36: Participating Teachers: Characteristics

Teacher	Experience (years)	Age	Gender
A	20+	50-55	F
B	10+	35-40	F
C	20+	50-55	F
D	10+	35-40	M

Two teachers were already involved in the research project, one as main researcher and the other as research supervisor. Two teachers were teachers of

---

[163] *Dutch: 4thgrade in secondary education.*
[164] *Dutch: categoraal gymnasium.*

Latin at the school that hosted the experiment. To ensure that teachers of the process-oriented lessons had a deep understanding of the lessons' content, the teachers involved in the research project taught the *process condition*, while the other two experienced teachers taught the *product condition.* Teachers taught one group each, either *Process* (A and B) or *Product* (C and D).

2.3 Measures

We implemented several instruments to collect data to assess fidelity of implementation. Additionally, we measured the dependent variables 1) quality of the translations (target-text coherence and ST-TT equivalence) and 2) the translation process. *Table 37* shows the instruments that were used to collect data.

Table 37: Overview of Instruments

	Session				
	1	3	5	6	7
Instrument	Pre-test	Instruction	Instruction	Post-test	Maintenance
Implementation Fidelity					
Questionnaire		x	x	x	x
Lesson books		x	x		
Translation Quality					
Target text	x			x	x
Translation process					
Questionnaire			x	x	x
Screencast	x			x	x

2.3.1   Dependent Variable: Implementation of Fidelity

The questionnaire covered five variables for fidelity, listed in *Table 38*. The first three variables indicate how participants felt about the lessons. I expected to find no significant differences between conditions, as I wanted to avoid that effects of the conditions could be attributed to differences in perception between conditions. Variable four measured self-reported activities during the translation process and variable five measured use of metalanguage through open questions. Systematic revision and the development of metalanguage was part of the instruction sessions in the *process condition*. In the *product condition* it was not. Therefore, I expected to find differences between conditions in the results on variables four and five.

Table 38: Fidelity of Implementation Measured

Variable	Instruction session	Test session
1. Expectations of the usefulness of the training	3, 5	6, 7
2. Attitude about the lessons	3	
3. Perception of the lessons	3, 5	
4. Translation process	5	6, 7
5. Use of metalanguage	3, 5	

Lesson books: Lesson books were used in all four lessons. At the start of each class, participants were each given a lesson book containing explicatory texts and assignments for that lesson only. They had to write down their responses in the book during the class. They were not asked to work on the assignments at home.

2.3.2   Dependent variable: Translation quality

The *process condition* aimed at improving target-text coherence, while the *product condition* aimed at improving equivalence of target text and source text. Therefore, we assessed the quality of the target texts participants produced in the three measurements on both aspects:
1. coherence of target text on text level,
2. equivalence of source text and target text on sentence level.

Text coherence. For holistic assessment of the text coherence in the target texts participants produced, we used the online tool for comparative judgment D-PAC.[165] Through the programme we instructed assessors as follows:

> You are offered two texts at a time for comparison: the texts are translations of fables. Sometimes the two texts are translations of the same fable, sometimes the fables differ. You are asked to judge which text seems to be written in more coherent Dutch. Is the text written fluently, do you understand the story that is told and the moral? Are relations between sentences logical? Your judgement is based on your gut feeling, or rather: on your expertise as a teacher.

---

[165] When the target texts are uploaded in D-PAC, the tool offers a randomly composed set of two texts next to each other on the computer screen and the assessor chooses which text is more coherent. See also Chapter 2.

Each target text received a coherence score through comparative assessment in random pairs, by means of D-PAC. The 480 target texts from the three measurement occasions were uploaded in D-PAC. For the assessment I approached 83 people in my network, of which 39 responded favourably: teachers of Classics, Dutch, Philosophy, History, Economics and English as well as teacher trainers and even a professional translator from English to Dutch. Due to workload and unknown other reasons, only 28 of the initial 39 assessors were able to perform assessments. The assessors made 4838 comparisons, on average 172.8 comparisons each. Reliability of scores was high ($r = .85$). This way each text received a coherence score. A screenshot of the ranking of the 480 texts is shown in *Figure 33*. As the tests consisted of two fables, two scores for each test occasion were available. In data-analysis the average of the two scores was taken as the score for each test occasion.

Figure 33: 480 Target Texts Ranked According to Coherence in D-PAC.
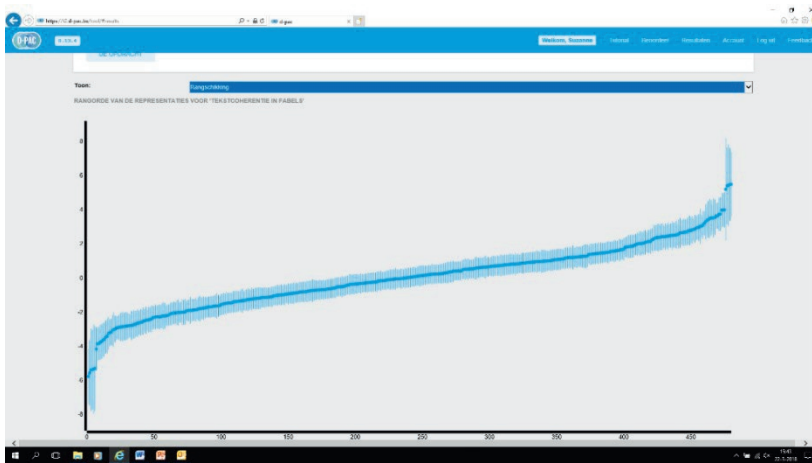
Text equivalence. We used colon-rating to assess ST-TT equivalence. As this rating method is considered to be highly reliable[166] one experienced rater is considered to be sufficient. Given that the workload of rating 480 target texts is considerable, we decided to split the target texts over two experienced raters, provided inter-rater reliability was accurate. Two skilled assessors scored a sample of thirty target texts with the colon score model. We selected target texts written by the first thirty participants in alphabetical order, equally spread over the six fables, which means that we selected the texts of the first

---

166 *See Chapter 2.2.*

five participants for the first fable, the texts of the second five participants for the second fable, and so on until the thirtieth participant. Inter-rater reliability was high (r = .89), hence target texts were assigned to one assessor: for each test assessor 1 scored the first fable (i.e. fable 1, 3, and 5) and assessor 2 scored the second fable (i.e. fable 2, 4, and 6) so that both assessors contributed equally to each measurement occasion. In data-analysis the average of the two scores was taken as the score for each test occasion, with maximum average scores ranging from 13,5-14 (*Table 39*).

Table 39: Maximum Colon-Scores per Test Occasion

	Pre-test	Post-test	Maintenance-test
Max score points 1st fable	14	16	15
Max score points 2nd fable	13	12	11
Max score points average	13,5	14	13

2.3.3   Dependent variable: Translation process

We collected data about the translation process by a questionnaire and a screencast of the target-text production, as a back-up for the self-reported translation process in the questionnaire.

Questionnaire. We collected translation process data on three occasions (after session 5, post-test and maintenance test) with questions about 1) orientation, 2) text production, and 3) revision. After session 5, a question in the *process condition* questionnaire asked about the use of the reminder: "I used the 'reminder' provided (no / yes). Other than that, the same questions were asked on all occasions:

1.   Orientation[167]
-    When I was translating today I thought about the meaning of the story BEFORE[168] I started writing (no / one fable no, one fable yes[169] / yes)
-    When I was translating today I marked the three building blocks of the fable in the text (no / one fable no, one fable yes / yes)

---

[167] The questionnaire did not explicitly mention "translation phases."
[168] Capitalised in the questionnaire.
[169] Answers in Italics were added in posttest and delayed test, as participants translated two fables.

2. When I was translating today I understood the story's moral (no / yes, after class discussion[170] / yes, while I was writing down my translation / yes, before I started writing down my translation).
3. Text production
   Process condition: When I was translating today I wrote a first and second draft of my target text (no / *one fable no, one fable yes* / yes).
   Product condition: When I was translating today I paid attention to mistakes I am known to make (no / *one fable no, one fable yes* /yes).
4. Revision
   Process condition: AFTER translating today, I revised my text with the +/- method[171] (no / yes / *one fable revised without +/-method, one not revised / one fable revised with +/-method, one not revised / both* revised without +/-method / yes).
   Product condition: AFTER translating today, I checked my translation for mistakes I am known to make (no / *one fable no, one fable yes* / yes, without attention to known errors / yes).

Screencasts. To observe the translation process, the activity on the computer screen was recorded by the online tool *Screencast-o-matic* on all three test occasions. We intended to record all target-text production activities through screencast as additional information to the translation process participants reported in the questionnaire. If needed, a sample of screencasts could be coded on text-revision activities e.g. the final five minutes of the translation process.

2.4 Procedure

Instructors. All teachers were provided well in advance with detailed lesson plans for each session. Teachers in the *product condition* were instructed to teach as much as possible in a 'business as usual' way, even if the lessons in the *product condition* were designed to be more focused on translating than usual. To avoid bleeding, teachers in the *product condition* were not made familiar with the content of the lessons in the *process condition*.

Teachers in both *process* and *product condition* reported in an informal setting that the lessons were conducted as planned. Some lessons in group D were finished by teacher A, as teacher D had to leave at a fixed time and some lessons took a little longer. This did not lead to any problems. Teacher C spontaneously reported that she found the teacher instructions and lesson plans very clear. Teacher A and C did not report problems with the execution of the lessons, despite the technical and scheduling difficulties described below.

---

[170] *The answer 'yes, after class discussion' was omitted in posttest and delayed test.*
[171] *For the +/- method, see Chapter 4.3.3 and Chapter 5.9.1.*

Instruction sessions. Participants were invited to come to the hosting school on Monday or Thursday afternoon, in accordance with their availability. Lessons in groups A and C were scheduled simultaneously on Mondays, lessons in B and D on Thursdays.[172] Mondays and Thursdays an assistant was present to provide drinks and snacks and to manage the administration of lesson books and test files. For missed sessions a catch-up session was provided, to make sure participants were able to attend all sessions.

The Monday groups (A and C) experienced the following logistic difficulties:

1. The 3rd session was postponed one week due to a change in the teacher's schedule.
2. The 5th session had to be cancelled due to a snowstorm.

Catch-up sessions were organised on Wednesday and Friday in the same week.

Three participants changed from group A to group B due to their school schedule. One participant changed from group B to group A in the course of the experiment. Forty-two participants attended all seven sessions in the *process condition* (see Appendix A for details). Two participants changed from group C to group D, two participants changed from group D to group C and one participant dropped out. Thirty-nine participants attended all sessions in the *product condition* (see Appendix A for details).

At the start of each lesson, participants were given a lesson book containing explicatory texts and assignments for that session only. The lesson books' wrappers were colour-coded[173] to avoid mistakes in distributing the books. Participants had to write their name on the cover each time. The teachers collected them at the end of each lesson and the assistant made sure the set was complete and then stored the lesson books.

Tests. On all three test occasions, participants translated two fables using a desktop computer in one of two computer rooms of the hosting school. They were provided with a login to a shared guest account to access the computer and a Latin-Dutch dictionary in hardcopy (Pinkster, 2008).[174] The activity on the computer screen was recorded by the online tool *Screencast-o-matic*. In the trial run of the experiment, the screencast programme could not be used

---

[172] *As observation in all classes presented a problem in a setting where two classes are scheduled simultaneously, lessons were audio recorded. The recording could be used as back-up, to check the implementation if inexplicable differences between groups or conditions would arise. This proved not to be necessary.*
[173] *A = red, B = blue, C = green, D = yellow.*
[174] *We used both the 5th and 6th edition of the dictionary.*

due to technical problems.[175] Again, we experienced some technical difficulties during the pre-test as the screencast programme was not functioning on several computers for the Monday groups (A and C). Furthermore, it appeared that the screencast files could not be saved directly to the guest account. It took some time to install all participants at functioning computers and to adapt the instructions for saving the files. These difficulties were solved the following Thursday for groups B and D. Participants saved the target texts as PDF files and the screencasts on the shared guest account. After each test the assistant checked whether all participants had completed the test and uploaded the files. All PDF files were uploaded at the same time in D-PAC for comparative judgement after the delayed test.

The period of time between post-test and delayed test varied due to scheduling problems. The Monday groups (A and C) were initially scheduled to perform the post-test on Monday 11/12 and the delayed test on Monday 15/01. As we had to postpone session 3 one week for these groups,[176] the whole schedule moved a week and the post-test was performed one week later than planned. The delayed test, however, was scheduled well in advance for the week of January 15th and could not be moved. Therefore, the period of time between post-test and delayed test for groups A and C was one week shorter (four weeks), than for groups B and D (five weeks). To avoid that activities were prompted by the presence of their familiar instructor the post-tests were not performed by the groups' teachers. Due to absence of teacher C teachers supervised the tests following the schedule shown in *Table 40*.

Table 40: Distribution of Teachers over Groups and Test Sessions

	Post-test Thu 14/12	Post-test Mon 18/12	Delayed test Mon 15/01	Delayed test Thu 18/1
Group/Teacher	B / D	A / B	A / B	B / D
	D / A	C / A	C / A	C / A

2.5 Data collection

Data collection on fidelity.
- *Questionnaire:* A questionnaire was set up in the online tool *Socrative* to perform the evaluations of sessions 3, 5, 6 and 7. Participants used their telephone (instruction sessions 3 and 5) or the desktop computer (test

---

[175] See Chapter 5.
[176] Due to a change in the teacher's schedule.

sessions 6 and 7) to complete the questionnaire. In sessions 3 and 5, the questionnaire was provided in print for those who had no telephone or could not connect it to the internet.

- *Lesson book.* Data were collected during the lessons where participants had to fill in their lesson book. The assistant scored whether the tasks and exercises were completed. The quality of the work was not analysed. We only included data in the analyses from the participants who completed the training and attended all seven sessions. The assistant reported that no lesson books were missing in both conditions and that all assignments were completed.

Data Collection on quality of translations and translation process.
- *Target texts.* In the pre-tests, two text files of fable 1 were not saved as instructed and could not be used in the data analysis (fable 1: N=79). As a result of the technical difficulties, not all participants in groups A and C were able to translate both texts of the pre-test in the given time and some of the pre-test files are missing (fable 2: N = 74). In post-test and delayed test all participants saved their text files correctly (N = 81).
- *Screencast.* As not all participants in groups A and C were able to translate both texts of the pre-test in the given time some of the pre-test screencasts are missing (fable 2: N=74). In post-test and delayed post-test, all files of participants that completed the experiment were successfully saved.

2.6 *Fidelity of implementation.*[177]

Observing O'Donnell's principles for fidelity of implementation (O'Donnell, 2008), we assumed that lessons with a clear purpose that are perceived as useful elicit high responsiveness in participants. To measure participants' responsiveness, we analysed items from the questionnaire that we administered at the end of session 3, session 5, post-test, and delayed test. We asked participants what they thought they had learned and how they rated the expected usefulness of what they had learned. We also asked whether they thought the lessons itself to be useful (session 3) and fun (session 5). Additionally, we asked participants to self-report on the performance of the activities they were taught in the lessons, to indicate fidelity.

Participant responsiveness. At the end of the third session, we asked participants in both conditions what they expected of the usefulness of what they had

---

[177] *To increase readability for readers with non-statistical backgrounds, I have moved most statistical references to footnotes.*

learned so far for improving their translations, specifying *genre knowledge* in both conditions and *revision* in the *process condition* and *error analysis* in the *product condition*. *Table 41* shows scores that indicate the expectations of 1) the usefulness of genre knowledge and 2) of the usefulness of what they had learned in session 3. In expectations of the usefulness of *genre knowledge* the difference between conditions is statistically significant.[178] Participants in both conditions on average had positive expectations of the usefulness of what they learned ($M > 1,5$ on a 3-point scale), though no significant difference[179] between the two conditions was observed in expectations of the usefulness of what they learned.

Table 41: Participants' Perceived Usefulness (Session 3)[180]

	Process		Product	
	M	*SD*	*M*	*SD*
Q1: How useful for translating do you expect your genre knowledge (e.g. three building blocks) to be?	1.88	.81	1.18	.68
Q2: How useful for translating do you expect *what you learned in session 3[181]* to be?	1.73	.63	1.51	.64

The lessons in the *process condition* presented genre knowledge embedded in the orientation phase of the translation process, combining it with an orientation assignment. Through this type of presentation, the usefulness of this particular type of knowledge was not only stated, but also argued and experienced through the assignment. The lessons in the *product condition* presented genre characteristics in a more isolated way, merely stating the fable's characteristics and importance and offering an assignment to apply the genre knowledge as such. The relation between genre knowledge and text comprehension was not discussed.

Table 42 presents the *development in expectations* in participants of the usefulness of their lessons' content measured on three occasions: at the end of the last instruction session (5), post-test (session 6) and delayed test (session 7).

---

[178] t*(78) = 4.15,* p *<.001.*
[179] t*(78) = 1.5,* p *= .22.*
[180] *Scale 0-3.*
[181] *What you learned in session 3: Process= revision and Product= error analysis.*

Table 42: Participants' Expectations of Usefulness[182]

Session	5		6		7	
	M	*SD*	*M*	*SD*	*M*	*SD*
Condition Product knowledge of personal translation errors	1.68	0.78	1.54	0.64	1.51	0.77
Condition Process knowledge of the translation process	1.64	0.62	1.57	0.55	1.46	0.55

The questions we asked about expectations differed between conditions: we asked what they expected of the usefulness of their knowledge of the translation process (*process condition*) or what they expected of the usefulness of their knowledge of their personal translation errors (*product condition*) for improving their translations in general, i.e. also outside the experiment. No participant reported at any time they expected the knowledge to be "not useful." On all three measurement occasions the expectations in both conditions were "somewhat useful" – "useful" (*M* > 1.5 on a 0-3 scale). The expected usefulness did not decrease during the experiment[183] and correlations were stable.

We also asked participants, after instruction sessions 3 and 5, in open questions whether they had missed anything in the lessons. Twenty-eight participants (almost 70%) in the *process condition* and nineteen participants (almost 50%) in the *product condition* reported they missed nothing in the lessons at the end of session three (*Table 43*). After completing all instruction sessions at the end of session 5, the number of participants that reported they missed nothing decreased in the *process condition* with four participants (11%), while in the *product condition* it increased with four (17%).

Participants in the *process condition,* after completing all instruction sessions, mainly reported having missed feedback on their translations (5), more challenging source texts (5), and grammar (4). Some had missed tips and tricks (2), Latin (1) or theory on translation (1). In the *product condition,* they missed mainly grammar (4) if they missed anything at all. Other aspects participants in the *product condition* reported to have missed were: theory on translation (2), tips and tricks (2) and more challenging source texts (1). It was to be expected that participants missed grammar in the *process condition*. However,

---

[182] *Scale 0-3.*
[183] *Process condition: session 5-6 (t (41) = .90, p =.37) and session 6-7 (t (40) = 1.95, p = .06). Product condition: session 5-6 (t (36) = 1.43, p =.16) and session 6-7 (t (36) = .27, p = .79).*

for the *product condition* it seems odd, as the content of the lessons in this condition specifically focused on Latin and grammar.

Table 43: Participants That Reported Having Missed Nothing

	Process				Product		
Session 3		Session 5		Session 3		Session 5	
number	%	number	%	number	%	number	%
28	68	24	57	19	49	23	66

Attitude and perception. We measured participants' attitude and perception of the lessons after sessions 3 and 5, using questions about clarity, experienced difficulty (perception) and experienced pleasantness (attitude). *Table 44* shows that participants perceived lessons in both conditions as clear[184] and the level of the assignment was not perceived as too easy or too difficult.[185] The lessons in both conditions were considered to be fairly pleasant.[186] Differences between conditions were not statistically significant,[187] so possible differences in effect between conditions cannot be attributed to differences in attitude and perception.

Both conditions featured group assignments, working in pairs and individual assignments. I wanted to know if strong preferences for one of these work forms existed, as this might influence participants' attitude; e.g. a strong preference for working in pairs could affect the fidelity of implementation of the experiment. Therefore, participants were asked about their preferred way of working after session three (*Table 45*). Differences between conditions were not statistically significant.[188]

---

[184] *M > 2 on a 3-point scale in both conditions.*
[185] *M > 1,9 on a 4-point scale, where 2 is 'exactly right.'*
[186] *M > 1,5 on a 3-point scale in both conditions.*
[187] *Clarity (t (78) = 1.69, p = .09); difficulty (t (78) = .37, p =.71); pleasantness (t (73) = -1.96, p = .05). The difference between the two conditions for clarity and pleasantness was marginally statistically significant.*
[188] t = (78) = -.76, p =.45.

Table 44: Participants' Perception of the Lessons

Ses-sion	Question	Process		Product	
		M	*SD*	*M*	*SD*
3	Q: Did you feel the lessons were clear? (not = 0 – very = 3)	2.29	0.56	2.08	0.58
3	Q: What do you think of the difficulty level of the assignment[189] (0 = far too easy – 4 = far too difficult)	1.98[190]	0.61	1.92	0.66
5	Q: Did you feel the lessons were pleasant? (not = 0 – very = 3)	1.54	0.55	1.81	0.62

Table 45: Preferred Work Form per Condition

	Process		Product	
Work format	Frequency	%	Frequency	%
Individual	16	39	10	25.6
Pair	14	34.1	21	53.8
Group composed by students	10	24.4	4	10.3
Group composed by teacher	1	2.4	4	10.3
Total	41	100	39	100

2.6.1 Programme differentiation

Another aspect of fidelity of implementation is whether participants performed the activities they were taught during the experiment. We used questions about the self-reported translation process to assess differences in the translation process between conditions, as a change in the translation process was part of the experiment in the *process condition*. We asked questions about the translation activities in three phases of the translation process: 1) orientation 2) text production and 3) revision at the end of the instruction sessions

---

[189] *Assignment: process = comparing eye-tracking film of revisor 1 and revisor 2; product = translation task.*
[190] *NB: 4-point scale.*

(session 5), post-test (session 6) and delayed test (session 7). We will discuss the results for each translation phase separately.

Phase 1: Orientation. The orientation phase was explicitly discussed in the *process condition* as part of the translation process. In the *product condition,* however, it was not an explicit subject. The three building blocks of the fable were discussed in both conditions, in *Process* participants experienced its relevance for the orientation phase through several assignments, in *Product* the three building blocks were introduced by the teacher as fable characteristics and the building blocks were marked in a fable (see Chapter 5.8).

I asked all participants the same three questions on orientation,[191] expecting different outcomes for each condition. The third question was asked once at the end of session 5, and twice, once for each translated fable, in test session 6 and 7. *Table 46* shows the percentage of participants that reported they performed the specified orientation activity before writing.

Table 46: Orientation Activities Before Writing (Percentages)

Performed activity		Process			Product		
		5	6	7	5	6	7
1. Think about the meaning of the story		72	75	77	34	49	49
2. Mark the building blocks		30	40	40	3	17	20
3. Understand the moral	Fable 1	24	0	0	8	0	0
	Fable 2	x[192]	0	0	x	0	0

The percentage of participants who reported they thought about the meaning of the story before writing (Q1) is higher on all three measurement occasions in *Process* than in *Product*. This difference was expected, as thinking about the story as an orientation activity was part of the experiment in *Process*. The difference between conditions is statistically significant.[193] It seems striking that in *Product* the percentage is the same in session 6 and 7 (49%). However, with a low correlation,[194] this behaviour is not as stable as the percentage

---

[191] *Q1: Before I started writing today, I thought about the meaning of the story BEFORE I started writing (yes / no).*
Q2: When I was translating today, I marked the three building blocks of the fable in the text (yes / no).
Q3: When I was translating today, I understood the story's moral before writing (yes / no).
[192] *Only one fable was offered for translation in session 5.*
[193] $\chi^2(1) = 9.92$ *(p =.002).*
[194] *r = .35 (p = .03).*

suggests, meaning that participants who thought about the meaning of the story before translating in the pre-test did not necessarily maintain that behaviour in the post-tests and participants who did not think about it in pre-test reported to have started thinking about it in post-test or delayed test. In the section below, regarding target-text production, I will discuss differences between conditions in understanding the moral in the text-production phase.

The percentage of participants who reported that they marked the three building blocks of the fable (Q2) in *Process* is higher than in *Product* on all three measurement occasions. In *Product,* we notice an increase in building-blocks marking from session 5 (3%) to 6 (17%) and 7 (20%). Possibly the question itself has triggered marking of the building blocks in next sessions. The consistency in *Process* of marking the building blocks in session six and seven is striking (40%). In this case the correlation is very high,[195] which indicates that this behaviour is highly stable: participants who marked the three building blocks in the post-test maintained that behaviour in the delayed test four to five weeks later, without further prompting.

Interestingly, none of the participants reported that they had understood the moral of the story *before* they started writing (i.e. as an orientation activity) in sessions 6 or 7. This suggests that thinking about the *moral*, resulting in understanding the moral *before* writing, stopped within a week after the lessons were finished or was not maintained during test sessions at least.

Phase 2: Text production. The text production phase was explicitly discussed in the *process condition* as part of the translation process. Participants were made aware that *translating* is essentially *writing* and that writing two drafts of a target text is part of the translation process. In the *product condition,* text production was not explicitly discussed. Here, participants learned to signal the translation errors they made through their error analyses. No explicit attention was given to writing a target text. To assess implementation, we asked a condition-specific question.[196] *Table 47* shows the percentage of participants per condition answering that they performed the instructed text-production activities through session 5, 6 and 7. We observe that directly after completing all instruction sessions (session 5) 83% -84% participants in both conditions report they do as they were instructed in the experiment. However, in the *process condition,* a sharp decrease occurs, as 49% of the participants in post-test (session 6) report that they are writing two drafts of the target text. In the

---

[195] r = .90 (p <.001).
[196] Process condition Q: When I was translating today I wrote a first and second draft of my target text (yes/no).
Product condition Q: When I was translating today I paid attention to errors I am known to make (yes / no).

delayed test (session 7) only 37% reported they were writing two drafts. In the *product condition* participants also report that they are paying less attention to their known errors while writing the target text, but this decrease is more gradual, to 76% in post-test and 51% in delayed test.

Table 47: Performed Text-Production Activities (Percentages)

Condition	Performed activity	Session		
		5	6	7
Process	Write a first and second draft of my target text	83	49	37
Product	Pay attention to translation errors	84	76	51

Additionally, I asked in both conditions whether participants understood the moral of the story *while* they were translating (*Table 48*), as understanding the moral may influence coherent text production. Thinking about the meaning of the story and trying to understand what you are translating was an explicit part of the experiment in *Process*, therefore I expected differences between conditions. The answers of participants were coded 0 (I did not understand the moral) 1 (I understood moral *while* writing) and 2 (I understood moral *before* writing). The difference in scores of participants who report they understood the moral, either *before* (orientation phase*, Table 13*) or *while* writing (*Table 15*) is significant between conditions at the end of session 5.[197] Differences between conditions are no longer significant[198] in sessions 6 and 7 for both fables. This suggests that understanding the moral or attempting to understand it before writing in *Process* is behaviour that wears off, as is writing two drafts of target text.

Table 48: Understanding the Story's Moral <u>while</u> Writing (Percentages)

	Process			Product		
Fable	5	6	7	5	6	7
1	76	92	76	89	89	79
2	X[1]	71	87	X	75	85

Table note[1]: in session 5, only one fable was offered for translation.

Phase 3: Revision. The revision phase was explicitly discussed in the *process condition* as part of the translation process. Participants practiced revising texts written by others before writing and revising their own target texts. The +/- method[199] was offered and practiced as a tool for revising, although it was not presented as the one and only way to revise. In the *product condition* revision was not explicitly taught. Here participants focused on their personal error repertoire and were encouraged to check whether they had avoided the errors they were known to make in each following translation task. In both conditions, we asked about activities performed *after* writing the target text. We expected conditions not to differ in reporting that they did what they were taught, i.e. use the provided method (fidelity). At the same time we expected more revision activities in *Process* (with or without method) as the importance of revision in the translation process was a prominent part of the experiment in *process condition*. *Table 49* presents the percentage of participants that reported whether they performed taught activities in both conditions in three categories.[200]

Table 49: Performances of the Taught Revision Activity (Percentages)

	Process			Product		
	5	6	7	5	6	7
No	2	14	24	13.5	5	13.5
Yes, without method	15	81	66	40.5	59	59.5
Yes, with method	83	5	10	46	36	27

By adding the percentages of reported revision (with and without method) we can observe performed activities as a binary (revision yes or no). Table 50 shows the percentage of participants who reported they revised or checked target text per condition. At the end of session 5, these activities are reported more frequently in *Process* than in *Product*, but at the end of session 6 and 7 we observe a decrease in reported activities in *Process*, while *Product* seems to remain more stable.

When we look at differences within conditions from one session to the next, the only significant difference in reported revision activities is in

---

[199] *For the +/- method, see Chapter 4.3.3 and Chapter 5.9.1.*
[200] *Recoding into three categories is described in Section 2.7.*

Process, where the decrease between session 5 and 6 is significant.[201] Other differences between sessions within conditions are not significant.[202]

Table 50: Revision or Checking Target Text (Percentages)

Session	5	6	7
Product	86.5	95	69.5
Process	98	86	76

2.6.2   Conclusions on Fidelity of Implementation

I wanted to establish accurate fidelity of implementation to be able to attribute effects to the experiment reliably. Therefore, I assessed that differences between conditions in expectations of usefulness, perception of clarity and pleasantness were not statistically significant. I wanted to assess that the differences in lessons between conditions were implemented. Here I was expecting statistically significant differences between conditions. The lessons in the *process condition* focused on developing new translation behaviour, while lessons in the *product condition* focused on error analysis, as a more common practice in the Latin classroom.

   Participants in both conditions considered the lessons to be useful, clear, and sufficiently entertaining at the end of sessions 3 and 5 (*Table 41:Q2 , Table 42, Table 43, Table 44*). The only significant difference between conditions is observed in the expectations of the usefulness of genre knowledge (*Table 41:Q1*). These were significantly higher in the *process condition*. The emphasis on the use of genre knowledge was part of the design of the lessons in the *process condition* and therefore a difference between conditions was expected here. Differences between conditions in the preferred work form were not significant (*Table 45*). Therefore, expectations, attitude, and perception in both conditions can be described as similar and positive, as a result of which it is plausible that differences in outcome between conditions are not related to these factors.

   Lessons in the *product condition* focused on error analysis and offered tools for error analysis in a structured manner. The lessons in the *process condition* focused on developing new behaviour in orientation, text production, and revision. We measured implementation through questions about the translation process. The differences between conditions in the orientation phase

---

[201] t *(40) = 8.709;* p *< 001.*
[202] *Process 6-7:* t *(40) =.530;* p *=.599 / Product 5-6:* t *(36) = .206;* p *= .838 / C6-7* t *(36) = 1.527;* p *= .136.*

(before writing) are, as anticipated, significant. In the *process condition,* participants report that they applied the orientation activities that were part of the experiment lessons.

In text production and revision phases, differences between conditions were not significant at the end of the instruction sessions (session 5). However, in post-test and delayed test, participants in *Product* remained more faithful to what they had learned than in *Process*, and differences between conditions were significant: in *Product,* most of participants' learned behaviour was significantly more stable than in *Process*. Only "marking the building blocks" and "thinking about the meaning of the text before writing the target text" were stable activities over session 5, 6 and 7 in both conditions. The decline in the practice of writing two drafts and revising in post-test and delayed test, however, is noticeable in *Process*. Possible explanations for the differences in stability of the learned behaviour between conditions are:

1. Participants in *Process* have not internalised what can be gained by writing two drafts and therefore writing the second draft could be experienced as superfluous.
2. Writing a target text on a computer does somehow influence the writing process: in *Process* the text of the first draft can easily be revised without making the text look 'messy,' which could make writing a whole new draft seem unnecessary.
3. In *Product,* the thought that paying attention to known pitfalls can improve your translation is in line with what students are taught in school. Therefore, the behaviour in *Product* is not entirely new, and could be more readily accepted as a means to improve your translation than writing two drafts, the behaviour taught in *Process*.
4. The questions on revision/checking the translation were phrased differently, as a result of which it is possible that "revising" in *Process* has become a concept that participants no longer related to "checking your translation errors."

Having established that fidelity of implementation in the experiment was accurate, we must look at scores in pre-test before we can look into effect measurements, as it must be established that differences in pre-test are not significant.

2.7 Analyses

We collected data on students' translation products and translation processes on three measurement occasions: pre-test, post-test and delayed test. The translation products were scored on two qualities of the texts: Equivalence and Coherence. Two issues are now of interest. First, we need to establish that

both scores have indeed some 'Latinity.' Second, we want to check whether the two conditions did not differ in the quality of Latin proficiency and other relevant variables at the pre-test. In the pre-test coherence and equivalence score correlated $r = .65, p < .001$. Coherence and equivalence score correlated significantly with self-reported Latin proficiency $r = .45, p < .001$ and $r = .39, p < .001$ respectively. So all three scores share something like 'Latinity'. The differences between conditions of the pre-test scores for Equivalence, Coherence, Latin and Dutch proficiency were not statistically significant *(Pillai's trace .429, df (4,238), p = .788)*. Therefore, in reporting effects of conditions we can safely attribute these effects to the experiment, as differences in the pre-test both in scores and in group composition are not significant, while programme differentiation is established.

Translation Quality. First, we tested the effect of the two learning conditions for the three measurement occasions, for two qualities of the translations students wrote: coherence and equivalence. These analyses reveal whether one of the two conditions had a different effect on one of the two outcome variables: Coherence and Equivalence, regardless of specific features of the learners. The data structure forms a hierarchical structure: participants' scores are nested in measurement occasions as visually represented in *Figure 34.*

Figure 34: Visual Representation of the Nested Data.
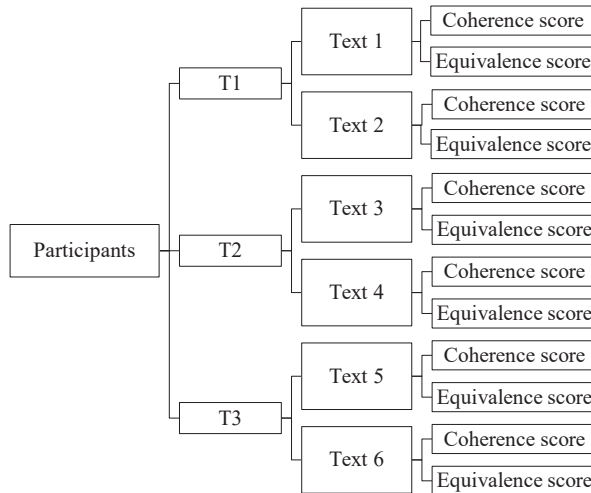
First, we compared a series of nested models, and we will present the estimates for the best fitting model. We started with a model without any explanatory factors: Model 0 contained the intercept and random components for

differences within and between students. Consecutively we added factors: Time (model 1), Condition (model 2), and the interaction between Time and Condition (model 3). These analyses provide an answer to the research question which of the two learning conditions was best for which of the two dependent variables: Coherence and Equivalence of the target text, the translation. To prepare these analyses we winsorised the scores in the dependent variables after boxplot analyses (10 out of 240 scores for Coherence, and one score out of 240 for Equivalence). We converted the extreme scores in the closest highest or lowest score respectively.

Secondly, we explored whether one of the conditions had an effect for a specific group of learners. These analyses provide answers on specific questions, for instance: do students with a relatively high level of Dutch proficiency profit more form the *process condition* regarding their coherence scores on the post- and delayed test than students with a relatively low level of Dutch proficiency? We ran analyses for each of the four learner variables available: the self-reported proficiency levels of Latin and Dutch and Coherence and Equivalence scores from the pre-test translations. For the two proficiency levels, we could analyse the effect for condition and three measurement occasions. For Coherence and Equivalence we had to analyse the effect of conditions on T2 and T3, while the coherence and equivalence scores were taken from the first (pre-test) measurement occasion. As explanatory variables we included 1) Time, 2) Condition, 3) Time*Condition, 4) Learner variable, 5) Time*Learner variable, 6) Condition*Learner variable, and 7) Learner variable*Time*Condition. When an interaction component with the learner variable in these models (model 6 and 7) proves to be significant, it indicates that the effect of condition is different for levels in the learner variable.

Translation processes. We analysed the development of the translation processes by scoring answers on orientation and revision activities and looking for significant differences between conditions as described on page 146. The systematic coding and analysis of the answers to the open questions in the questionnaire measuring the development of metalanguage proved to be too time-consuming for the scope of this dissertation.

3   RESULTS

3.1 Effects of Conditions on Coherence and Equivalence Scores

We started by looking whether our hypotheses were supported by the results. We expected that coherence scores in the *process condition* would improve significantly compared to the *product condition* and that equivalence scores in the *product condition* would improve significantly compared to the *process*

condition in the post-test (T2). *Table 51* provides the observed scores for the two qualities of translations measured at three measurement occasions for both learning conditions under study.

Table 51: Observed Scores for Two Conditions

Variable	Condition	T1		T2		T3	
		M	*SD*	*M*	*SD*	*M*	*SD*
Equivalence	Product	6.17	1.67	9.06	2.25	8.33	1.70
	Process	6.17	2.11	9.54	1.80	7.43	1.65
Coherence	Product	-0.96	1.40	0.25	1.39	0.42	1.46
	Process	-1.00	1.29	0.64	1.23	0.38	1.39

Note: Equivalence scores: theoretical range: 0-13.5 (T1); 0-14 (T2); 0-13 (T3); Coherence scores are logit scores, with a mean of ca. 0.

It shows that in pre-test (T1), differences between conditions on both variables were not significant.[203] Observed effects of condition can therefore be interpreted as pure effects of learning conditions.

Having established this, we need to provide a comparison of post-test (T2) and delayed test (T3) by comparing a series of nested models. We present the estimates for the best fitting model. *Table 52* provides the comparisons between the nested models, showing the effects of time (i.e. post-test and delayed test) and condition on both Coherence (upper panel) and Equivalence scores (lower panel).

The upper panel shows that for Coherence only an effect of time was observed.[204] The effect of time was significant for the post-test (T2)[205] and the delayed test (T3),[206] compared to the pre-test (T1), but not for the delayed test (T3) compared to the post-test (T2).[207] This implies that the effects of both conditions on T2 were maintained at T3. In other words: participants significantly improved the coherence of their translation as a result of the lessons in both conditions. This improvement lasted in these tests. Moreover, the *process condition* did not have a differential effect on Coherence. This contradicts our

---

[203] *The scores per measurement occasion are collected with two translations per test session. Reported in table 18 are the mean scores per test session. The scores on the pre-test session (T1) did not differ significantly* (Pillai's Trace (2, 78) = 0,016, p = .984).
[204] *(X^2) = 77.19, p < .001.*
[205] *(t(162) = -1,430, p <.001, ES = 1.37).*
[206] *(t(162) = -1.377, p < .001, ES = 1.32).*
[207] *(p = .75).*

hypothesis that the coherence scores in the *process condition* would significantly improve compared to the *product condition*.

Table 52: Effect of Condition and Time on Coherence and Equivalence

Coherence						
Models	-2loglikelihood	Df	Comparison	X^2	df	*p*
(0) Y = C + [variances]	875.272	3				
(1) + Time	798.078	5	0 vs 1	77.194	2	<.001
(2) + Condition	797.888	6	1 vs 2	0.190	1	.663
(3) + Time * Condition	795.452	8	2 vs 3	2.436	2	.296

Equivalence						
Models	-2loglikelihood	Df	Comparison	X^2	df	*p*
(0) Y = C + [variances]	1088.228	3				
(1) + Time	974.533	5	0 vs 1	113.695	2	<.001
(2) + Condition	974.311	6	1 vs 2	0.222	1	.638
(3) + Time * condition	965.960	8	2 vs 3	8.351	2	.015

The lower panel shows a significant effect of an interaction between condition and time for Equivalence.[208] This implies that the effect of time varied for conditions. At T2, both conditions scored significantly higher on Equivalence than at T1.[209] At T3, the interaction effect between time and condition[210] indicates that the effect of time differs significantly between both conditions:[211] *process condition* scored significantly lower on Equivalence than *product condition* (see *Table 51*: ES = .54). However, both conditions improved their initial equivalence scores, as both scored significantly higher at T3 than at T1.[212] This supports the hypothesis that, though Equivalence in both conditions improved as a result of the lessons, the *product condition* has an extra effect on equivalence scores at T3.

---

[208] *(X²) = 8,351; p < .015.*
[209] *(t (81) =-9.14, p < .001).*
[210] *(F (1,81) = 10.88, p = .001).*
[211] *(F (1, 81) = 46,26, p < .001).*
[212] *(F (1, 81) = 51,44, p < .001).*

3.2 Explorations: Do the Effects of Learning Conditions Differ for Specific Learner Characteristics?

To explore whether effects of conditions were different for learner variables, we analysed the effects of time and condition for T2 and T3. Four learner scores were available as indicators for estimating differential effects: Coherence-score at T1, Equivalence-score at T1, and the self-reported levels of Latin and Dutch proficiency. Here, we report the effects of the two proficiency scores.

3.2.1   Effects of proficiency levels Dutch and Latin.

In this subsection we report the effects of the two proficiency scores as indicators of rather general proficiency, based upon all kinds of tests at school. When an interaction component that includes the learner variable in these models proves to be significant, it would indicate that the effect of condition is different for levels in the learner variable, as we mentioned above (Section 2.7). Appendix D (table LXIII and LXIV) provides the model comparisons.

  Dutch proficiency scores did not contribute to the Coherence scores nor to the Equivalence scores on T2 and T3, which implies that Dutch proficiency does not predict the variance in both outcomes. Moreover, proficiency in Dutch did not interact with Condition and Time, indicating that the effect of condition on Coherence and Equivalence did not depend on proficiency in Dutch.

  Latin proficiency scores, however, affect Coherence as well as Equivalence scores, indicating that proficiency in Latin predicts scores on both dependent variables. The effect of Latin proficiency is a main effect: it does not interact with Condition or Time.[213] The effect is positive, which means that the higher the score on Latin proficiency, the higher the scores on Coherence (*Figure 35A* for T2) as well as on Equivalence (*Figure 35B* for T2) in *both* conditions.

---

[213] *On Coherence: β = .50,* t *(81) = 4.43,* p *<.001; on Equivalence; β = .59,* t*(81) = 3,58,* p *=.001.*

Figure 35: Contribution of Proficiency Latin to Condition Effects at T2.

35A. Effects on Coherence scores
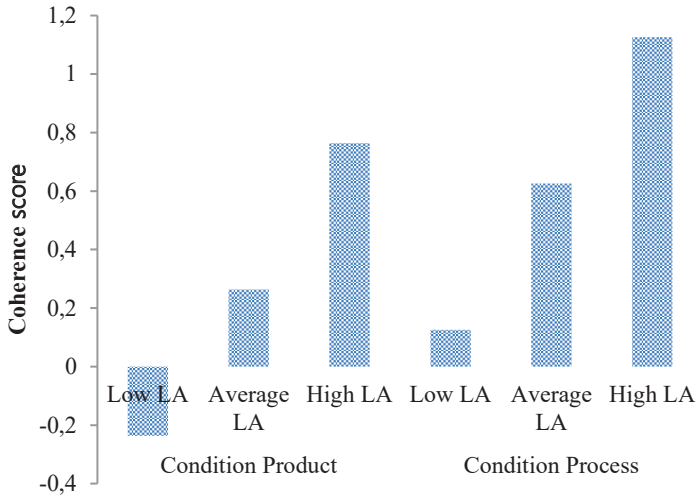
35B. Effects on Equivalence scores
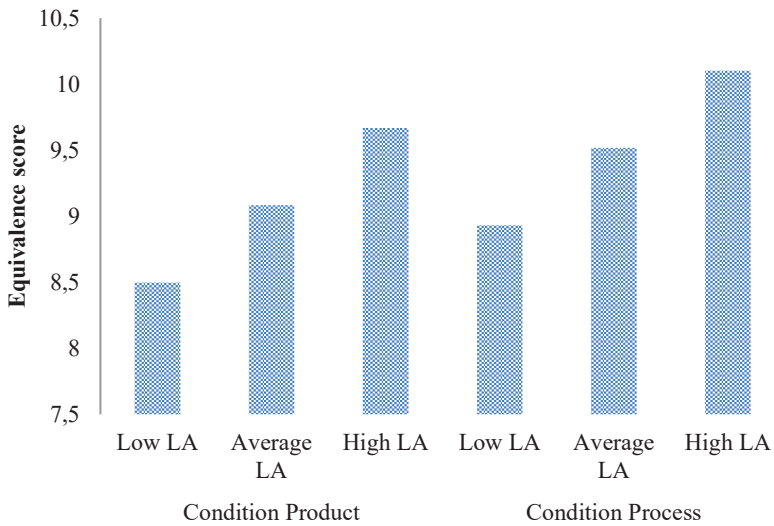
To conclude, the condition effects are not moderated by Dutch and Latin proficiency, which would have meant that students with a particular level of Dutch or Latin would profit more from one or the other condition. The effects reported in Section 3.1 can be generalised across levels of Dutch and Latin.

3.2.2   Differential effects for initial performance on Coherence and Equivalence.

In the pre-test (T1), all participants received a coherence score and an equivalence score, based on two translations they produced during test session 1. We will discuss the effects of the two learning conditions on T2 and T3 with the Equivalence score of T1 and the Coherence score of T1 added as learner variable separately. The model comparisons are reported in Appendix E.

Effects of Coherence at T1 on Coherence. For the scores of Coherence at T2 and T3 we found main effects of Coherence at T1, which means that Coherence scores in the pre-test (T1) were an indication of Coherence scores at T2 and T3 irrespective of conditions and measurement occasion.[214] *Figure 36* shows that effects of the T1-Coherence score on Coherence at T2 and T3 are not condition specific. The effects do not differ at T2: those who started with higher coherence scores on T1 performed better at T2 and T3. The level of T1-Coherence played the same role in both conditions for Coherence.

Effect of Equivalence at T1 on Coherence. For the effect of Equivalence T1 scores on Coherence on T2 and T3 we found main effects, which indicate that in both conditions participants with higher initial (pre-test) scores for Equivalence scored significantly higher on T2 as well as on T3 on Coherence.[215] We did not expect a differential effect of initial Equivalence scores on Coherence for both conditions. *Figure 37* shows that the initial scores on Equivalence affect the scores on coherence not differentially. In both conditions, participants with higher scores on Equivalence at T1 scored higher on Coherence than participants with lower scores. This indicates that in both conditions, the initial level of Equivalence contributed to Coherence.

---

[214] *(t(81) = 5,18,* p *< .001, β = .38).*
[215] *EquivalenceT1 (*t*(81) = 3 .49,* p *= .001, β = .21.*

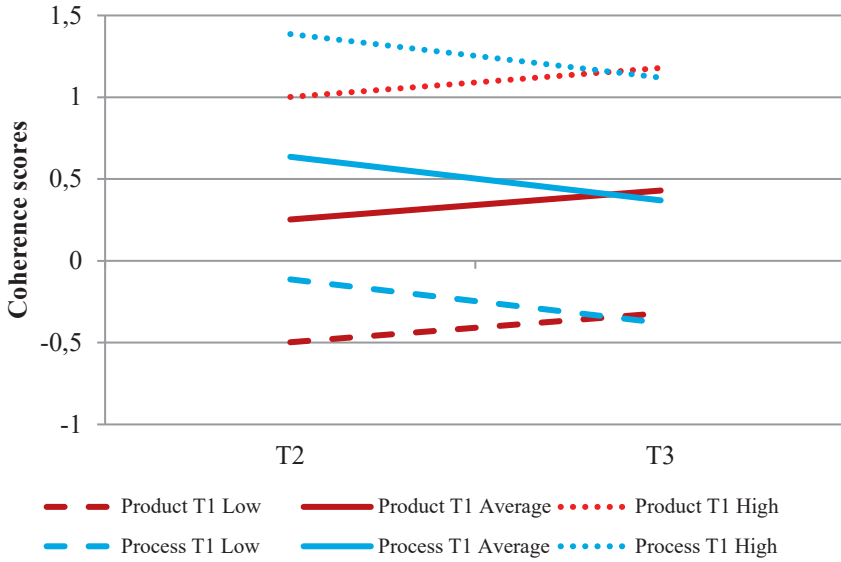*Figure 36: Initial Coherence Proficiency on Condition Effects (Coh).*

Figure 37: Initial Equivalence Proficiency on Condition Effects (Coh).
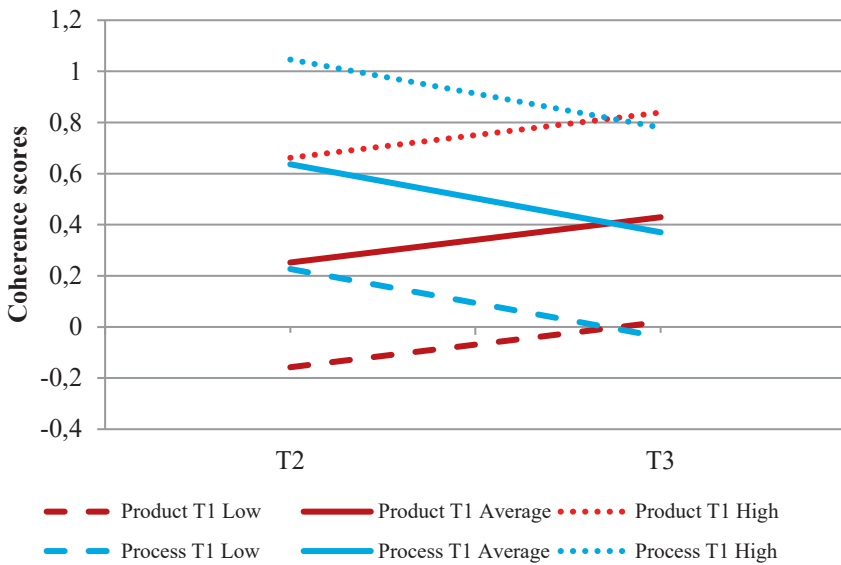
Effect of Equivalence at T1 on Equivalence. For Equivalence, the analysis with Equivalence T1-scores as learner variables showed that the initial score on Equivalence was involved in three effects:[216]

1.   A main effect of EquivalenceT1, which means the scores on Equivalence in pre-test (T1) were an indication of Equivalence scores at T2 and T3.
2.   An interaction between this effect and condition indicating that the effect of the initial level of Equivalence varied across conditions, and
3.   A three-way interaction between the initial Equivalence score, time and condition, indicating that the contribution of initial Equivalence scores varied across measurement occasions and conditions.

Figure 38 shows the effect of initial Equivalence scores on Equivalence in T2 and T3 per condition.

At T2 the effect of the learning conditions is non-significant for participants who scored initially average on EquivalenceT1: whether these participants followed the *product* (red) or the *process* (blue) *condition*, the scores on Equivalence at T2 do not differ.

Within the *product condition*, we observed that participants who scored relatively high at T1 still scored relatively high at T2 and low scoring participants at T1 scored relatively low on T2: the rank order at T1 on Equivalence has not been changed at T2 as a result of the learning condition.

In the *process condition* however, the effect of initial Equivalence level is absent at T2. We observe that this condition, which focused on coherent target-text production, has changed the rank order on Equivalence between T1 and T2 as a result of the learning condition: the scores at T2 have evened out and are approaching each other closely. This finding is unexpected; an experiment aimed at improving *coherence*, has had an effect on the rank order of *equivalence* scores.

T3 indicates whether effects have persisted 4-5 weeks after completing the experiment. At T3 the effect of EquivalenceT1 is present again: high scoring participants in T1 scored relatively high in T3 and low scoring participants in T1 scored relatively low again in T3. For the average scoring participant on EquivalenceT1 the effect of the learning condition *Product* is still not significant.

---

[216] 1) $F_{(1, 81)} = 13,383$, $p < .001$, $\beta = .23$.
2) *EquivalenceT1*Condition:* $F_{(1, 81)} = 5,849$, $p < .018$.
3) *EquivalenceT1*Time*Condition* ($F_{(1,81)} = 12.556$, $p = .001$, $\beta = .74$.

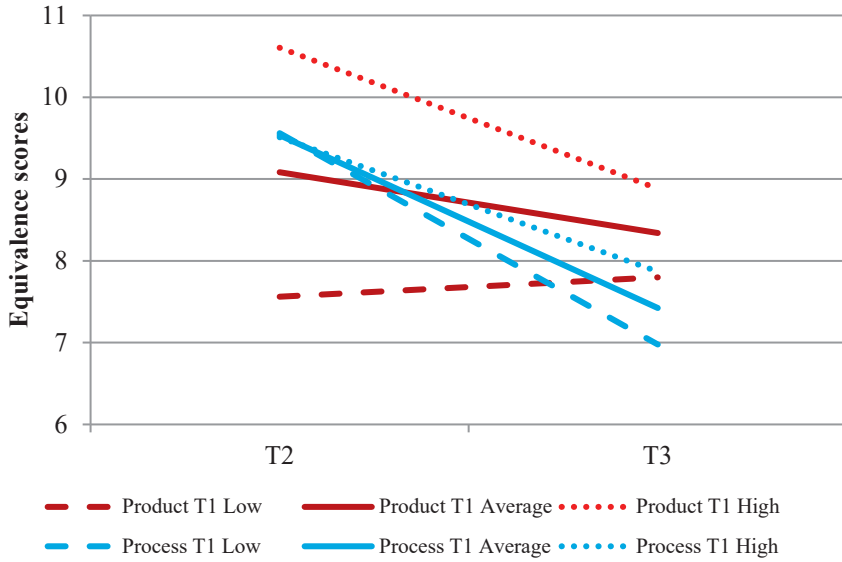*Figure 38: Initial Equivalence on Condition Effects (Equi).*

Figure 39: Initial Coherence on Condition Effects (Equi).
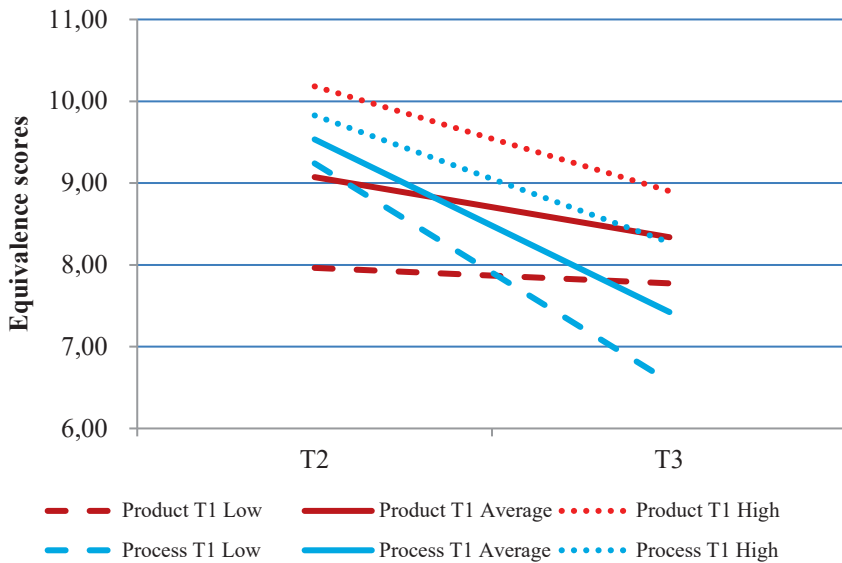
Effect of CoherenceT1 on Equivalence in both conditions. We found a more complex pattern of the effect of CoherenceT1 on the scores of Equivalence, the contrasting T1-variable. We observed that on T3 scores are significantly lower than on T2, which is an effect of Time,[217] and we observed the product condition scoring higher on Equivalence than the process condition on T2, which is an effect of the interaction of time and condition.[218] Furthermore, we observed a main effect of CoherenceT1[219] and a three way interaction between CoherenceT1*Time*Condition.[220] This effect indicates that the role of CoherenceT1 varies across conditions and measurement occasions. We explore this variation below. *Figure 39* shows the effect of initial Coherence scores on Equivalence at T2 and T3 per condition.

For average and high-scoring students on Coherence at T2, no significant effect of conditions on Equivalence was observed. However, for low-scoring participants we noticed a significant difference between the two conditions at T2. Participants scoring low on CoherenceT1, scored higher on Equivalence at T2 when they were trained in the process condition than in the product condition, indicating that these weak students benefitted from instructional attention towards Coherence and improved the ST-TT Equivalence at T2. This indicates transfer from Coherence training to Equivalence results for this particular group of students.

At T3, the initial lower scoring students on CoherenceT1 seem to be best of in the product condition for this variable.

3.3 Summary of Interaction Effects

Table 53 presents the summary of the interaction effects for two categories: similar and contrasting input-output variables.

The most remarkable interaction effect is the effect on Equivalence at T2 (*Figure 38* and *Figure 39*). The differences between students at the start of the experiment on Equivalence were muted at T2 and T3 in the *process condition* only. Students with relatively low scores on Equivalence or Coherence at the start of the lessons scored higher on Equivalence at T2 than their counterparts in the product condition, which aimed at improving Equivalence. So, for relatively weak students regarding Equivalence, opting for a course on *coherence* would be a better instructional choice than intensive error analysis. Regrettably, this effect is absent in the delayed post-test.

---

[217] F *(1,81) = 48,908*, p *< .001;* β *= 2,1.*
[218] *Time*Condition:* F*(1,81) = 11,435*, p *=.001,* β *= -1.38 (*t *(81)= -1,38*, p *=. 001.*
[219] F*(1,81) = 10,320*, p *= .002,* β *= .44.*
[220] F*(1,81) = 4.161*, p *= .045,* β *= .56.*

Table 53: Summary of Effects of T1-Level on Post-Test Scores

Outcome variable	Learner variable	Effects per Condition	Figure
Coherence	Latin	Effect of Latin similar (both conditions)	35A
	Dutch	No effect	
	Coherence	Main effect of T1-level: differences at	36
	Equivalence	T1 are reflected at T2/T3, for both conditions, for both T1-variables	37
Equivalence	Dutch	No effect	
	Latin	Effect of Latin similar (both conditions)	35B
	Equivalence	In the process condition, the differences	38
	Coherence	due to T1-level are muted, for both T1-variables. Students with relatively low T1-scores were better off in the process condition than in the product condition.	39

4   CONCLUSION

In the pre-test, correlation between both coherence score and equivalence score with self-reported Latin grade is significant in both conditions, while correlation between both scores and Dutch grade is not significant in both conditions. This supports the assumption that the Latin grade reflects proficiency in Latin translation, while the grade for Dutch is not related to the quality of a translation.

Effects in post-test and delayed test can reliably be attributed to the experiment, as fidelity of implementation is accurate, and scores in pre-test do not significantly differ between conditions. Therefore, the three questions I posed in the introduction to the present chapter can now be reliably answered.

1. Do participants in the *process condition* produce more coherent target texts in post-test and delayed test, compared to participants in the *product condition?*

No, Coherence in the *process condition* does not significantly improve in post-test and delayed test compared to the *product condition*. However, coherence improves significantly at T2 in both conditions, which effect remains at T3 (*Table 52*, upper panel).

2. Do participants in the *product condition* produce more equivalent target texts in post-test and delayed test, compared to participants in the *process condition*?

No, Equivalence in the *product condition* compared to the *process condition* does not significantly improve in post-test and delayed test. However, Equivalence in both conditions improves significantly at T2. The *product condition* does have an extra effect on Equivalence scores at T3 (*Table 52*, lower panel).

3. To what extent do the effects of conditions depend on initial proficiency in Latin or Dutch (self- reported and scores at pre-test)?

Dutch proficiency had no effect, *self-reported Latin proficiency* had a similar effect in both conditions. *Coherence T1* score was an indication for Coherence T2 and T3 in both conditions (*Figure 36*). The effects of CoherenceT1 on Equivalence vary across conditions and measurement occasions. We found that students with relatively low Coherence T1-scores were better off in the *process condition* than in the *product condition to* improve their Equivalence score at T2. This indicates transfer from Coherence training to Equivalence results for these students (*Figure 39*).

    EquivalenceT1 contributed to Coherence at T2 and T3 (*Figure 37*). EquivalenceT1 also was an indication of Equivalence at T2 and T3. However, we observed that initially weak, average and strong Equivalence scorers approached each other at T2 in the *process condition* (*Figure 39*). This effect was not maintained at T3.

    We observed a decline of effects of CoherenceT1 and EquivalenceT1 at T3 in the *process condition (Figure 38 and Figure 39)*. This decline of effects could be explained by the assumption that participants need more time to grow accustomed to the process-oriented translation strategy. The content of the lessons in the *process condition* was new for all participants. The *product condition*, on the other hand, provided lessons that were connected to the activities participants were more familiar with.

    This may seem to be a disappointing outcome at first glance, but it actually is a very promising outcome for Latin teaching practice. In fact, four well-designed lessons focusing on either translation *product* or translation *process* significantly improved Coherence as well as Equivalence of the target texts participants produced. The design principles that were the foundation of the lessons in both conditions can therefore be qualified as successful.

These results seem to contradict the common belief that more focus on linguistic analysis of the source text or more practice is the only way to learn to produce more coherent or more equivalent translations. It is also clear that

writing coherent texts for translations is not a result of writing skills in the target language, as no significant correlation between Dutch grade and quality scores was found.

CHAPTER 7
GENERAL DISCUSSION

1   PROBLEMS IN TEACHING LATIN TRANSLATION

The problem this dissertation addressed was that Dutch students in upper secondary education generally struggle to show their understanding of a Latin source text when they translate it into Dutch, although this is a requirement formulated by the programme for the Dutch central final examinations (domain A, sub-domain 1). Their lack of understanding is demonstrated by the incoherent target texts students produce when they are translating Latin (*Table 54*).

Table 54: Example of an Incoherent Target Text[221]

De hond die vlees over de rivier draagt.
Hij verliest terecht het eigene die het van een ander na streeft.
De hond, terwijl hij het vlees over de rivier draagt, ziet terwijl hij zwemt van de heldere wateren in de spiegel zijn spiegelbeeld, menend dat door een ander met een wapen de buit werd geroofd, maar nadat hebzucht hem had bedrogen en dit met de mond te voelen verloor hij voedsel, niets verlangde hij zozeer om dit te kunnen aanraken.

The dog who carries meat across the river.
He rightly loses his own, who strives for it belonging to another. The dog, while carrying the meat across the river, sees while swimming of the clear waters in the mirror its reflection, thinking that by another with a weapon the booty was stolen, but greed having deceived him and to feel this by/with the mouth he lost food, nothing he desired as deeply to be able to touch this.

To investigate possible solutions to this problem, the overall research goal of this dissertation was to design and test an intervention that would teach students to produce more coherent texts when they translate Latin into Dutch. The target-text coherence that we wanted to improve was defined in chapter one as *intratextual coherence*.

I analysed the problem of insufficient coherence of students' target texts by connecting it to four problems in the teaching practice of Latin translation

---

[221] *Dutch translation by participant in the experimental study (Chapter 6) [English translation below by SL].*

in relation to translation studies and educational science. The first problem was the *translation assignment* students are given. Translation studies provided the insight that part of the problem lies in the insufficient clarity of the *skopos* of the translation assignment in Latin class (Chapter 1). Students are asked to translate Latin source texts to demonstrate their proficiency in Latin, not to make content available to a public that needs a translation to understand the source text. As such, the translation assignment in Latin class hardly ever reflects a real-world problem. To address this issue in this dissertation, I formulated a translation assignment for the experiment that is more specific as to the expected outcome of the translation action than usual, considering that the production of a coherent target text has a purpose in the real world, while the production of an *in*coherent target text serves no fathomable purpose whatsoever. I turned the assignment into a more or less real word problem, by asking students to "*produce a coherent target text.*" explaining to them that a person without knowledge of the Latin source text should be able to understand the text. How to achieve this goal remained to be seen.

The second problem in the teaching practice of Latin translation was related to the content and instruction method of lessons teaching translation. As such, this problem was related to educational sciences. I did not know *what* I had to teach students to make them produce more coherent target texts. This problem was connected to an underlying problem: I had little knowledge of the translation process in high-school students leading to coherent target texts. Learning activities for producing a coherent as well as an equivalent target text were not yet defined, as a result of which a method or strategy to translate source texts into coherent target texts simply was not available. Again, I looked into translation studies to gain insight in the components of translation competence and translation process of adult and professional translators (Chapter 1). This led to the similarities between translation and writing competences (Schrijver, Van Vaerenbergh, Leijten & Van Waes, 2018) and the four phases of the translation process: orientation, target text production 1, target text production 2, and revision (Breedveld, 2002). I then wondered what the translation activities of successful *student* translators of Latin were and what successful strategies these *students* used to produce a coherent target text. This problem, the gap in existing knowledge of the translation process that leads to coherent target-text production by students, was addressed in chapter three. Looking through the eyes into the head of successful student translators via eye-tracking, I tentatively identified common characteristics of their translation activities. A successful translation process is characterised by 1) source text comprehension, 2) a broad range of types of knowledge, 3) the ability to consciously switch both between these types of knowledge and between problem-solving strategies, and 4) the ability to reflect on the translation

process. The use of metalanguage to phrase this conscious decision-making was observed in successful students. Based on literature on the position of revision in the writing process (Van den Bergh, Rijlaarsdam & Breetvelt, 1994) which is, as I argued in Chapter 1.4.2, in many ways similar to the translation process, I started looking into students' revision activities and found that all successful student translators put effort into revising drafts. I operationalised these findings as five skills:

-   the skill to reflect on the translation process,
-   the skill to use metalanguage to guide the translation process,
-   source-text comprehension skills,
-   target-text production skills,
-   target-text revision skills.

I combined the development of these skills in four components, similar to the four phases of the translation process of professional translators (Breedveld, 2002), and connected each component with the translation activities observed by Breedveld (2002). Thus, I developed a new process-oriented translation strategy (*Figure 40*) that formed the content of the lessons I designed for teaching to produce coherent target texts.

Figure 40: Process-Oriented Translation Strategy.[222]

---
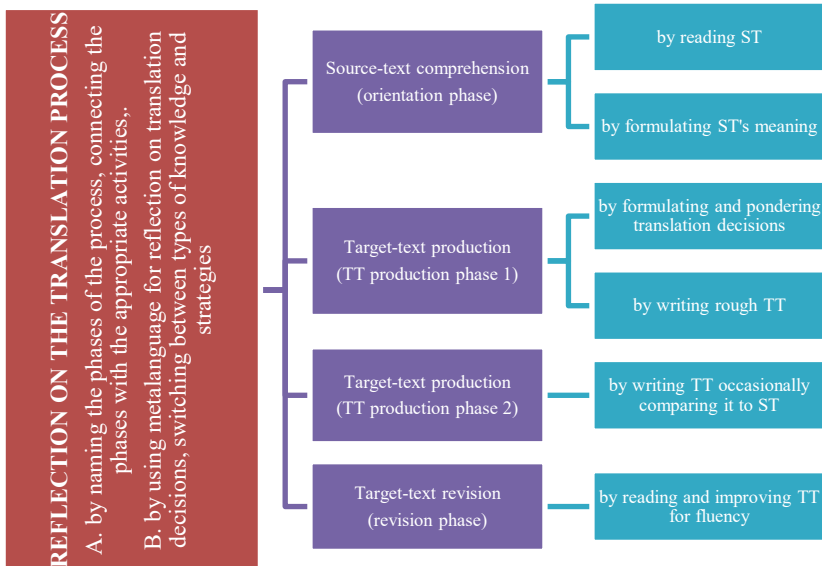
[222] *Presented in Chapter 4 as Figure 12.*

The third problem in the teaching practice of Latin translation was that an evidence-based method to teach the translation of a Latin source text into a coherent target text is not available. Again, the similarity between the translation and writing competences was helpful in addressing this problem, as was the connection with educational sciences (Chapter 1). Studies on efficient methods to teach writing were amply available, leading to the decision to use strategy instruction, which is considered to be most effective for writing instruction (Graham & Perin, 2007) and apply it to translation instruction (Chapter 4). I designed the lessons aimed at teaching the process-oriented translation strategy based on three design principles:

- Self-Regulated Strategy Development (Graham & Harris, 1996, 2002) was a framework for the design of the lessons instructing a process-strategy based on four translation phases.[223]
- Four of Graham & Perin's evidence-based recommendations (2007) in the design of lessons teaching coherent target-text production were used by setting clear and specific goals for what participants were to accomplish with their translation product. Thus, the lessons addressed the issues with the translation assignment and *skopos*, and offered activities to gather and organise ideas about the content of the source text before students wrote a first target text. Graham & Perin's recommendations overlap with Self-Regulated Strategy Development in that they also recommend strategy instruction and modelling for writing instruction.
- I used learning activities aimed at raising awareness of elements of the process-oriented translation strategy to regulate the development of two main sub-competences (Göpferich, 2008): 1) strategic competence concerning the *process* of translating, and 2) communicative competence concerning *meaning and coherence* of source text and target text.

The fourth problem in the teaching practice of Latin translation in the Netherlands is the rating method. Colon rating, the most commonly used method for rating translations, is a reliable method to assess ST-TT equivalence on colon level. However, it provides insufficient assessment of the coherence of targets texts. Using the definition of target-text coherence as *intratextual coherence* I opened the way to assessing the target-text coherence by rating its coherence independently from the source text. I established that comparative holistic assessment is a reliable method to assess intratextual coherence quality. Comparative assessment is fairly easy to apply using D-PAC as a tool for comparative judgment, and highly reliable, provided that sufficient comparisons are made by a varied team of experienced assessors (Chapter 2).

---

[223] *Orientation, text production 1, text production 2, revision.*

2 DESIGNING LESSONS TO TEACH LATIN TRANSLATION

Having thus established for the first time a more specific translation assignment, a process-oriented translation strategy, an instruction method, and a method to reliably assess target-text coherence, I designed lessons aimed at improving target-text coherence. Other than the traditional product-oriented approaches to translation, these lessons started by raising awareness of the translation process by scoring, categorising and comparing translation activities. Focus on meaning was achieved by visualisation of the text through drawing in the orientation phase and awareness of translation activities (text production and revision) by verbalising the activities, formulating strategies, giving and receiving feedback on translation activities and classroom discussion. The lessons raised awareness of text coherence by comparing target texts to assess coherence, revising the least coherent text into a more coherent text, comparing the revised texts, and revising a self-produced target text.

I designed an experiment in two conditions, experimental and control, with pre-test and post-test (Chapter 4), tested this experiment (Chapter 5) and encountered a serious problem regarding the stability of the control condition. It became clear that 'business as usual' in combination with the instruction of error analysis did not provide sufficient direction for the lessons. Therefore, I decided to reconsider the position of the control condition and to redesign the experimental study as a competition between two interventions with different effects on two qualities of a translation. The first focused on improving target-text coherence through improving the translation process, while the second focused on improving the translation product, more specifically ST-TT equivalence through error analysis. I renamed the conditions *process condition* (the former experimental condition) and *product condition* (the former control condition). While all lessons needed some refurbishing, the design principles in the newly named *process condition* were valid and could remain the same. The lessons themselves were redesigned in both conditions to address the issues that were experienced in the testing with lacking fidelity of implementation by reducing the complexity of the assignments and improving instructions for teachers. Additionally, and specifically in the newly named *product condition*, I reduced the number of fables that had to be translated and designed more varied lessons. This way an experimental study was set up to find an answer to the initial research question in the *process condition*:

> Will instruction of a strategy that focusses on the improvement of the translation *process* and that is aimed at coherent target-text production, revision, and metalanguage lead to the production of more coherent target texts than a traditional instruction, focusing on the translation *product* and aimed at developing knowledge and linguistic understanding of Latin?

Additionally, the research question for the *product condition* was formulated:

> Will an instruction, focusing on the translation *product* and aiming at de-
> veloping knowledge and linguistic understanding of Latin through error
> analysis lead to better ST-TT equivalence than instruction of a strategy that
> focusses on the improvement of *process* and that is aimed at the develop-
> ment of coherent target-text production, revision, and metalanguage on
> translation?

3   RESULTS OF THE EXPERIMENT

The results of the experiment showed that after an intervention of four lessons
(eight hours in total), participants significantly improved both target–text co-
herence and ST-TT equivalence in both conditions. As was expected, the les-
sons in the *process condition* teaching a process-oriented translation strategy
resulted in a significantly improved coherence in target texts in post-test.
However, in the *product condition,* target–text coherence improved signifi-
cantly as well, although in this condition lessons were aimed at improving ST-
TT equivalence by error analysis. The difference of coherence scores at T2
was not significant between conditions. The effect on Coherence remained in
the delayed test for both conditions. The improved coherence score in the
product condition concurs with the belief that an equivalent translation of a
coherent source text leads to a certain level of coherence in the target text.

In the *product* condition*, ST-TT equivalence improved significantly in the
post-test, as expected. ST-TT equivalence improved significantly in the *pro-
cess condition* as well, which was surprising because no specific instruction
was given on equivalence of translations or on Latin in the lessons. These
effects in post-test appear to be in stark contrast with the belief that more focus
on Latin is the only way to teach students to produce more equivalent transla-
tions. The effects on equivalence in the *product condition* remained in the de-
layed test. The effects on equivalence in the *process condition* decreased in
relation to the post-test, while remaining higher than in the pre-test. Based on
the observed decline of effects in the delayed test in the *process condition*, it
seems that participants may need more time to grow accustomed to (elements
of) the process-oriented translation strategy. It would be interesting to see
what happens when the process-oriented translation strategy is more inte-
grated in the regular lessons.
Additionally, I found that writing coherent texts for translations is not a result
of proficiency in the target language, as no significant correlation between
Dutch grade and coherence scores was found in the tests.[224]

---

[224] *The reported grades for Dutch were quite homogenous.*

Summarising, this experimental study has established that the process-oriented translation strategy offers a new and effective approach to translation. The correlation between Latin proficiency and pre-test scores confirms that translation always also relates to source-language proficiency. Therefore, the focus on the Latin source text should not be neglected, as is supported by the positive effects on both coherence and equivalence in the *product condition*.

4   VALIDITY

The first concern was to establish whether the process-oriented translation strategy or error analysis would have effects. Based on the design pilot I decided to conduct the experimental study in a laboratory setting, to guarantee internal validity. Therefore, the setting in which the experiment was conducted was not a natural classroom setting, as a result of which the ecological validity could arguably suffer: the chosen setting might affect participants' motivation, making it higher than usual, as well as classroom practice, which could become more task-oriented than usual. I did not perceive this as a problem, as these were exactly the aspects of the lessons I wanted to improve based on the pilot. As to motivation, I designed the recruiting of participants to obtain some diversity in motivation as follows:

-   I targeted *weak translators* by presenting the experiment as a turbo-translation training.
-   I targeted *strong translators* by stressing the importance of participating in a scientific experiment.
-   I targeted students with *low intrinsic motivation* by offering payment for participation.

This has resulted in a diversely motivated group of participants, as is indicated by the reasons given to participate varying from "I like the payment" to "I need tutoring for Latin and might as well get paid for it" and "I want to help develop methods for teaching Latin translation." Furthermore, the mix of participants from different schools ensured that effects could not be attributed to participants' shared experiences with learning Latin translation. Additionally, this mix minimised social interaction and thus contributed to high time on task activity during the classes. Over the weeks, participants became acquainted and socialised during the break, but this did not seem to influence performance or classroom discipline.

    To avoid that teachers' performance would vary over condition if the same teacher would teach both conditions, the four groups were each taught by a specific teacher. All teachers were highly motivated to participate in the experiment.

Summarising, as previous evidence-based research on how to teach Latin translation was not available, I have for the first time reliably established effects on target-text coherence or equivalence by using a highly controlled laboratory setting.

5   HOW TO USE THESE RESULTS
FOR TEACHING LATIN TRANSLATION

The process-oriented translation strategy developed in this dissertation is new for Latin translation class. Before this study, learning activities for producing a coherent target-text were not yet clearly defined, so progress was made in teaching Latin translation by designing a strategy, linking the strategy to learning activities and using a tested method to instruct the strategy. A tested strategy to translate source texts into coherent target texts is now available. It is established that both target-text coherence and ST-TT equivalence improved significantly after a relatively short intervention of four lessons. It is also established that even participants with low initial equivalence scores improved significantly in the condition that taught the process-oriented translation strategy. This suggests that a shift in focus from source text to target-text production through the process-oriented translation strategy could lead to the production of translations with higher ST-TT equivalence. The waning effect in the delayed test suggests that a longer period of practice with the strategy is needed for weaker students.

As to visible changes in the translation process of participants, firm conclusions cannot be drawn as all data on the translation process I discussed were self-reported by answers to the questionnaire. The analysis of the screencast data that I collected as a back-up proved to be too time-consuming for the scope of this dissertation. However, the data are still available and could provide interesting material for a future study to provide background for the translation activities participants reported in the questionnaire.

Now that I have reliably established effects, the next step is to research the generalizability of the outcomes and test the lessons in regular classes. Following that, the lessons could be expanded to other source-text genres. The experiment used fables as source texts, and it is possible that the improvement is genre specific: participants may have improved their translation of fables as opposed to Latin in general. Therefore, a possible subject for further research is to test the method we designed on different types of texts with different genre characteristics.

Secondary education in the Netherlands is moving towards programme differentiation.[225] The findings of this study could easily fit into this

---

[225] *Whether this is a good thing is a debate outside the scope of this dissertation.*

development. Teachers could start by labelling lessons a "translation train-ing." The translation lessons could be further fine-tuned by intertwining pro-cess lessons and product lessons and integrating references to the process-ori-ented translation strategy in lessons following the training. Programme differ-entiation could be achieved by starting the training programme with a transla-tion task to assess the translation process for each student individually, using the interval assignment as performed in the lessons in the *process condition*. Additionally, each student would receive an initial target-text quality level by coherence and equivalence scores. Based on an analysis of their target-text quality and translation process, students could determine individual learning goals to improve their translation process, target-text coherence or equiva-lence and follow a personalised route to reaching their goal.

In the past five years, I conducted workshops on translating Latin in class, introducing the interval assignment to Dutch teachers of Latin on several oc-casions. As a result of these workshops, I have seen that some teachers who participated in my workshops on translating are taking steps toward the pro-cess-oriented approach. They have responded enthusiastically to the interval-scoring assignment and reported that they used it in their classes, successfully raising awareness of all activities that translation involves. Students reportedly appreciate the process-oriented approach as well, as it offers a whole new per-spective to the translation assignment.

6   THE FUTURE FOR THE LATIN TRANSLATION ASSIGNMENT

Discussing the future of the Latin translation assignment is a somewhat peri-lous enterprise, as it touches on deeply rooted sensitivities in Dutch classicists. Suggestions to change the translation assignment in the central national exam-inations (Kroon & Sluiter, 2010) caused heated debate. Personal friendships were ended as irreconcilable differences concerning the "cornerstone of Latin secondary education" – as I called the translation assignment in chapter one – drove a wedge between opponents.[226] Therefore, I proceed with great caution as I phrase my hopes and concerns for the future of the Latin translation as-signment in Dutch education.

First and foremost, I would like to state that I personally like the act of translating. I like to think about the meaning of a source text and to phrase and rephrase my translation in an attempt to capture the meaning of the original text. I like to ponder the differences between the Ancient world the source text originates from and my own modern world. I like to try and capture both the content and the style of the original text and to discuss all these aspects of translating with my students. This discussion and the decisions we make based

---

[226] *Source: personal communication with classicists in the Netherlands (2011).*

on this discussion, to me, are the reason to translate texts as opposed to merely reading them.

However, I do not specifically like translations. I would like to illustrate the importance I attribute to reading texts in the source language by an anecdote. I once was looking for a present for an English friend, and obviously I turned to a bookstore. As I wanted to share some of my Dutch heritage, I looked into translated Dutch classics and found a translation of the Dutch novel by Gerard Reve *De Avonden* ("The Evenings"). I opened the book and read the first sentence. I immediately decided to buy my friend some other gift: the translation simply did not reflect what I thought Reve had written.

Having stated my position on translating and translations, I have to address the question: why do we translate Latin in school? I believe the aim is reflected by the phrasing of the programme for the Central Final Examination in Latin: "The candidate can demonstrate his comprehension of Latin (…) by translating an unseen passage."[227] The ultimate aim of translating Latin in secondary education is to reach a profound understanding of what is written, how it is phrased, and what it means to us by discussing with our students the source texts, translation options and decisions, and how to capture this source text in a target text. This is no easy task, as we use literary texts originating from a distant era, with customs we are not familiar with. Therefore the student as well as, in many cases, the teacher needs scholarly guidance and a keen interest. This practice leads to a type of close reading that the mere reading of a text never could accomplish. This may be the reason that Latin teachers hardly ever felt the need to specify the *skopos* of the translation assignment: the act of translation itself legitimised the assignment.

The described translation activity occurs in the classroom, in dialogue with the students. But how does this aim of translating Latin relate to using the translation assignment as a test in upper secondary education? In the Dutch national final exam, a Latin source text of approximately one hundred and twenty words has to be translated into Dutch in approximately ninety minutes. This type of assignment, under time pressure, creates a setting where pondering the meaning of the source text, discussing translation decisions, and carefully writing and revising the target text become very difficult, if not impossible. The produced target text is subsequently rated in chunks on equivalence only. In many cases, an adequate and complete rendition of the meaning of the source text is even penalised by rewarding zero points, if the linguistic aspects of the source text were treated too freely.

I feel compelled to stress that, in my opinion, a mere translation assignment is not fit as a test for source-text comprehension in upper secondary

---

[227] Domain A, sub-domain 1 "examenprogramma Latijnse taal en cultuur" (www. examenblad.nl).

education. Based on the current study, I cannot reach any other conclusion than that the future of Latin translation lies in the classroom, in the collaborative effort to discuss the meaning of the source text, and in creating a coherent target text as an outcome of a translation process based on writing and revising.

REFERENCES

Adema, S.M. Latin learning and instruction as research field, *Journal of Latin Linguistics.* Article in preparation.

Adema, S.M., Faber, T., De Groot, R., De Jonge, F., Langedijk, I., Langerak-Wakker, M., … Winterstein, D. (2008). PSOLMO rules. *VCN Bulletin, 127*, 43-47.

Adema, S. M., & Van Gils, L. (2015). Epistularum genera multum - een talige benadering van teksttypen in brieven van Cicero en Plinius [Epistularum genera multum – a language-oriented approach to types of texts in letters by Cicero and Pliny]. *Lampas, 48,* 18-38.

Balbo, A. (2011). Pragmatic aspects of teaching translation methods from Latin to Italian. In R. Oniga, R. Iovino, & G. Giusti, (Eds.), *Formal Linguistics and the Teaching of Latin: Theoretical and Applied Perspectives in Comparative Grammar*, (pp. 371-392). Cambridge Scholars Publishing.

Barkaoui, K. (2010). Variability in ESL essay rating processes: the role of the rating scale and rater experience, *Language Assessment Quarterly*, *7* (1), 54-74, doi:10.1080/15434300903464418

Bartelds, D. (2018). Wat is er moeilijk aan woordenboekgebruik? [What is difficult about using a dictionary?]. *Lampas*, *51,* 235-255.

Bekker, R., & Van Oeveren, C.P.D. (2008). Het slepen voorbij [Beyond dragging]. *Lampas, 41,* 263-273.

Blok, H. (1986). Essay rating by the comparison method, *Tijdschrift voor onderwijsresearch, 11*, 169-176.

Braaksma, M.A.H. (2002). *Observational learning in argumentative writing*. (Doctoral Dissertation). Universiteit van Amsterdam, Amsterdam.

Breedveld, H. (2002). Translation process in time. *Target, 14* (2), 221-240.

Breedveld, H., & Van den Bergh, H. (2002). Revisie in vertaling: wanneer en wat [Revision in translation: when and what]. *Linguistica Antverpiensia 1,* 327-345.

Buswell, G.T. (1928). *A laboratory study of the reading of modern foreign languages*. New York: Macmillan.

Case, L., Harris, K.R., & Graham, S. (1992). Improving the mathematical problem-solving skills of students with learning disabilities: Self-regulated strategy development. *The Journal of Special Education, 26,* 1-19.

Commissie Modernisering Leerplan. (1971). *Memorandum van de Commissie modernisering leerplan klassieke talen* [Memorandum of the committee to modernise the curriculum for classical languages]. *Lampas, 4,* 38-64.

Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction, 2,* 109–142.

De Man, A.G., Te Riele, G.J.M.J., Sanders, J. (1971). *Redde rationem.* Groningen: Noordhoff.

Derix, H., Felten J., Van Gelder, H., Van Gessel, H., Van Mourik, Chr., Tielens, F., Verhoeven, P. (1983). *Klassieken in het voortgezet onderwijs, een discussienota* [Classics in secondary education – a discussion paper]. Enschede: Stichting Leerplanontwikkeling.

Diercks, G.F., & Van Loenen, N.H.C. (1955). *Tirocinium Latinum*. Amsterdam: Unieboek/Het Spectrum.

Examenblad. Retrieved from www.examenblad.nl.

Florian, L. (2015). *Heimliche Strategien, wie Übersetzen Schülerinnen und Schüler?* [Secret strategies, how do students translate?]. Göttingen: V&R Unipress.

Flower, L., & Hayes, J.R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, *32*(4), 365-387. doi: 10.2307/356600.

Futch, O. (1935). A study of eye-movements in the reading of Latin. *The Journal of General Psychology, 13*(2), 434–63. doi:10.1080/00221309.1935.9917896.

Göpferich, S. (2008). *Translationsprozessforschung: Stand—Methoden—Perspektiven* [Translational process research: State of affairs—Methods—Perspectives]. Tübingen: Narr.

Göpferich, S. (2009). Towards a model of translation competence and its acquisition: The longitudinal study 'TransComp.' In S. Göpferich, A. L. Jakobsen, & I. Mees (Eds.), *Copenhagen studies in language 37. Behind the mind: Methods, models and results in translation process research* (pp. 11–37). Frederiksberg: Samfundslitteratur.

Göpferich, S., Jakobsen, A.L., & Mees, I.M. (Eds.). (2009). *Looking at eyes. Eyetracking studies of reading and translation processing. Copenhagen Studies in Language 36*. Frederiksberg: Samfundslitteratur.

Goris, M. (1999). *Special experience in een experiment "lineaire aanpak van vertalen": een passage van Euripides' Bacchanten* [Special experience in an experiment: "linear approach to translation": a passage from Euripides' Bacchae]. Enschede: Stichting Leerplanontwikkeling. Retrieved from www.stilus.nl.

Goris, M. (2002). *Vakdossier klassieke talen* [State of affairs in classical languages]. Enschede: Stichting Leerplanontwikkeling.

Goris, M. (2007). *Vakdossier 2e fase, klassieke talen* [State of affairs 2nd phase, classical languages]. Enschede: Stichting Leerplanontwikkeling.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*(3), 445-476.

Harris, K., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation* (2nd ed.). Cambridge, MA: Brookline Books.

Harris, K., Graham, S., Mason, L., & Saddler, B. (2002). Developing self-regulated writers. *Theory   Into   Practice, 41*(2),   110-115.   Retrieved   from http://www.jstor.org/stable/1477462.

Holz-Mänttäri, J. (1984). *Translatorische Handeln. Theorie und Methode* [Translational action: theory and method]. Helsinki: Suomalainene Tiedeakatemia.

Huisman, A., Rijpstra, C., De Rooij, D., & Simons, M. (2012). Correctie vertaling CE, discussiestuk [Grading the translation of the national exam, discussion paper]. *VCN Bulletin, 147*, 7-12.

Hulstaert, K. (2016). *Latijn: lezen zien begrijpen, de positionele methode in de klas* [Latin: reading, seeing, understanding, the positional method in class]. Didactica Classica Gandensia nr. 52. Gent: Skribis.

Hunt, S. (2018). Latin is not dead, the rise of communicative approaches to the teaching of Latin. In A. Holmes-Henderson, S. Hunt, & M. Musié. (Eds.), *Forward with classics: Classical languages in schools and communities* (pp. 89-107). London: Bloomsbury Academic.

Hvelplund, K.T. (2014). Eye-tracking and the translation process: Reflections on the analysis and interpretation of eye-tracking data. *MonTI: special Issue 1: Minding translation | Con la traducción en mente*, 201-223.

Jansen, A., Struyk, F., & Hunink, V. (2014). *Lectori salutem* (CE Latijn 2015) [Central examination of Latin 2015]. Houten: Uitgeverij Hermaion.

Janssen, D., Van Waes L., & Van den Bergh, H. (1996). Effects of thinking aloud on writing processes. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 233–250). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Janssen, T., Braaksma, M., & Rijlaarsdam, G. (2006). Literary reading activities of good and weak students: A think-aloud study. *European Journal of Psychology of Education, 21*(1), 35-52.

Johnson, L.A., Graham, S., & Harris, K.R. (1997). The effects of goal setting and self-instruction on learning a reading comprehension strategy: A study of students with learning disabilities. *Journal of Learning Disabilities, 30,* 80-91.

Judd, Ch.H., & Buswell, G.T. (1922). *Silent reading: A study of the various types*. Chicago: University of Chicago Press.

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Krings, H.P. (2005). Wege ins Labyrinth — Fragestellungen und Methoden der Übersetzungsprozessforschung im Überblick [Path into the labyrinth — Overview of questions and methods of translation process research]. *Meta: journal des traducteurs / Meta: Translators' Journal*, *50*(2), 342-358.

Kroon, C.H.M. (2007). *Inleiding tot de Latijnse syntaxis* [Introduction to Latin syntax]. Amsterdam: Amsterdam University Press.

Kroon, C.H.M., & Sluiter, I. (2009). *Tussenrapport van de verkenningscommissie klassieke talen* [Interim report of the classical languages exploratory committee]. Enschede: Stichting Leerplanontwikkeling.

Kroon, C.H.M., & Sluiter, I. (2010). *Het geheim van de blauwe broer*: *Eindrapport van de verkenningscommissie klassieke talen* [The secret of the blue brother: Final report by the classical languages exploratory committee]. Leiden/Enschede: Stichting Leerplanontwikkeling.

Lesterhuis, M., Van Daal, T., Van Gasse, R., Coertjens, L., Donche, V., & De Maeyer, S. (2018). When teachers compare argumentative texts. Decisions informed by multiple complex aspects of text quality. *L1-Educational Studies in Language and Literature*, *18*, (1-22). doi: https://doi.org/10.17239/L1ESLL-2018.18.01.02.

Lesterhuis, M., Verhavert, S., Coertjens, L., Donche, V., & De Maeyer, S. (2017). Comparative judgement as a promising alternative to score competences. In E.

Cano, & G. Ion (Eds.), *Innovative practices for higher education assessment and measurement* (pp. 119-138). Pennsylvania: IGI Global.

Lingua Latina. Retrieved from https://lingualatina.dk/wp/.

Luger, S. (2015). Latijn vertalen in de bovenbouw: een mission impossible? [Translating Latin in upper secondary education: a mission impossible?]. *Lampas*, *48*, 212–225.

Luger, S. (2016). Vertalingen beoordelen in de bovenbouw: Een cola-tic of een glijdende schaal? [Assessing translation in upper secondary education: A Coca Cola with a twist or a sliding scale?]. *Lampas, 49*, 3-19.

Luger, S. (2018). How do Dutch adolescents translate Latin into coherent Dutch? A Journey into the Unknown. *Journal of Latin Linguistics, 17(2)*, 333–365. doi: https://doi.org/10.1515/joll-2018-0015.

Merrill, M.D. (2002). First principles of instruction. *Educational Technology Research and Development, 50*(3), 43-59. https://doi.org/10.1007/BF02505024.

Messick, S. (1994). Alternative modes of assessment, uniform standards of validity. *ETS Research Report Series*, *1994*(2), 1-22. doi: https://doi.org/10.1002/j.2333-8504.1994.tb01634.x.

Nord, C. (1997). *Translating as a purposeful activity. Functionalist approaches explained.* Manchester: St Jerome.

Nvivo.nl. Retrieved from https://www.nvivo.nl/.

O'Brien, S. (2006). Eye-tracking and translation memory matches. *Perspectives-Studies In Translatology*, *14*(3), 185-205.
doi: http://dx.doi.org/10.1080/09076760708669037.

O'Brien, S. (2008). Processing fuzzy matches in translation memory tools: An eye-tracking analysis. In S. Göpferich, A. Jakobsen, & I. Mees (Eds.) *Looking at eyes— eye tracking studies of reading and translation processing* (pp. 79-102). *Copenhagen studies in language 36.* Frederiksberg: Samfundslitteratur.

O'Donnell, C.L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, *78*(1), 33-84.

Olive, T. (2010). Methods, techniques, and tools for the on-line study of the writing process. In N.L. Mertens (Ed.), *Writing: Processes, tools and techniques* (pp. 1-18). Hauppauge, New York: Nova Science Publishers.

Olive, T., Kellogg, R. T., & Piolat, A. (2002). The triple task technique for studying the process of writing. In: *Contemporary tools and techniques for studying writing* (pp. 31-59). Springer, Dordrecht. doi: https://doi.org/10.1007/978-94-010-0468-8_3.

Online Woordenboek Latijn – Nederlands [Online Latin – Dutch Dictionary]. Retrieved from http://www.latijnnederlands.nl/.

Perseus under Philologic, 2018 edition. Retrieved from http://perseus.uchicago.edu/.

Pinkster, H. (2018). *Woordenboek Latijn–Nederlands* [Latin–Dutch Dictionary]. Amsterdam: Amsterdam University Press.

Pollmann, E., Prenger, J., & Glopper, K. (2012). Het beoordelen van leerlingteksten met behulp van een schaalmodel [Grading students' texts using a scale model]. *Levende Talen Tijdschrift*, *13*(3), 15-24.

Pym, A. (2003). Redefining translation competence in an electronic age. In defence of a minimalist approach. *Meta: journal des traducteurs / Meta: Translators' Journal*, *48*(4), 481-497. doi: https://doi.org/10.7202/008533ar.

Raap, J. (1992). Vertalen op school: een ramp? [Translation at school: a disaster?]. *Lampas, 25*, 149-151.

Reiss, K., & Vermeer, H.J. (1984). *Grundlegung einer allgemeinen Translationstheorie* [Founding of a general theory of translation]. Tübingen: Niemeyer.

Rijlaarsdam, G., Couzijn, M., & Van den Bergh, H. (2004). The study of revision as a writing process and as a learning-to-write process. In *Revision Cognitive and Instructional Processes* (pp. 189-207). Springer, Dordrecht.

Rijlaarsdam, G., Van den Bergh, H., Couzijn, M., Janssen, T., Braaksma, M., Tillema, M., … Raedts, M. (2011). Writing. In K.R. Harris & S. Graham (Eds.), *APA educational psychology handbook: Volume 3*, (pp. 189 -227). Washington, D.C.: American Psychological Association.

Robert, I. S., & Van Waes, L. (2014). Selecting a translation revision procedure: do common sense and statistics agree? *Perspectives: Studies in Translatology*, *22*(3), 304-320. doi: http://dx.doi.org/10.1080/0907676X.2013.871047.

Schrijver, I. (2014). *The Translator as a text producer, the effects of writing training on transediting and translation performance.* (Doctoral Dissertation). Universiteit van Antwerpen, Antwerp.

Schrijver, I., Van Vaerenbergh, L., Leijten, M., & Van Waes, L. (2018). De vertaler als schrijver: Tekstproductiecompetentie als onderdeel van vertaalcompetentie [The translator as a writer: Text production competence as part of translation competence]. *Tijdschrift voor Taalbeheersing, 40*(3), 269-301. doi: 10.5117/TVT2018.3.001.SCHR.

Schunk, D.H. (1995, March–April). *Social origins of self-regulatory competence: The role of observational learning through peer modeling*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Indianapolis, IN.

Sicking, C.M.J. (1968). Tweede interimrapport commissie modernisering leerplan klassieke talen [Second interim report of the committee to modernise the curriculum for classical languages]. *Lampas, 1*, 166-176.

Stichting Leerplanontwikkeling. (2015). *Curriculumspiegel deel A: Generieke trendanalyse* [National curriculum overview part a: generic trend analysis]. Enschede: Stichting Leerplanontwikkeling.

Stichting Leerplanontwikkeling. (2015). *Curriculumspiegel deel B: Vakspecifieke trendanalyse* [National curriculum overview part b: subject-specific trend analysis]. Enschede: Stichting Leerplanontwikkeling.

Stilus dynalyse uitleg [Stilus explanation of dynalysis]. Retrieved from http://www.stilus.nl/dynalyse/uitleg.htm.

Stilus woord vooraf bij de examens Grieks en Latijn [Stilus introduction to the Latin and Greek exams]. Retrieved from http://www.stilus.nl/examen/#10.

Van Amerongen R., Budé, A., Derix, H., Van Gessel, H., Goossens, T., Mak, … Zwart, J. (1986). Leerplan Latijn [Curriculum for Latin]. *VCN Bulletin, 31*, 1-39.

Van Daal, T., Lesterhuis, M., Coertjens, L., Donche, V., & De Maeyer, S. (2019). Validity of comparative judgement to assess academic writing: examining

implications of its holistic character and building on a shared consensus. *Assessment in Education: Principles, Policy & Practice*, *26*(1), 59-74.

Van den Bergh, H., Rijlaarsdam, G., & Breetvelt, I. (1994). Revision process and text quality: An empirical study. In G. Eigler & Th. Jechle, (Eds.), *Writing: Current Trends in European Research* (pp. 133-148). Freiburg: Hochschul Verlag.

Van de Walle, E., & Van Houdt, T. (2004). Bouwstenen voor een vernieuwde leesdidactiek: over de plaats en functie van leesstrategieën in het leesproces [Building blocks for renewed reading didactics: about the place and function of reading strategies in the reading process]. *Kleio: Tijdschrift voor Oude Talen en Antieke Cultuur, 33*(2), 72-86.

Van Houdt, T. (2008). The strategic reading of Latin (and Greek) texts: a research-based approach. In R. Lister (Ed.), *Meeting the challenge: international perspectives on the teaching of Latin* (pp. 54-70). Cambridge: Cambridge University Press.

Van Krieken, R. (1981). De klassieke vertaling en het actueel dilemma: taalkennis of tekstbegrip [Classical Translation and the current dilemma: language skills or text comprehension]. *Levende Talen, 363*, 558-577.

Van Krieken, R. (1982). Vertalen en begrijpen [Translating and understanding]. *Toegepaste taalwetenschap in artikelen, 13*, 128-147.

Van Oeveren, C.D.P. (2019). *Ithaka gaf je de reis* [Ithaka gave you the journey]. (Doctoral Dissertation). Vrije Universiteit, Amsterdam.

Verbaal, W. (2015). Leren lezen als Romeinen: niet enkel een kwestie van syntaxis. [Learning to read like Romans: Not just a matter of syntax]. *Taal in tekst*, 7-38.

Verhavert, S., De Maeyer, S., Donche, V., & Coertjens, L. (2018). Scale separation reliability: what does it mean in the context of comparative judgment? *Applied Psychological Measurement*, *42*(6), 428-445.

Verhoeven, P. (1996). *Tekstbegrip in het onderwijs klassieke talen.* (Doctoral Dissertation). Universiteit Leiden, Leiden.

Vermeer, H.J., (1990). Skopos und Translationsauftrag - Aufsätze [Skopos and translation assignment – essays]. Heidelberg: Universitätsdruckerei.

Wesdorp, H. (1981). *Evaluatietechnieken voor het moedertaalonderwijs: een inventarisatie van beoordelingsmethoden voor de stelvaardigheid, het begrijpend lezen, de spreek-, luister- en discussievaardigheid* [Evaluation techniques in mother-tongue education: an inventory of assessment methods for essaywriting skills reading comprehension, speaking skills, listening skills and discussion skills]. 's-Gravenhage: Stichting voor Onderzoek van het Onderwijs.

AUTHOR INDEX

SUMMARY
LOST IN LATIN TRANSLATION

Teaching Students to Produce Coherent Target Texts

The problem this dissertation addressed was that Dutch students in upper secondary education generally struggle to show their comprehension of a Latin source text when they translate it into Dutch, although this is a requirement formulated by the programme for the Dutch central final examinations (domain A, sub-domain 1). Their struggle is demonstrated by the incoherent target texts students produce when they are translating Latin. The research question this dissertation addressed is how to improve the target-text coherence in texts upper secondary students produce when translating Latin.

In this dissertation I identified four problems in Dutch upper secondary education related to translating Latin:

1. The translation assignment itself. The *skopos* of the translation is unclear and the translation assignment for Latin does not represent a real-world problem.
2. Learning activities for producing a coherent as well as an equivalent target text are not defined. Teachers do not really know *what* to teach their students to make them translate Latin into a coherent target text.
3. A base of shared knowledge for teaching Latin translation is missing Teachers do not really know *how* to teach their students to translate Latin into a coherent target text.
4. The rating method used in Dutch upper secondary education provides insufficient assessment of the coherence of targets texts.

Before an experimental study aimed at improving target-text coherence in students translations could be designed and performed, these problems had to be addressed. The first problem was addressed in chapter one, where the production of a coherent target text understandable for a person unfamiliar with the source text was formulated as translation assignment.

Problem two was addressed in chapters three and four: through previous research and the performance of an eye-tracking study with proficient students, translation activities leading to coherent target-text production could be identified. These activities formed the basis for the newly developed process-oriented translation strategy presented in chapter four (*Figure 12*). Thus, learning activities for producing a coherent target text were defined.

The third problem was solved in chapter four, where Self-Regulated Strategy Development (Graham & Harris 1996, 2002), an effective instruction method used in the teaching of writing, was adapted to teaching the process-oriented translation strategy.

The fourth problem was addressed in chapter two, where holistic assessment was successfully tested for reliably rating target-text coherence.

Having solved these four problems, I was able to develop lessons, as described in chapter four, to address my initial research question: how to improve the target-text coherence in texts upper secondary students produce when translating Latin. I designed an experimental study for two learning conditions:

1. an experimental condition, teaching the process-oriented translation strategy, following SRSD as instruction method,
2. a control condition, teaching Latin using error analysis, following 'business as usual' as instruction method.

The testing of these lessons brought to light that a control condition featuring 'business as usual' was problematic, as this business is not well defined which caused issues with the stability of the control condition, where ʻbusiness as usualʼ did not provide sufficient direction for the lessons (Chapter 5). Therefore, I redesigned the experimental study (Chapter 5), which resulted in an experiment consisting of two competing experimental conditions:

1. a *process condition* focusing on the development of text production and revision skills as well as the development of metalanguage, with instruction characterised as strategy instruction, and
2. a *product condition* focusing on error analysis and source text-target text equivalence, where the instruction method could still be characterised as *business as usual.*

The results of the experiment (Chapter 6) showed that after an intervention of four lessons (eight hours in total), participants significantly improved both target–text coherence and ST-TT equivalence in both conditions. The *process condition* did not have a differential effect on Coherence. A surprising effect was that ST-TT equivalence improved significantly in the *process condition* as well, although no specific instruction was given in the lessons either on equivalence of translations or on the source language (Latin). It also became clear that writing coherent target texts is not a result of proficiency in Dutch (the target language), as no significant correlation between Dutch grade and coherence scores was found in the tests.

This dissertation has established that the process-oriented translation strategy offers a new and effective approach to teaching translation. The correlation between Latin proficiency and pre-test scores confirms that translation always also relates to source-language proficiency. Therefore, the attention to

the Latin source text should not be neglected, as supported by the positive effects on coherence and equivalence in the *product condition*. However, to teach translation and to translate are separate activities and a clear distinction between them could significantly improve the practice of Latin translation in the Dutch classroom and thereby the coherence of the target texts students write.

SAMENVATTING
VERTAALD OF VERDWAALD

Leerlingen leren coherente doelteksten te produceren bij het vertalen van Latijn

Veel leerlingen die Latijn vertalen in de bovenbouw van het gymnasium produceren incoherente doelteksten. Dat is niet alleen een probleem omdat het de docenten droevig stemt (Kroon & Sluiter, 2010), maar vooral ook omdat volgens de eindtermen de kandidaat door een onbekende Latijnse tekst in het Nederlands te vertalen begrip van die tekst moet demonstreren (domein A, subdomein 1) en dat lukt vaker niet dan wel. De onvrede over deze situatie leidde tot de onderzoeksvraag van dit proefschrift: hoe kan de coherentie van de doelteksten die leerlingen in de bovenbouw van het gymnasium produceren als ze Latijn vertalen verbeterd worden?

In het eerste hoofdstuk heb ik vier problemen vastgesteld in relatie tot het vertalen van Latijn in het Nederlandse gymnasiumonderwijs:

1. De vertaalopdracht die de leerling krijgt. De *skopos* van de vertaling is niet duidelijk doordat de vertaalopdracht geen authentiek probleem is. De leerling vertaalt immers voor een docent die de brontekst veel beter kan lezen, begrijpen en vertalen dan de leerling zelf.
2. Het is niet vastgesteld welke leeractiviteiten leiden tot de productie van een coherente (of equivalente) vertaling. Docenten weten dus niet goed *wat* ze moeten onderwijzen om leerlingen een coherente doeltekst te laten produceren.
3. Er is geen algemeen aanvaarde kennisbasis voor een methode om Latijn te leren vertalen. Docenten weten dus niet goed *hoe* ze hun leerlingen kunnen onderwijzen een coherente doeltekst te produceren.
4. De beoordelingsmethode voor vertalingen die in de bovenbouw gangbaar is (colon-beoordeling) geeft onvoldoende ruimte voor het beoordelen van doeltekstcoherentie.

Deze problemen moesten worden aangepakt voordat het mogelijk was een experimentele studie te ontwerpen en uit te voeren om een antwoord te vinden op de hoofdvraag hoe de doeltekstcoherentie bij leerlingen kan worden

verbeterd. Als oplossing voor het eerste probleem definieer ik in het eerste hoofdstuk een nieuwe vertaalopdracht. De vertaalopdracht moet luiden dat de leerling een coherente doeltekst produceert die begrijpelijk is voor iemand die de brontekst niet kent.

Het tweede probleem, de vertaalactiviteiten die leiden tot coherente doeltekstproductie is het onderwerp van de hoofdstukken drie en vier. Ik heb allereerst in hoofdstuk drie de vertaalactiviteiten vastgesteld die bijdragen aan de productie van een coherente doeltekst op basis van eerder onderzoek (Breedveld, 2002, Göpferich, 2008) en door middel van een eye-tracking studie die ik heb uitgevoerd bij succesvolle leerlingvertalers (hoofdstuk 3). Deze vertaalactiviteiten per fase van het vertaalproces vormen de basis voor de procesgerichte vertaalstrategie die ik presenteer in hoofdstuk vier (*Figure 12*).

In het eerste hoofdstuk heb ik beargumenteerd dat doeltekstproductie en tekstschrijven verwant zijn (Schrijver, 2014). Voor schrijfdidactiek bestaat een geteste en effectief gebleken instructiemethode: de Self-Regulated Strategy Development (Graham & Harris 1996, 2002). Het ontbreken van een betrouwbare methode om coherente doeltekstproductie te onderwijzen, heb ik aangepakt door deze methode voor schrijfdidactiek aan te passen voor vertaaldidactiek (hoofdstuk 4).

De beoordelingsmethode was het vierde en laatste probleem en komt in hoofdstuk twee aan de orde. Ik heb een holistische beoordelingsmethode getest en vastgesteld dat dit een betrouwbaar meetinstrument is voor tekstcoherentie.

Toen deze vier kwesties succesvol waren opgelost kon ik overgaan op het ontwerpen van de lessen om de procesgerichte vertaalstrategie te onderwijzen (hoofdstuk vier) en heb ik een experimentele studie ontworpen voor twee condities:

1. een experimentele conditie, waarin de procesgerichte vertaalstrategie wordt onderwezen volgens de principes van SRSD,
2. een controleconditie waarin het vertalen van Latijn wordt onderwezen met behulp van foutenanalyse volgens de principes van 'de gebruikelijke praktijk.'

Bij het proefdraaien van deze lessen openbaarde zich een nieuw probleem op het gebied van de instructiemethode 'de gebruikelijke praktijk' (hoofdstuk 5). De twee docenten in de controleconditie verschilden te zeer in hun aanpak om van een stabiele controleconditie te kunnen spreken. Met andere woorden, de praktijk bleek nog minder duidelijk gedefinieerd te zijn dat verwacht. Daarom heb ik de opzet van de experimentele studie aangepast en herontworpen als een experiment met twee concurrerende experimentele condities:

1. een *procesconditie* die zich richt op de ontwikkeling van tekstproductie- en revisievaardigheden en de ontwikkeling van metataal, met strategieinstructie als instructiemethode, en
2. een *productconditie* gericht op foutenanalyse en brontekst-doeltekst equivalentie, met een instructiemethode die nog steeds geldt als *de gebruikelijke praktijk* vanwege de focus op brontekst en equivalentie.

De resultaten van dit experiment (hoofdstuk 6) tonen aan dat na een interventie van vier lessen (een totaal van acht uur les) de deelnemers zowel doeltekst-coherentie als brontekst-doeltekst equivalentie in beide condities significant hebben verbeterd. De *procesconditie* had geen differentieel effect op Coherentie. Verrassend was dat de brontekst-doeltekst equivalentie ook in de *procesconditie* significant verbeterde, terwijl in die lessen geen gerichte aandacht was voor de equivalentie van vertalingen of voor Latijn. Voor de suggestie dat de productie van een coherente doeltekst vooral het resultaat is van taalvaardigheid in het Nederlands is geen aanwijzing gevonden: tussen het cijfer dat deelnemers voor Nederlands behaalden en hun coherentiescores is geen correlatie vastgesteld.

Met dit proefschrift het ik aangetoond dat de procesgerichte vertaalstrategie een nieuwe en effectieve benadering biedt om leerlingen te leren vertalen. De correlatie tussen het zelf gerapporteerde cijfer voor Latijn op school en de scores in de voormeting bevestigt dat er inderdaad een relatie bestaat tussen vertalen en kennis van de brontaal. Gezien de positieve effecten van de *productconditie* op zowel equivalentie als coherentie houd ik ook geen pleidooi voor het verwaarlozen van de aandacht voor de Latijnse brontekst. Maar wel wil ik benadrukken dat het *leren* en het *vertalen* van Latijn twee aparte vaardigheden zijn en dat een duidelijk onderscheid daartussen de praktijk van het Latijnse vertaalonderwijs aanzienlijk kan verbeteren. Met als resultaat een verbetering van de coherentie en dus de leesbaarheid van de doelteksten die de leerlingen schrijven.

APPENDIX A - CHAPTER 5

Table LV: Assessors: Subject Taught and Number of Comparisons

Assessor	Subject	Number of comparisons
1	Dutch	1
2	Dutch	1
3	Dutch	77
4	Dutch	6
5	English	73
6	History	73
7	History	73
8	Latin	73
9	Latin	73
10	Latin	73
11	Latin	30
12	Latin	73
13	Latin	73
14	Latin	76
15	Latin	76
16	Latin	39
17	Latin	73
18	Philosophy	73

Table LVI: Distribution of Lessons over Classes in Experimental Condition.

	Group A- classes	Group B - classes
Lesson 1 / Pre-test	Thursday 01/06/2017 8.30-9.40 Thursday 01/06/17 14.15-15.25	Monday 12/06/2017 9.45-10.55
Lesson 2	Wednesday 07/06/17, 12.40-13.50 Thursday 08/06/17 8.30-9.40: part 1	Wednesday 14/06/17 8.30-9.40 Friday 16/06/17 12.40-13.50: part 1
Lesson 3	Thursday 08/06/17 8.30-9.40: part 2 Thursday 08/06/17 14.15-15.25 Wednesday 14/06/17 12.40-13.50	Friday 16/06/17 12.40-13.50: part 2 Monday 19/06/17 9.45-10.55)
Lesson 4	Thursday 15/06/17 8.30-9.40 Thursday 15/06/17 14.15-15.25	Wednesday 21/06/17 8.30-9.40
Lesson 5	Wednesday 21/06/17 12.40-13.50 Thursday 22/06/17 8.30-9.40)	Friday 23/06/17 12.40-13.50
Lesson 6 / post-test and learner report	Thursday 22/06/17 14.15-15.25	Monday 26/06/17 9.45-10.55

Note: All Classes Highlighted in Grey Were Observed by the Researcher, All Classes in Non-Highlighted B Were Taught by the Researcher.

.

Table LVII: Distribution of Lessons over Classes in Control Condition.

	Group C - classes	Group D -classes
Lesson 1 / pre-test	Thursday 08/06/2017 9.45-10.55	Tuesday 06/06/2017 12.10-13.10
Lesson 2	Friday 09/06/17 14.50-15.25	Wednesday 07/06/17 11.25-12.35
Lesson 3	Tuesday 13/06/17 09.35-10.35	Friday 09/06/17 14.50-15.25
Lesson 4	Thursday 15/06/17 09.45-10.55	Tuesday 13/06/17 12.10-13.10
Lesson 5	Friday 16/06/17 5th period (14.50-15.25) Tuesday 20/06/17, 2nd period (9.35-10.35)	Wednesday 14/06/17 11.25-12.35
Lesson 6 / post-test and learner report	Thursday 22/06/17, 2nd period (09.45-10.55)	Friday 16/06/17 14.50-15.25

Note: All Classes Highlighted in Grey Were Observed by the Researcher

APPENDIX B - CHAPTER 5

The components of the lessons we designed and redesigned after testing are shown in the following tables. *Table LVIII* shows the redesign of the common features in both conditions (right) compared to the original design (left). *Table LIX* compares specific features of the designed (left) and redesigned lessons (right) in the experimental / *process condition* and

Table LX compares specific features of the designed (left) and redesigned (right) lessons in the control / *product condition. Table LXI* compares the features of the two conditions, *process* and *product*, in the redesigned experiment.

Table LVIII: Designed and Redesigned Lessons: Common Features [228]

	Design	Redesign
Session 1 Pre-test	Translation of fable Ovis, cervus et lupus or De vulpe et uva on computer (I) Translation of fable Vulpes ad personam tragicam (I) Evaluation (I)	Translation of fable Canis per fluvium carnem ferens on computer (I) Translation of fable Ovis, cervus et lupus on computer (I) break Translation of fable Vulpes ad personam tragicam (I) Evaluation (I)
Session 2	Assignment on types of translation: equivalent, literary, communicative (fable Vacca et capella, ovis et leo) (P) Peer-comparison (P) Feedback (P) Working in pairs / individually (P/I)	Assignment on types of translation: equivalent, literary, communicative (fable Vacca et capella, ovis et leo) (P) Introduction on fable as a genre (moral, situation, action) (C) Peer-comparison (P) Feedback (P) Classroom discussion (C) Working in pairs / individually (P/I)
Session 3	Reflection on translation habits (I) Working individually (I)	Reflection on translation habits (I)

[228] *(I) = Individual; (P) = Pair work; (C) = whole class; (G) = group work*

Table LVIII: Designed and Redesigned Lessons: Common Features [228]

	Carrousel assignment (G/I)	Peer comparison (P) Classroom discussion (C) Working in pairs / individually (P/I) Evaluation 1 in Socrative (I)
Session 4	Reflection on translation habits (I) Classroom discussion (C) Working individually (I)	Reflection on translation habits (I) Peer comparison (P) Classroom discussion (C) Working in pairs / individually (P/I)
Session 5	Reflection on translation habits (I) Working individually and in pairs (I/P)	Reflection on translation habits (I) Translation of Lupus et Agnus (I) Evaluation 2 in Socrative (I)
Session 6 Post-test	Translation of fable De vulpe et uva or Ovis, cervus et lupus on computer (I) Learner report	Translation of fable Musca et Mula on computer (I) Translation of fable Canes Famelici on computer (I) Evaluation 3 in Socrative (I)
Session 7 De-layed test	-	Translation of fable Canes et Corcodilli on computer (I) Translation of fable De vulpe et uva on computer (I) Evaluation 4 in Socrative (I)

Table LIX: Both Conditions: Specific Features [229]

	Designed / Experimental condition	Redesigned / Process condition
Session 1 Pre-test	1. Scoring the translation-activities of Vulpes ad personam tragicam in 30s intervals (I)	same
Session 2	1. Analysis of interval scores (I) 2. Comparing own translation process with translation process of professional translators (I) 3. Comparing interval scores, giving and receiving feedback (P) 4. Participating in classroom discussion on translation process (C) 5.1 Listening to introduction of the concept 'text coherence' (C) 5.2. Assessing text coherence of 4 translations of Vulpes ad personam tragicam (P) 5.3 Ranking the translation on coherence (G) 5.4 Discussion of the ranking (C)	1 – 4 same 5. Listening to introduction of the concept 'text coherence' (C) 6. Analysis of Mustela et homo / Asinus ad senem pastorem: 6.1 Marking moral, situation, action in Dutch translation (I). 6.2 Marking linguistic characteristics in Latin text (I). 6.3 Drawing assignment (I) 7. Peer comparison of marking and drawing (P) 7. Participating in classroom discussion on text coherence (C)
Session 3	1. Text comprehension Carrousel (G) 1.1 Listening to introduction on genre characteristics: three building blocks (moral, situation, action) (C) 1.2 Marking moral, situation, action in Dutch translation of fable Asinus ad senem	1. Observation of text revision (screencast) (I) 2. Formulating feedback on revision activities in screencast (P) 3. Participating in classroom discussion on text revision (C) 4. Assessing text coherence of 4 translations of Vulpes ad personam tragicam (P)

[229] *(I) = Individual; (P) = Pair work; (C) = whole class; (G) = group work*

Table LIX: Both Conditions: Specific Features [229]

	pastorem, or Mustela et homo, or Asinus et leo venantes (I) (I). 1.3 Drawing a cartoon in three pictures of the translation of the same fable as 1.2 (I) 1.4 Peer comparison of marking and drawing (G) 1.5 Formulating expectations of text coherence (G). 2.1 Observation of text revision (screencast) (I) 2. Formulating feedback on revision activities in screencast (P) 3. Participating in classroom discussion on text revision (C) 4.1 Revision of least coherent translation (P) 4.2 Comparing revised texts (P) 4.3 Peer feedback on revised text (P) 5. Participating in classroom discussion on revision (C) 6. Online evaluation in Socrative (I)	4.1 Revision of least coherent translation (P) 4.2 Comparing revised texts (P) 4.3 Peer feedback on revised text (P) 5. Participating in classroom discussion on revision (C)
Session 4	1.Observation of text production (screencast) (I) 2. Formulating feedback on text production activities in screencast (P) 3. Participating in classroom discussion on text production. (C) 4. Applying knowledge of the full translation process (guided: marking building blocks, drawing, comparing, writing chunks, writing text, revising) translating one fable (Mulier parturiens / Canes	1.Observation of text production (screencast) (I) 2. Formulating feedback on text production activities in screencast (P) 3. Participating in classroom discussion on text production. (C) 4. Applying knowledge of the full translation process (guided) translating one fable Mulier parturiens / Duo calvi (I & P) 4.1 Revision of one fable Mulier parturiens / Duo calvi (I & P)

Table LIX: Both Conditions: Specific Features [229]

	Famelici / Canes et Corcodilli / Duo calvi (I & P) 5. Participating in classroom discussion of both fables (C) 6. Listening to explanation of linguistic characteristics in Latin text related to the three building blocks. (C)	5. Participating in classroom discussion of both fables (C)
Session 5	1.2 Marking moral, situation, action in fable Muli Duo et Latrones (with translation given) (I) 1.2 Marking linguistic characteristics in Latin text (I). 2. Applying the knowledge of the full translation process (with a reminder) translating Lupus et Agnus (I) 3. Using a (projected) slide as a reminder of phases in translation process (when needed) (I) 4. Peer feedback on revised text (P)	1. Applying the knowledge of the full translation process (with a reminder) translating fable Lupus et Agnus (I) 2. Using written reminder of phases in translation process (when needed) (I) 3. Participating in classroom discussion on the translation process and on the moral (C)

Table LX: The Control / Product Condition; Specific Features[230]

	Designed / Control condition	Redesigned / Product condition
Session 1	pre-test	pre-test
Session 2	1. Checking translation fable Vulpes ad personam tragicam (I)	1-4 same
	2. Error analysis of translation of fable Vulpes ad personam tragicam (I)	5. Participating in classroom discussion on moral, situation, action in Vulpes ad personam tragicam and on genre characteristics in general (C)
	3. Comparing error analysis, giving and receiving feedback (P)	
	4. Participating in classroom discussion on commonly made errors, discussion of relevant grammar (C)	
	5. Comparing results of assignment on types of translation (G)	
Session 3	1.Translation Carrousel (G)	1. Translation of Asinus ad senem pastorem (I)
	1.1 Translation of fable Asinus ad senem pastorem / Mustela et homo / Asinus et leo venantes (I)	1.1 Comparison of translation product with a peer and discussing differences (P)
	1.2 Checking translation of another fable by a peer who is not familiar with the source text and making error analysis of that translation (I)	2. Answering written questions on text comprehension (P)
	1.3 Drawing a cartoon of the third fable in three pictures based on the translation by another peer of the translation (I)	3. Participating in classroom discussion reviewing the translation and the written questions, discussion of relevant grammar (C)
	1.4 Discussion of the errors that were made in the three fables'	4. Translation of Mustela et homo (I)
		4.1 Checking translation with model (I)

---

[230] *(I) = Individual; (P) = Pair work; (C) = whole class; (G) = group work*

Table LX: The Control / Product Condition; Specific Features[230]

	translations: formulation of tips and tricks (G)	4.2 Making an error analysis (I) 5. Peer comparison of errors: formulation of tips and tricks (G) 6. Presenting tips and tricks to the whole class (G)
Session 4	1.1 Translation of fable Lupus et Agnus (I) 1.2 reviewing the translation and discussing moral & relevant grammar (C) 1.3 Making an error analysis (I) 2.1 Translation of fable Duo calvi (I) 2.2 Checking own translation with a model (I) 2.3 Making error analysis (I)	1. Translation of Duo calvi (I) 1.1 Checking translation with model (I) 1.2 Making an error analysis (I) 1.3 peer comparison & peer feedback on error analysis 2. Formulating translation tips and tricks (C) 2.1 Participating in classroom discussion classifying tips and tricks. 3. Translation of Mulier parturiens (according to preference: I / P) 4. Participating in classroom discussion reviewing the translation and discussing relevant grammar (C) 5. Answering written questions on text comprehension (I)
Session 5	1. Translation of fable Mulier parturiens / Canes Famelici / Canes et Corcodilli (P) 1.1 Participating in classroom discussion reviewing the translation and discussing moral & relevant grammar (C) 1.2 Making an error analysis of the translation of one of the three translated and discussed fables (I)	1. Participating in classroom discussion reviewing the translation and discussing relevant grammar (C) 2. Drawing assignment OR error analysis (according to preference) (I)

Table LXI: Both Conditions; Common and Specific Features [231]

	Process condition	Product condition
Pre-test Common	Translation of fable 1 on computer (I) Translation of fable 2 on computer (I) break Translation of fable Vulpes ad personam tragicam (I) Evaluation (I)	
Specific	1. Scoring the translation-activities of Vulpes ad personam tragicam in 30s intervals (I)	
Session 2 Common	Assignment on types of translation: equivalent, literary, communicative (Vacca et capella, ovis et leo) (P) Introduction on fable as a genre (moral, situation, action) (C) Peer-comparison (P) Feedback (P) Classroom discussion (C) Working in pairs / individually (P/I)	
Specific	1. Analysis of interval scores (I) 2. Comparing own translation process with translation process of professional translators (I) 3. Comparing interval scores, giving and receiving feedback (P) 4. Participating in classroom discussion on translation process (C) 5. Listening to introduction of the concept 'text coherence' (C) 6. Analysis of Mustela et homo / Asinus ad senem pastorem: 6.1 Marking moral, situation, action in Dutch translation (I). 6.2 Marking linguistic characteristics in Latin text (I). 6.3 Drawing assignment (I)	1. Checking translation of Vulpes ad personam tragicam (I) 2. Error analysis of translation of Vulpes ad personam tragicam (I) 3. Comparing error analysis, giving and receiving feedback (P) 4. Participating in classroom discussion on commonly made errors, discussion of relevant grammar (C) 5. Participating in classroom discussion on moral, situation, action in Vulpes ad personam tragicam and on genre characteristics in general (C)

[231] *(I) = Individual; (P) = Pair work; (C) = whole class; (G) = group work*

	7. Peer comparison of marking and drawing (P)	
	8. Participating in classroom discussion on text coherence (C)	
Session 3 Common	Reflection on translation habits (I) Peer comparison (P) Classroom discussion (C) Working in pairs / individually (P/I) Evaluation 1 in Socrative (I)	
Specific	1. Observation of text revision (screencast) (I) 2. Formulating feedback on revisionactivities in screencast (P) 3. Participating in classroom discussion on text revision (C) 4. Assessing text coherence of 4 translations of Vulpes ad personam tragicam (P) 4.1 Revision of least coherent translation (P) 4.2 Comparing revised texts (P) 4.3 Peer feedback on revised text (P) 5. Participating in classroom discussion on revision (C)	1. Translation of Asinus ad senem pastorem (I) 1.1 Comparison of translation product with a peer and discussing differences (P) 2. Answering written questions on text comprehension (P) 3. Participating in classroom discussion reviewing the translation and the written questions, discussion of relevant grammar (C) 4. Translation of Mustela et homo (I) 4.1 Checking translation with model (I) 4.2 Making an error analysis (I) 5. Peer comparison of errors: formulation of tips and tricks (G) 6. Presenting tips and tricks to the whole class (G)
Session 4 Common	Reflection on translation habits (I) Peer comparison (P) Classroom discussion (C) Working in pairs / individually (P/I)	
Specific	1.Observation of text production (screencast) (I) 2. Formulating feedback on text production activities in screencast (P) 3. Participating in classroom discussion on text production (C) 4. Applying knowledge of the full translation process (guided) translating one fable Mulier parturiens / Duo calvi (I & P) 4.1 Revision of one fable Mulier parturiens / Duo calvi (I & P)	1. Translation of Duo calvi (I) 1.1 Checking translation with model (I) 1.2 Making an error analysis (I) 1.3 peer comparison & peer feedback on error analysis 2. Formulating translation tips and tricks (C) 2.1 Participating in classroom discussion classifying tips and tricks. 3. Translation of Mulier parturiens (according to preference I / P)

	5. Participating in classroom discussion of both fables (C)	4. Participating in classroom discussion reviewing the translation and discussing relevant grammar (C) 5. Answering written questions on text comprehension (I)
Session 5 Common	Reflection on translation habits (I) Translation of Lupus et Agnus (I) Evaluation 2 in Socrative (I)	
Specific	1. Applying the knowledge of the full translation process (with a reminder) (I) 2. Using written reminder of phases in translation process (when needed) (I) 3. Participating in classroom discussion on the translation process and on the moral (C)	1. Participating in classroom discussion reviewing the translation and discussing relevant grammar (C) 2. Drawing assignment OR error analysis (according to preference) (I)
Post-test Common	Translation of fable Musca et Mula on computer (I) Translation of fable Canes Famelici on computer (I) Evaluation 3 in Socrative (I)	
Delayed post-test Common	Translation of fable Canes et Corcodilli on computer (I) Translation of fable De vulpe et uva on computer (I) Evaluation 4 in Socrative (I)	

APPENDIX C - CHAPTER 6

Table LXII: A-D Data Collection Specified for Each Group

VIII-A: Process Condition: Group A (Monday): n = 28 initial/20 final

Ses-sion	Date	Pre-sent	Catch up	Data	Particulars
1	6/11	20	Thu. 9/11 (4) Fri. 10/11 (1)	Translations 1&2 Screencasts 1&2 Interval scores Written evaluation	Due to technical problems not all participants had time to complete the 2nd translation. 3 participants no show (28-3 = 25) One participant moved to group C and had to stay there for the experiment. 1 to group C (25-1 = 24) 2 changed A > B (24-2 = 22)
2	13/11	20	Thu 16/11(1) Fri 17/11(1)	Lesson book	
3	20/11	-	-	-	Postponed

217

APPENDIX C

3	27/11	19	Thu 30/11(1) Fri 1/12 (2)	Lesson book Evaluation in Socrative	1 changed A>B (22-1 = 21) 1 participant changed B>A (21+1 = 22)
4	04/12	17	Thu 30/11(1) Fri 8/12 (2)	Lesson book	1 drop out (22-1 = 21) 1 drop out (21-1 = 20)
5	11/12	-	-	-	Cancelled due to snowstorm
5_1	13/12	11	-	Lesson book Evaluation in Socrative	Catch up session (Wednesday)
5_2	15/12	9	-	Lesson book Evaluation in Socrative	Catch up session (Friday)
6	18/12	18	Thu 21/12(2)	Translations 3&4 Screencasts 3&4 Evaluation in Socrative	
7	15/01	19	Thu 18/1 (1)	Translations 5&6 Screencasts 5&6 Evaluation in Socrative	

VIII-B: Process Condition

Group B (Thursday): n= 23 initial/22 final

Session	Date	Present	Catch up session	Data	Particulars
1	9/11	20	-	Translations 1&2 Screencasts 1&2 Interval scores Written evaluation	3 participants no show (23-3 = 20)
2	16/11	20	20/11 (2)	Lesson book	2 participants changed A > B (20+2 = 22)
3	23/11	18	Mon 27/11(3) Thu 30/11(1)	Lesson book Evaluation in Socrative	1 participant changed B>A (22-1 = 21) 1 participant changed A>B (21+1 = 22)
4	30/11	22		Lesson book	
5	7/12	20	Fri 8/12 (1) Wed 13/12 (1)	Lesson book Evaluation in Socrative	
6	14/12	20	Mon 18/12 (2)	Translations 3&4 Screencasts 3&4 Evaluation in Socrative	
7	18/01	22		Translations 5&6 Screencasts 5&6 Evaluation in Socrative	

APPENDIX C

VIII-C: Product Condition

Group C (Monday): n=29 initial, 24 final

Session	Date	Present	Catch up session	Data	Particulars
1	6/11	23	Thu. 9/11 (1) Fri. 10/11 (1)	Translations 1&2 Screencasts 1&2 Written evaluation	Due to technical problems not all participants had time to complete the 2nd translation. 5 participants no show (29-5 = 24) One participant joined group C after the break and was placed in group C for the rest of the experiment. 1 to group C (24+1 = 25)
2	13/11	23	Thu 16/11(1) Fri 17/11 (1)	Lesson book	*2 participants changed D>C (25+2 = 27)* 1 participant changed C>D (27-1 = 26) 1 participant dropped out (26-1 = 25)
3_0	20/11	-	-	-	*Postponed*
3_1	27/11	22	Fri 1/12 (3)	Lesson book Evaluation in Socrative	
4	04/12	21	Fri 8/12 (3)	Lesson book	*1 participant changed C>D (25-1 = 24)*
5_0	11/12	-	-	-	Cancelled due to snowstorm

5_1	13/12	13	-	Lesson book Evaluation in Socrative	Catch up session (Wednesday)
5_2	15/12	10	Mon 18/12(1)	Lesson book Evaluation in Socrative	Catch up session (Friday)
6	18/12	21	Thu 21/12 (3)	Translations 3&4 Screencasts 3&4 Evaluation in Socrative	
7	15/01	20	Thu 18/1 (3) Fri 19/1 (1)	Translations 5&6 Screencasts 5&6 Evaluation in Socrative	

APPENDIX C

VIII-D: Product Condition

Group D (Thursday): n= 22 initial, 15 final

Session	Date	Present	Catch up session	Data	Particulars
1	9/11	15	-	Translations 1&2 Screencasts 1&2 Written evaluation	*7 participants no show (22-7=15)*
2	16/11	13	Fri 17/11 (1)	Lesson book	*2 participants changed D>C (15-2=13)* *1 participant changed C>D (13+1=14)*
3	23/11	10	Mon 27/11 (4)	Lesson book Evaluation in Socrative	
4	30/11	14	Mon 4/12 (1)	Lesson book	*1 participant changed C>D (14+1= 15)*
5	7/12	14	Wed 13/12(1)	Lesson book Evaluation in Socrative	
6	14/12	13	Mon 18/12(2)	Translations 3&4 Screencasts 3&4 Evaluation in Socrative	
7	18/01	15		Translations 5&6 Screencasts 5&6 Evaluation in Socrative	

APPENDIX D - CHAPTER 6

Table LXIII: Effect on Coherence for Initial Latin and Dutch Proficiency

Models Learner Variable: Latin	X^2	df	Comparison Models	X^2	df	p
(0) Y = C + [variances]	545.366	3				
(1) + time	545.264	4	0 vs 1	0.102	1	0.749
(2) + condition	544.795	5	1 vs 2	0.469	1	0.493
(3) + time * condition	543.000	6	2 vs 3	1.795	1	0.180
(4) + Learner variable	525.445	7	3 vs 4	17.555	1	0.000
(5) + LV * time	525.328	8	4 vs 5	0.117	1	0.732
(6) + LV * condition	525.324	9	5 vs 6	0.004	1	0.950
(7) + LV * time * condition	524.751	10	6 vs 7	0.573	1	0.449

Models Learner Variable: Dutch	X^2	df	Comparison Models	X^2	df	p
(0) Y = C + [variances]	545.366	3				
(1) + time	545.264	4	0 vs 1	0.102	1	0.749
(2) + condition	544.795	5	1 vs 2	0.469	1	0.493
(3) + time * condition	543.000	6	2 vs 3	1,795	1	0,180
(4) + Learner variable	542,115	7	3 vs 4	0,885	1	0,347
(5) + LV * time	541,999	8	4 vs 5	0,116	1	0,733
(6) + LV * condition	538,830	9	5 vs 6	3,169	1	0,075
(7) + LV * time * condition	538,063	10	6 vs 7	0,767	1	0,381

Table LXIV: Effect on Equivalence for Initial Latin and Dutch Proficiency

| Models | | | Comparison | | | |
Learner Variable: Latin	X^2	df	Models	X^2	df	p
(0) Y = C + [variances]	680.272	3				
(1) + time	646.228	4	0 vs 1	34.044	1	0.000
(2) + condition	645.849	5	1 vs 2	0.379	1	0.538
(3) + time * condition	635.639	6	2 vs 3	10.210	1	0.001
(4) + Learner variable	623.722	7	3 vs 4	11.917	1	0.001
(5) + LV * time	622.375	8	4 vs 5	1.347	1	0.246
(6) + LV * condition	622.371	9	5 vs 6	0.004	1	0.950
(7) + LV * time * condition	621.057	10	6 vs 7	1.314	1	0.252
Models			Comparison			
Learner Variable: Dutch	X^2	df	Models	X^2	df	p
(0) Y = C + [variances]	680.272	3				
(1) + time	646.228	4	0 vs 1	34.044	1	0.000
(2) + condition	645.849	5	1 vs 2	0.379	1	0.538
(3) + time * condition	635.639	6	2 vs 3	10.210	1	0.001
(4) + Learner variable	633.183	7	3 vs 4	2.456	1	0.117
(5) + LV * time	633.183	8	4 vs 5	0.000	1	1.000
(6) + LV * condition	630.027	9	5 vs 6	3.156	1	0.076
(7) + LV * time * condition	629.901	10	6 vs 7	0.126	1	0.723

APPENDIX E - CHAPTER 6

Effect of learner variables: initial levels of Coherence and Equivalence on similar and contrasting output variables. Table LXV: Effects on similar output variable; Table LXVI: Effects on contrasting output variable. Output variable: Winsorised scores.

Table LXV: Effect of Initial Coherence Score on Coherence and Initial Equivalence score on Equivalence

Models Learner Variable Coherence	X^2	df	Comparison Models	X^2	df	p
(0) Y = C + [variances]	545.366	3				
(1) + time	545.265	4	0 vs 1	0.101	1	0.751
(2) + condition	544.795	5	1 vs 2	0.470	1	0.493
(3) + time * condition	543.000	6	2 vs 3	1.795	1	0.180
(4) + LV	519.844	7	3 vs 4	23.156	1	0.000
(5) + LV * time	519.827	8	3 vs 5	0.017	1	0.896
(6) + LV * condition	519.824	9	4 vs 6	0.003	1	0.956
(7) + LV * time * condition	518.336	10	5 vs 7	1.488	1	0.223

Models Learner Variable Equivalence	X^2	df	Comparison Models	X^2	df	p
(0) Y = C + [variances]	680.272	3				
(1) + time	646.228	4	0 vs 1	34.044	1	0.000
(2) + condition	645.849	5	1 vs 2	0.379	1	0.538
(3) + time * condition	635.639	6	2 vs 3	10.210	1	0.001
(4) + LV	627.374	7	3 vs 4	8.265	1	0.004
(5) + LV * time	627.357	8	4 vs 5	0.017	1	0.896
(6) + LV * condition	621.71	9	5 vs 6	5.647	1	0.017
(7) + LV * time * condition.	610.037	10	6 vs 7	11.673	1	0.001

Table LXVI: Effect of Initial Coherence score on Equivalence and Effect of Initial score of Equivalence on Coherence

Models			Comparison			
Learner Variable: Coherence	X^2	Df	Models	X^2	df	p
(0) Y = C + [variances]	680.272	3				
(1) + time	646.228	4	0 vs 1	34.044	1	0.000
(2) + condition	645.849	5	1 vs 2	0.379	1	0.538
(3) + time * condition	635.639	6	2 vs 3	10.210	1	0.001
(4) + LV	625.870	7	3 vs 4	9.769	1	0.002
(5) + LV * time	625.869	8	4 vs 5	0.001	1	0.975
(6) + LV * condition	625.506	9	5 vs 6	0.363	1	0.547
(7) + LV * time* condition	621.449	10	6 vs 7	4.057	1	0.044
Models			Comparison			
Learner Variable: Equivalence	X^2	Df	Models	X^2	df	p
(0) Y = C + [variances]	545.366	3				
(1) + time	545.265	4	0 vs 1	0.101	1	0.751
(2) + condition	544.795	5	1 vs 2	0.470	1	0.493
(3) + time * condition	543.000	6	2 vs 3	1.795	1	0.180
(4) + LV	531.634	7	3 vs 4	11.366	1	0.001
(5) + LV * time	531.614	8	4 vs 5	0.020	1	0.888
(6) + LV * condition	529.684	9	5 vs 6	1.930	1	0.165
(7) + LV * time* condition	526.878	10	6 vs 7	2.806	1	0.094

LIST OF RELATED PUBLICATIONS
AND PRESENTATIONS

Publications

Luger, S. (2015). Latijn vertalen in de bovenbouw: een mission impossible? [Translating Latin in upper secondary education: a mission impossible?]. *Lampas*, *48*, 212–225.

Luger, S. (2016). Vertalingen beoordelen in de bovenbouw: Een cola-tic of een glijdende schaal? [Assessing translation in upper secondary education: A Coca Cola with a twist or a sliding scale?]. *Lampas*, *49*, 3-19.

Luger, S. (2018). How do Dutch adolescents translate Latin into coherent Dutch? A Journey into the Unknown. *Journal of Latin Linguistics*, *17*(2), 333–365. doi: https://doi.org/10.1515/joll-2018-0015.

Presentations and Workshops

Luger, S. (2015). *Vertalingen beoordelen, (hoe) kan het anders?* [Translation Assessment, (how) can it be performed differently?] Workshop presented at the VCN Nazomerconferentie, Nunspeet.

Luger, S. (2016). *Eye-tracking en onderzoek naar vertalen.* [Eye-tracking and Translation Research] Presentation at the OIKOS PhD/ReMa presentation days, Ravenstein.

Luger, S. (2016). *Coherente tekstproductie bij vertalen* [Coherent Textproduction in Translation.] Workshop presented at the Regionale Burendag, Heerhugowaard.

Luger, S. (2016). *Begrijp je wat je vertaalt, of vertaal je wat je begrijpt?* [To Understand what is Translated or to Translate what is Understood?] Workshop presented at the VU nascholingsdag, Amsterdam.

Luger, S. (2017). Paper. *How do Dutch adolescents translate Latin into coherent Dutch? A Journey into the Unknown*. Presentation at the International Colloquium on Latin Linguistics, Munich, Germany.

Luger, S. (2017). *Intervention Study in Upper Secondary Education on the Translation of a Latin Source Text into a Coherent Dutch Target Text*. One-Slide presentation at the Mini Conference on Writing, Gand, Belgium.

Luger, S. (2017). *Lost in Translation: How to teach students to translate Latin into a coherent target text.* Workshop presented at the VCN Nazomerconferentie / Euroclassica Conference, Leiden.

Luger, S. and Gerbrandy, P.S. (2018). *Vertalen als schrijven in de klas* [Translation as Writing in the Classroom]. Workshop for Illustere School, University of Amsterdam.

Luger, S. (2018). *Aan mij is een verlangen… mijn vertaling te redigeren tot een co-herente tekst* [To Me Is the Desire to Revise my Translation into a Coherent Text]. Workshop presented at the VCN Nazomerconferentie, Nunspeet.

Luger, S. (2018). *Vertalen en tekstcoherentie* [Translation and Target-Text Coher-ence]. Presentation for the Landelijke LIO-dag [National Day for Teachers in Training], Amsterdam.

Luger, S. (2019). *Vertalen, hoe gebeurt dat eigenlijk?* [Translation, How Does it Hap-pen?]. Presentation for the Landelijke LIO-dag [National Day for Teachers in Training], Amsterdam.

Luger, S. (2019). *Vertalen is business as usual. (Hoe) leren we onze leerlingen verta-len?* [Translation is Business as Usual. (How) Do we Teach Translation?]. Work-shop presented at the VU nascholingsdag, Amsterdam.

Luger, S. (2019). *Vertalen is schrijven* [Translation is Writing]. Lesson presented at the Publieksmiddag Geesteswetenschappen, Vleuten.

CURRICULUM VITAE

Suzanne Luger (1966) obtained her degree in Latin and Greek Languages and Cultures from the University of Amsterdam in 1991. She completed her postgraduate teacher training at the Instituut voor de Lerarenopleiding (ILO) of the University of Amsterdam in 1992 and has worked as a teacher of Classics ever since. She has worked at Coornhert Lyceum (Haarlem), Johan van Oldenbarnevelt Gymnasium (Amersfoort), St. Ignatiusgymnasium (Amsterdam), Het 4e Gymnasium (Amsterdam) and is currently working at St. Ignatiusgymnasium again.

In addition to her teaching job in secondary education, she became a teacher trainer for Classics at the Graduate School of Child Development and Education of the University of Amsterdam in 2007. In 2009, she qualified as a University teacher (BKO) and in 2010 she obtained her VELON certificate as a teacher educator.

Suzanne served two terms (2009-2015) on the board of the Association of Dutch Classicists (VCN), she participated in the development of a Golden Standard for the Dutch Gymnasium (2012-2013) and has been a member of the editorial committee of Lampas, the Dutch peer-reviewed journal on Classics, since 2010. She was co-editor in chief of the two issues of Lampas on domain-specific teaching and learning (2011, 2015).

Suzanne participated in many conferences held by the Association of Dutch Classicists (VCN) and presented workshops on various subjects, from Translation to Integration of Language and Culture and Quality Assessment.

In September 2014, she started her part-time PhD research at the University of Amsterdam under the supervision of Caroline Kroon, Gert Rijlaarsdam and Suzanne Adema as part of the Dudoc Alfa project. In the years devoted to research and teaching, she remained active as an editor for Lampas and hosted workshops on several occasions, always focusing on the subject of her research.

Suzanne is married, lives in Amsterdam and is the mother of two adult children.

DANKWOORD

Een goedgekeurd onderzoeksvoorstel en de gelegenheid om vier jaar lang, drie dagen per week onderzoek te doen naar een kwestie waar ik me al lang over verbaasd had: de onzin die leerlingen kunnen produceren onder het mom van een vertaling. Wat een prachtige kans, mogelijk gemaakt door Dudoc Alfa.

De allereerste officiële dag in september 2014 zal ik niet snel vergeten. Suzanne, Kokkie en ik hadden afgesproken op de VU om een vliegende start te maken. In plaats van stralend op te stijgen werd ik overmand door de enorme last en verantwoordelijkheid die ik op mij voelde drukken. Hoe zou ik dit ooit tot een goed einde kunnen brengen? What was I thinking?

In de jaren die volgden heb ik het nooit meer als zo zwaar ervaren als op die eerste dag. Dat is te danken aan de mensen die me hebben geholpen, opgebeurd, gesteund, geadviseerd, afgeleid en die me – als dat nodig was – tot de orde hebben geroepen. De mensen die ik hier wil bedanken vallen in drie categorieën – die ook weer deels overlappen – uiteen: onderzoek, school en privé.

Caroline, jouw rol is cruciaal geweest in de totstandkoming van dit proefschrift. Als medevoorzitter van de verkenningscommissie die in 2010 haar spraakmakende rapport publiceerde sta jij in alle opzichten aan de wieg van dit onderzoek. Je was op de achtergrond altijd aanwezig en bereid tot meedenken. Het meest bijzonder vind ik wel de grondige en kritische manier waarop jij alles gelezen hebt. Als ik een impressionist ben, ben jij een fijnschilder. Jij wist van het begin tot het eind mijn weidse gebaren tot meer precisie te dwingen, en daarvoor wil ik je bedanken.

Het begeleiden van een ervaren docente, maar onervaren onderzoekster is best een klus. Daarom wil ik Suzanne en Gert bedanken, die dat op een bijna dagelijkse basis hebben gedaan. Suzanne, je hebt me altijd met raad en daad bijgestaan en geen moeite was je teveel: van waardevolle literatuursuggesties tot het samen zitten worstelen met de statistische analyses. Het was een ontdekkingstocht waarin jij de weg wees, maar die we voor een deel ook samen ondernomen hebben. Ik durf wel te zeggen dat onze werkverhouding zich heeft ontwikkeld tot een vriendschap.

Gert, ik weet niet waar ik moet beginnen met je lof te zingen. Je opmerkzaamheid over wat wel en wat niet zo lekker liep en de terloopse manier waarop je -soms heikele- kwesties ter sprake bracht dwingen bewondering af. Het eindeloze geduld waarmee je me door SPSS hebt geleid. Dankzij jou

ontdekte ik ook dat ik in de winter echt veel minder productief ben en schafte ik De Lamp aan, die me door de duistere dagen van 2017 en 2018 geholpen heeft.

Dank ook Huub, dat je hulp bood als wij er echt niet meer uitkwamen. Dank Marco, voor de hands-on assistentie met de eye-tracker en de nazorg.

Tanja, Daphne, Liselore, Anouk, Saskia, Chelsea, Magdalena, Tran, Marloes, Marie-Thérèse en Marrit, dank voor jullie steun en fijne lablunches. Uddhava en Martijn, we hebben lekker geschreven! Dank voor jullie scherpe en kritische blik. Tessa, dank voor je praktische adviezen in de eindfase. Brenda, dank voor je support op afstand in de voor jou ook uitdagende jaren. Mijn taken op de ILO zijn tijdens mijn afwezigheid vier jaar lang voortreffelijk waargenomen door Xander en het laatste jaar door Riemer.

Het International Colloquium of Latin Linguistics in München zie ik als een scharnierpunt in het onderzoekstraject. Rodie en Lidewij, dank voor jullie gezelligheid en support daar, naast Caroline en Suzanne, en ook op het Turfdraagsterpad. De lange dagen op BG1 werden opgefleurd door Anna, Aniek, Loes, Lindsay, Anne, Marijn, Jan, Elon, Elianne, Isolde en natuurlijk de hangmat, de plantencollectie, de kaasfondue en de pannenkoeken.

Ik vond het heel bijzonder dat ik van twee oud-leerlingen hulp en assistentie mocht krijgen bij de uitvoering van mijn onderzoek: Dana, dank je voor het beoordelen van de vertalingen en Hannah, dank je voor het assisteren bij het experiment. Alle collega's classici dank ik, voor jullie deelname aan de workshops, de interesse voor mijn onderzoek en voor het inzetten van de piepjestest in jullie lessen. In het bijzonder dank ik Mannus, voor het lezen van een vroege versie van de eerste hoofdstukken. Alle leerlingen die mijn piepjestesten hebben ondergaan, die proefkonijn of proefpersoon waren: dank voor jullie enthousiasme en kritiek "het ligt niet aan uw materiaal hoor, ik ben gewoon heel moe."

Ik wil graag iedereen bedanken die eindeloos veel meer of minder coherente fabels heeft gerangschikt in D-PAC: Jacqueline, Henri, Antoinette, Xander, Thor, Susanne, Renske, Daniel, Renata, Clara, Sylvia, Gerke, Veronique, Linde, Otto, Pieterjan, Michelle, Paula, Eric, Wander, Edgar, Wilma, Johan, Pauline, Tessa en Nienke. Anouk, Liselore en Martijn ook hiervoor dank.

Alle collega's van de twee scholen waar ik de afgelopen jaren werkte wil ik bedanken voor hun belangstelling, de secties Klassieke Talen van het 4e gymnasium en het Ignatius, de ICT-ondersteuning op beide scholen, maar ook de rectoren die – ondanks al het gedoe van een docente die er gedeeltelijk niet is – toch dit traject blijmoedig hebben gesteund.

Na een uurtje roeien nemen we elke zaterdag het leven door: dank voor jullie luisterend oor en oprechte enthousiasme Renata, Suzan en Barbara.

Iedereen die mijn beperkte beschikbaarheid voor activiteiten, etentjes en gezelligheid heeft geaccepteerd: dank voor jullie begrip.

Susanne, jij was bij het onderzoek, op school en privé een steun en toeverlaat en een echte vriendin: dank je wel voor je hulp, je kritische blik en je vriendschap.

Mijn zus Tessa en mijn oudste vriend Hans, dank dat ik jullie mocht doorzagen in de afgelopen jaren, die ook voor jullie niet makkelijk zijn geweest.

Van mijn moeder mag ik haar nooit bedanken, maar dat doe ik toch. Voor het voorbeeld, de zorg, de liefde en de belangstelling.

Cornelia, moeder van de Gracchen, zag haar kinderen als haar mooiste sieraden. Julia, toen ik hiermee begon, stond jij aan het begin van het studentenleven. Inmiddels ben je op weg naar je Master. Het was bijzonder om 'samen' student te zijn. Onze lunches, tussen jouw colleges en mijn ijverig geschrijf door waren lichtpuntjes in de week. Dat je de correctie van het hele manuscript op je hebt genomen en voortreffelijk hebt uitgevoerd vervult mij met grote trots: met je eigen kind academisch verkeren, ik vind dat heel bijzonder. Roeland, toen ik hieraan begon "een kwart van je leven geleden" was je een nog maar een kleine knaap. Jij hebt me op zoveel manieren geholpen: je ging als 15-jarige in je kerstvakantie toch maar met je moeder mee naar Antwerpen om Latijn te vertalen met een eyetracker. Inmiddels kijk je met een kritische blik naar mijn grafieken en doet suggesties om ze te verbeteren. Je sprak me moed in als ik het niet meer zag zitten en ik mocht je bakpartner zijn in ons grote Holtkamp project. Je bent een gouden gozer.

Een promotietraject van vier of vijf jaar verloopt eigenlijk nooit zonder ingrijpende gebeurtenissen voor jezelf en de mensen die je het meest na staan, je familie en je beste vrienden. Dat was voor mij de afgelopen jaren niet anders. Dat ik iemand naast me heb met wie ik dat allemaal kan delen en met wie ik alles aan kan, is een voorrecht en een groot geluk: Robert, dank. En daarnaast natuurlijk dank dat je elke avond voor me kookt, en dat je me 's avonds laat doorwerken, maar ook wel eens terugfluit. Net als ik jou. Dat we ondanks onze volle werkdagen toch nog zoveel fijne films en series hebben gekeken en dat we binnenkort de noodgedwongen korte vakanties of stedentripjes weer eens kunnen uitbreiden. En dat we eigenlijk altijd lol kunnen hebben, zelfs als het tegenzit.

Amsterdam, december 2019