

Binnendifferenzierte Sachtexte als Unterstützung für den Erwerb von domänenspezifischen Wissensbeständen und Konzepten

Texts tematics differenziads a l'intern sco med per sustegnair l'acquisiziun dal savair e da concepts da domenas specificas

La différenciation interne de textes factuels comme moyen de promouvoir l'acquisition de savoirs et de concepts relevant de domaines spécifiques

Testi differenziati quale sostegno all'acquisizione di conoscenze e concetti specifici

Internally differentiated topical texts as a means to support learning

Executive Summary

Andreas Imhof, Eva Engeli, Markus Kübler

Herausgeber | Publié par

Institut für Mehrsprachigkeit
www.institut-mehrsprachigkeit.ch

—
Institut de plurilinguisme
www.institut-plurilinguisme.ch

AutorInnen | Auteurs

Andreas Imhof, Eva Engeli, Markus Kübler

Übersetzung | Traduction

Mary Carozza, Barbla Etter, Pro-verbial Särl, Joël Rey - Traduzioni e redazioni


Das vorliegende Projekt wurde im Rahmen des Arbeitsprogramms 2012-2014 des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit durchgeführt. Für den Inhalt dieser Veröffentlichung sind die AutorInnen verantwortlich.

Le projet dont il est question a été réalisé dans le cadre du programme de travail 2012-2014 du Centre scientifique de compétence sur le plurilinguisme. La responsabilité du contenu de la présente publication incombe à ses auteurs.

Freiburg | Fribourg, 2016

Layout

Billy Ben, Graphic Design Studio

 Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Département fédéral de l'intérieur DFI
Dipartimento federale dell'interno DFI
Departament federal da l'intern DFI
Bundesamt für Kultur BAK
Office fédéral de la culture OFC
Ufficio federale della cultura UFC
Uffizi federal da cultura UFC

Index

Deutsch	Binnendifferenzierte Sachtexte als Unterstützung für den Erwerb von domänenspezifischen Wissensbeständen und Konzepten	3
	1 Theoretische Fundierung und Fragestellung	4
	2 Methodisches Design	7
	3 Ergebnisse der Befunde	12
	4 Interpretation der Ergebnisse ... Bibliographie	16 83
Rumantsch	Texts tematics differenziads a l'intern sco med per sustegnair l'acquisiziun dal savoir e da concepts da domenas specificas	19
	1 Fundament teoretic e dumondas	20
	2 Design metodic	23
	3 Resultats	28
	4 Interpretaziun dals resultats ... Bibliografia	32 83
Français	La différenciation interne de textes factuels comme moyen de promouvoir l'acquisition de savoirs et de concepts relevant de domaines spécifiques	35
	1 Fondement théorique et hypothèse	36
	2 Modèle méthodologique	39
	3 Résultats	44
	4 Interprétation des résultats ... Bibliographie	48 83
Italiano	Testi differenziati quale sostegno all'acquisizione di conoscenze e concetti specifici	51
	1 Base teorica e formulazione	52
	2 Metodologia	55
	3 Risultati	60
	4 Interpretazione dei risultati... Bibliografia	64 83
English	Internally differentiated topical texts as a means to support learning	67
	1 Theoretical basis and hypotheses	68
	2 Methodological design	71
	3 Findings	76
	4 Interpretation of the findings ... Bibliography	80 83

Binnendifferenzierte Sachtexte als Unterstützung für den Erwerb von domänenspezifischen Wissensbeständen und Konzepten

Executive Summary

—

Andreas Imhof, Eva Engeli, Markus Kübler

1

Theoretische Fundierung und Fragestellung

Das Projekt *Binnendifferenzierte Sachtexte als Unterstützung für den Erwerb von domänenspezifischen Wissensbeständen und Konzepten* (BiST) basiert auf der Feststellung, dass Lehr- und Lernmittel (und die darin enthaltenen Sachtexte) immer noch den Kern des Unterrichts darstellen (Oelkers, 2010, S. 39; Kahlert, 2010), diese aber dem Anspruch eines individualisierenden Unterrichts nicht genügen (Moser Opitz, 2010). Die Diskussion versteht sich vor dem Hintergrund der Anforderungen an eine inklusive Schule und der organisatorischen Konsequenz der Binnendifferenzierung. Der Befund, dass Lehrmittel in der Regel viel zu schwierig für das mittlere Leseniveau sind, wird in vielen Untersuchungen in verschiedenen Domänen gestützt (u.a. Lässig, 2010; Leisen, 2006; von Borries, 2010). Es scheint plausibel, dass fremdsprachige Kinder besonders von diesem Umstand betroffen sind (Moser, 2001) und dass ein Bedarf in der Herstellung von niveaumässig angepassten Sachtexten besteht mit der begründeten Vermutung, dass binnendifferenzierte Lehr- und Lernmaterialien die Verstehensleistungen von Inhalten (Begriffe, Kategorien, Zusammenhänge, Konzepte) steigern können.

Aus diesen Feststellungen leiten wir zwei Hypothesen ab:

1. Hypothese: Binnendifferenzierte Lehr- und Lernmaterialien können die Verstehensleistungen von Inhalten (Begriffe, Kategorien, Zusammenhänge, Konzepte) sowohl bei leistungsstarken wie auch bei schwachen Leserinnen und Lesern steigern.

2. Hypothese: Binnendifferenzierung ist vor allem für schwache Leserinnen und Leser nützlich.

Ein *Indikator* für die Richtigkeit der Hypothesen wäre der grössere Lernzuwachs von Lerngruppen mit binnendifferenzierten Sachtexten gegenüber Lerngruppen mit nicht binnendifferenzierten Sachtexten. Unter Lernzuwachs verstehen wir den Erwerb von fachlichen Konzepten. Dabei gehen wir von konstruktivistischen Ansätzen aus, gemäss welchen Lernende neues Wissen auf der Grundlage ihres Vorwissens selbst konstruieren müssen (z.B. Duit & Widodo, 2004). Dies bedeutet, dass Lernumgebungen an das bestehende Wissen der Schülerinnen und Schüler anknüpfen müssen, insbesondere auch an Vorstellungen, die nicht mit der wissenschaftlichen Sicht übereinstimmen. Ein Nichtaufgreifen dieser Vorstellungen wird als eine der

Hauptursachen für Lernschwierigkeiten angesehen (z.B. Bransford, Brown & Cocking 2000; Duit, 1995; Treagust, Duit & Niewsandt, 2000). Hardy et al. (2010) gehen in ihrer Kompetenzmodellierung von einer allgemeinen Beschreibung des Konzeptwechsels (conceptual change, Möller, 2015; Sodian, 2008; Carey & Spelke, 1994; Carey, 2000; Carey, 2011; Möller, 2012; Anohina et al., 2009) von naiven Konzepten über Zwischenvorstellungen bis hin zu wissenschaftlich gültigen Vorstellungen aus. In einer vergleichbaren Art wird der Erwerb von fachlichen Konzepten im historischen Lernen auch im Projekt *Historisches Denken* beschrieben (Bietenhader & Kübler, 2012). Aufbauend auf die Vorarbeiten von Engelen et al. (2002) und Hardy et al. (2010) und den fachwissenschaftlichen Befunden der Altsteinzeit (Höhneisen, 1998; Parzinger, 2014; Osterwalder & Robert, 1990) formulierte das Projektteam eine Beschreibung historischer Konzepte von *Steinzeit*; dies trotz der Schwierigkeit, dass Geschichtswissenschaft eine „ill-defined“ Domäne darstellt (Günther-Arndt, 2006).

Das breite Wissen aus der Lesbarkeitsforschung hat sich bisher kaum auf die Produktion von Lehrmitteln ausgewirkt (Leitner, 2008, S. 708; Kahlert, 2010, S. 49f.). Textverständlichkeit im Sinne von Textlogik, Sachwortschatz und Lebensweltbezug ist nicht unabhängig vom inhaltlichen Bezug, sondern ganz wesentlich auch domänenspezifisch definiert und deshalb nur bedingt von Fachbereich zu Fachbereich transferierbar (Lindauer & Schmelletin, 2011; Sodian, 2008). Auf den Grundlagen der Forschung zu Textver-

ständlichkeit und Textverstehen (Wellenreuther, 2008; Schmölzer-Eibinger & Egger, 2012; Akhondi, 2010) entwickelten wir ein heuristisches Modell zu Textschwierigkeit und aufgrund dieses Modells eine Sacherzählung zur Altsteinzeit auf drei Leseniveaus, jeweils in Deutsch und Rätoromanisch. Als Faktoren für die Textschwierigkeit gelten die Textlänge, die Syntaxebene und die Wortebene. Grundlage für die Formulierung der Sacherzählung war die deutsche Fassung. Da sich Textschwierigkeit im Romanischen teilweise anders gestaltet (z.B. zusammengesetzte Wörter), wurden das heuristische Modell und die Sacherzählung dementsprechend angepasst. Um den Fluss der Geschichte beizubehalten, wurden Textfenster mit zusätzlichen Sachinformationen auf der Grundlage des heuristischen Modells eingebaut.

1.1 Sprachliche Ausgangslage

Die Schweiz ist ein viersprachiges Land mit einem föderalistischen Schulsystem. Das Rätoromanische ist eine Minderheitensprache im Kanton Graubünden, die verwendeten Lehrmittel sind (mit Ausnahme der Sprachlehrwerke) meist aus dem Deutschen übersetzt.

Abbildung 1
Sprachensituation in der Schweiz (Quelle: Bundesamt für Statistik)



Es gibt im Rätoromanischen fünf historisch gewachsene Schriftsprachen, Idiome genannt. Anders als bei der üblichen Verwendung des Begriffs *Idiom* in der Linguistik, wird er hier nicht für einen Dialekt gebraucht, sondern für die regionalen, standardisierten Schriftsprachen. Daneben gibt es seit 1982 eine gemeinsame Schriftvariante, das Rumantsch Grischun, die auf Bundes- und Kantonsebene offiziell anerkannt ist.

Die an der Studie beteiligten rätoromanischsprachigen Schülerinnen und Schüler stammen aus drei Regionen mit unterschiedlichen Idiomen. Das Übersetzen der Texte in die einzelnen Idiome war unter anderem aus dem Grund komplex, dass die

lokalen Idiome oft sehr unterschiedliche Wörter für dieselbe Sache verwenden. Exemplarisch sei hier das Wort Heidelberg erwähnt. Bereits das Grundwort Beere ist in den beiden Idiomen Vallader (pommaraida) und Sursilvan (coclas) völlig unterschiedlich. Das Wort Heidelberg, welches im Deutschen als Differenzierung des Worts Beere leicht erkennbar ist, lässt sich weder in Vallader (uzuns) noch in Sursilvan (izuns) vom Stammwort Beere ableiten. Wie in diesem Beispiel gibt es in jeder Schriftvariante des Rätoromanischen viele Wörter, die nicht wie im Deutschen vom Oberbegriff hergeleitet werden können.

2 Methodisches Design

2.1 Stichprobe und Sacherzählung

Zur Untersuchung der Hypothesen wurde ein quasi-experimentelles Design verwendet und in sechs Klassen aus dem Kanton Schaffhausen und neun Klassen aus dem Kanton Graubünden mit insgesamt 221 Kindern (davon 108 Mädchen und 113 Knaben) der 4. Primarstufe durchgeführt.

Es konnten 204 Datensätze ausgewertet werden, wobei ein Datensatz aus zwei Audiodateien der Interviews (Prätest, Posttest), einer Zeichnung und zwei Fotografien der Strukturlegetechnik SLT (Prätest, Posttest) besteht. 17 Datensätze mussten ausgeschieden werden, da sie unvollständig waren oder technische Mängel aufwiesen.

Für die Untersuchung im Kanton Schaffhausen wurden Schulen mit einem mittleren bis hohen Anteil an Kindern ausgewählt, deren Muttersprache nicht Deutsch ist, sich also von der Schulsprache unterscheidet. Im Kanton Graubünden wurden Schulen der Regionen Surselva und Unterengadin/Münstertal ausgewählt. Ein besonderes Auswahlkriterium war nicht nötig, ausser dass alle Schulen in diesen Regionen von Schulbeginn bis zur vierten Klasse vollständig in Romanisch unterrichtet werden.¹

Für das Treatment wurden Sachtexte auf drei Schwierigkeitsniveaus in Deutsch verfasst und für den Kanton Graubünden in die rätoromanischen Idiome Sursilvan und Vallader (bzw. in die Standardsprache Rumantsch Grischun für das Münstertal) übertragen und angepasst.

1 | Dies im Unterschied zu anderen Schulgemeinden im romanischen Sprachgebiet, welche von der ersten Klasse an zweisprachig geführt werden. Allgemein ist in rätoromanischen Gebieten ein mittlerer bis hoher Anteil an Kindern zu verzeichnen, deren Muttersprache nicht der Schulsprache entspricht.

2.2 Untersuchungsablauf und Befragung der Kinder

Die Datenerhebung wurde zeitnah mit einem Prätest, dem Treatment und einem Posttest durchgeführt und lief nach folgendem Schema ab:

Tabelle 1
Schema Untersuchungsablauf

Experimentalgruppe		
Prätest	Intervention 25-30 Min. Lesen des Sachtextes	Posttest
Interview zu Zeichnung und Strukturlegemethode (SLT)	starke Leserinnen und Leser hohes Niveau im Sachtext	Interview zu Sachtext und SLT
	mittlere Leserinnen und Leser mittleres Niveau im Sachtext	
	schwache Leserinnen und Leser tiefes Niveau im Sachtext	
Kontrollgruppe		
Prätest	Intervention	Posttest
Interview zu Zeichnung und Strukturlegemethode	hohes Niveau im Sachtext	Interview zu Sachtext und SLT

Zu Beginn mussten die Kinder eine Zeichnung zum Thema Steinzeit anfertigen. Diese diente dazu, das Vorwissen zu aktivieren. Der Prätest bestand aus dem anschließenden Prä-Interview und der Strukturlegemethode. Im Interview wurden die Kinder zuerst zu ihrer Zeichnung befragt, danach mussten sie laut denkend (vgl. Fromann, 2005) vorgegebene Bild- und Wortkarten zum Thema Steinzeit den drei Kategorien „gehört zur Steinzeit“, „gehört nicht zur Steinzeit“ und „Begriff unbekannt oder unklar“, zuordnen

(siehe Tabelle 2) und damit eine Legeordnung erstellen (Strukturlegemethode SLT). In einem zweiten Schritt mussten die Kinder die Karten zueinander in Beziehung bringen, dies möglichst begründen und die Karten entsprechend legen. Im Posttest wurden die Kinder als erstes nach der Geschichte im Sachtext befragt, danach legten sie erneut die Bild- und Wortkarten mit der Aufforderung, allfällige Änderungen gegenüber der ersten Legeordnung möglichst zu begründen.

Um einheitliche Untersuchungen in Schaffhausen und Graubünden zu garantieren, wurde ein Interviewleitfaden mit genauen Hinweisen für die Interviewführung erstellt. Die Interviewenden wurden vorgängig in ihrer Aufgabe geschult.

2.3 Auswertungsmethoden

a) Auswertung der Interviews mit Strukturlegetechnik (Legeordnung)

Die einzelnen Items konnten von den Kindern in drei Felder gelegt werden, wobei dies aus fachlicher Sicht jeweils korrekt oder falsch sein konnte. Daraus ergeben sich für jedes Item sechs mögliche Kategorisierungen (siehe Tabelle 2).

Tabelle 2
Auswerteschema und kognitionspsychologische Interpretation aufgrund der Lage der Items. Die weissen Felder wurden als korrekt gewertet, die eingefärbten als falsch. Die leicht eingefärbten Felder wurden je nach Skala anders gewertet (siehe unten).

Fachliche Beurteilung des Items		Gehört zur Steinzeit	Gehört nicht zur Steinzeit	Weiss nicht
	Korrekte Items (Anzahl=22)	„Treffer“: korrektes Wissen bzw. korrekt im Gedächtnis vernetztes Wissensselement	„Verpasser“: falsches Wissen bzw. falsch im Gedächtnis vernetztes Wissensselement	„Unsicherheit falsch“: falsches „Nichtwissen“, noch nicht vernetztes Wissensselement bzw. bekanntes Wissensselement und Unsicherheit über die Zugehörigkeit
Falsche Items (Anzahl=6)	„Falscher Alarm“: Fehlkonzept, falsch im Gedächtnis vernetztes Wissensselement	„Korrekte Ablehnung“: korrektes Wissen bzw. korrekt nicht vernetztes Wissensselement	„Unsicherheit korrekt“: korrektes „Nichtwissen“, noch nicht vernetztes Wissensselement und Unsicherheit über die Zugehörigkeit	

Für jede Schülerin und jeden Schüler kann so ein Wert errechnet werden, welcher ihre bzw. seine Leistung widerspiegelt, wobei höhere Werte eine bessere Leistung aufzeigen:

*Skala Strukturlegetechnik =
Treffer + Korrekte Ablehnung = 28
– Verpasser – Falscher Alarm
– Unsicherheit, korrekt – Unsicherheit,
falsch*

Die Skala nimmt an, dass die unter „weiss nicht“ abgelegten Items den Schülerinnen und Schülern unbekannt sind und somit als „Fehler“ angerechnet werden, da die Versuchspersonen die Zugehörigkeit nicht explizit angeben können.² Für die Werte der Skala wurde in der Folge eine zweifaktorielle Varianzanalyse mit Messwiederholung berechnet (unabhängige Variablen: Experimentalgruppe, Leseniveau), um die Hypothesen der Studie zu testen.

Durch die Legeordnung kann die kognitive Organisation des Wissens bei den Versuchspersonen beurteilt werden. Diese Analyse beschränkt sich allerdings auf die bloße Zuordnung der Items zum Fachthema Altsteinzeit. Die konzeptuellen Vorstellungen der Schülerinnen und Schüler zur Altsteinzeit können durch die Analyse der

Legeordnung nicht beurteilt werden.³ Zwei Punkte können aufgrund der SLT beurteilt werden: erstens, ob ein Begriff den Lernenden bekannt ist und zweitens, ob sie diesen kognitiv mit dem Inhaltsbereich Altsteinzeit in Beziehung setzen. Durch die Analyse der Legeordnung wird das Lernen deshalb auf basalem Niveau, d.h. der Bildung von assoziativen Verknüpfungen von einzelnen Gedächtnisinhalten gemessen.

b) Auswertung der Interviews mittels Matrix des Conceptual Change

Die Auswertungen der Interviews erfolgten ab den Audio-Dateien, indem die Bewertung der Aussagen direkt in der entsprechenden Ebene im Raster vermerkt wurde. Die Definition der vier Niveaus geschah in einer ersten Phase heuristisch beziehungsweise domänenspezifisch und wurde durch zwei Pilotuntersuchungen Schritt für Schritt verfeinert und mit Ankerbeispielen operationalisiert.

3 | Eine weiterführende Untersuchung, welche stärker das inhaltliche Verständnis der Schülerinnen und Schüler testen kann, wurde im Rahmen der Studie angedacht, aber aufgrund des Aufwandes nicht weiter verfolgt. Die Grundidee liegt darin, nicht nur die Kategorisierung (gehört zu; gehört nicht zu), sondern vor allem die Positionierung der Begriffe untereinander zu analysieren.

2 | Eine weitere Skala, welche die unter „weiss nicht“ abgelegten Items mit berücksichtigt, wurde ebenfalls getestet. Da sich die Resultate für die beiden Skalen nur geringfügig unterscheiden, wird nur die Skala „Strukturlegetechnik“ beschrieben.

Tabelle 3

Modell zur Abbildung der Konzeptebenen historischen Wissens (und Auswertungsraster mit den vier definierten Konzeptstufen (nach Engelen et al., 2002 und Moser et al., 2007)

D2	Konzepte historischen Denkens im Sinne einer Einsicht in Re- und Dekonstruktion Konzeptebene 2: Bezug zu heute, Vergleich heute-früher
D1	Vorkonzepte historischen Denkens im Sinne konzeptueller Zugänge zur Geschichte, zeitliche Einordnungen Konzeptebene 1: Vollständiges Konzept bei konkreten Begriffen, annäherndes Konzept bei abstrakten Begriffen
C2	Ausbaufähige Konzepte bzw. Schülervorstellungen (Zusammenhänge etc. komplexerer Art) Zusammenhänge 2: Kategorien, Oberbegriffe, Abstrakta; Satzkonstruktionen mit Konjunktionen (da, weil, obschon etc.)
C1	Ausbaufähige Konzepte bzw. Schülervorstellungen (Zusammenhänge etc. einfacher Art) Zusammenhänge 1: Kategorien, Oberbegriffe, Abstrakta; finale Aussagen (um zu...); Funktion mit passendem Verb
B2	Fachwissenschaftlich nicht belastbare Konzepte bzw. Schülervorstellungen (differenzierte Aufzählungen, Alltagstheorien, Gegenwartsanalogien, Vermutungen) Spezifische Begriffe und Konkreta; differenziertere Verben (benutzen, wohnen); Nennung > 2 Begriffe; konkretere Vermutung
B1	Fachwissenschaftlich nicht korrekt vernetzte Wissens Elemente bzw. Schülervorstellungen (einfache Aufzählungen, Alltagstheorien, Gegenwartsanalogien, Vermutungen) Allgemeine Begriffe und Konkreta; Verb haben; vage Vermutungen
A	Kein korrektes historisches Konzept Allgemeine Aussagen aus der Lebenswelt der Kinder

Die in den Interviews gewonnenen Daten dienten zur Messung von Veränderungen im Konzeptlevel. Äusserungen der unteren Konzeptebenen kamen dabei häufiger vor als Äusserungen der höheren Ebenen. Um nun zu höheren Konzeptebenen gehörenden Aussagen mehr Gewicht zu geben, wurden diese mit einem Index gewichtet und bekamen so einen höheren Wert. Wenn ein Kind im Interview beispielsweise nur Äusserungen zu Konzeptebene D2 macht, ergibt dies einen höheren Wert als wenn ein Kind gleichviele

Äusserungen auf der Konzeptebene B1 macht.

c) Interrater-Reliabilität

Zur Quantifizierung der Interrater-Reliabilität wurden vier Interviews ausgewählt, separat gewertet und miteinander verglichen. Es wurde eine Übereinstimmung von 0.804 erzielt, was als gut bezeichnet werden kann. Diese Ratingart dient auch als Grundlage für die Auswertung der Daten (siehe Kap. 2.3 a).

3

Ergebnisse der Befunde

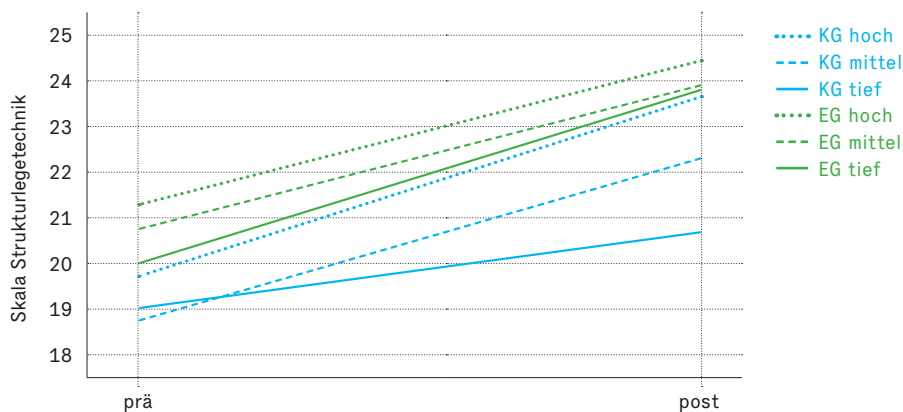
3.1 Veränderung kognitiver Strukturen durch Sachtexte

Der Prä-/Posttestvergleich der Legeordnungen (SLT) ergibt eine Verbesserung der Leistungen im Begriffswissen sowohl für die Experimentalgruppe als auch für die Kontrollgruppe (siehe Abbildung 2). Da beinahe alle Schülerinnen und Schüler eine Verbesserung beim Score der Strukturlegetechnik erreichen und diese Leistungssteigerung relativ markant ist, ergeben sich aus dem Vergleich der Prä- und der Posttests der beiden Gruppen sehr hohe Effektstärken:

Experimentalgruppe $d=1.38$, Kontrollgruppe $d=1.07$. Ein Grossteil der Leistungssteigerung kann dadurch erklärt werden, dass bei der Aussage „weiss nicht“ gelegte Items erst in der Posttestung korrekt in die Legeordnung integriert wurden. Aufgetrennt auf das Leseniveau zeigen ebenfalls alle Gruppen eine Leistungssteigerung.

Für das Begriffswissen lassen sich signifikante Haupteffekte für die Experimentalgruppe ($F(1/205)=13.104$, $p<0.001$), das Leseniveau ($F(2/205)=3.656$, $p=0.027$) und den Testzeitpunkt ($F(1/205)=330.242$, $p<0.001$) nachweisen. Ebenfalls lässt sich eine signifikante Dreifachinteraktion der Ex-

Abbildung 2
Veränderung Begriffswissen Experimentalgruppe (EG) und Kontrollgruppe (KG)



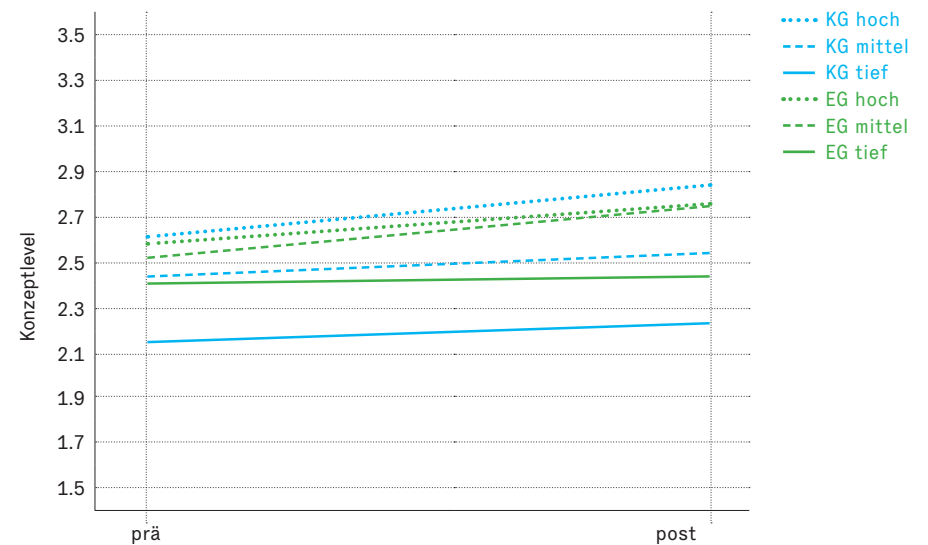
perimentalgruppe, des Leseniveaus und des Testzeitpunkts nachweisen ($F(2/205)=5.133$, $p=0.007$).

Als Hauptunterschied zwischen Experimentalgruppe und Kontrollgruppe kristallisiert sich die Leistungssteigerung bei den schwachen Leserinnen und Lesern heraus: Bei einem angepassten Lesetext (tiefes Niveau) der Experimentalgruppe zeigen die schwachen Leserinnen und Leser einen stärkeren Leistungsanstieg als die mittleren und starken. Bei der Kontrollgruppe ist der Leistungsanstieg bei den schwachen Lesenden deutlich kleiner als bei den anderen beiden Leistungsgruppen.

3.2 Veränderungen im Konzeptwissen durch Sachtexte

Für das Konzeptwissen lassen sich signifikante Haupteffekte für das Leseniveau ($F(2/192)=14.856$, $p<0.001$) und den Testzeitpunkt ($F(1/205)=22.425$, $p<0.001$) nachweisen. Hingegen erreicht keine der möglichen Interaktionen das notwendige Signifikanzniveau. Die untenstehende Abbildung 3 zeigt die Entwicklung zwischen Prä- und Posttest im Vergleich zwischen der Experimentalgruppe und der Kontrollgruppe in grafischer Darstellung.

Abbildung 3
Veränderung Konzeptwissen Experimentalgruppe (EG) und Kontrollgruppe (KG)



Die Mittelwertzunahmen in Tabelle 4 zeigen, dass die untersuchten Kinder insgesamt eine (leichte) Steigerung ihres historischen Verständnisses verzeichnen. Evident ist auch, dass sich die Leseleistungsniveaus (hoch, mittel, tief) deutlich abbilden. Tabelle 4 beschreibt alle Kinder; Tabelle 5 unterscheidet zwischen den Sprachgruppen. Der Effekt der Zunahme ist allerdings eher schwach und bestätigt unsere Ausgangshypothese nicht. Es bestehen nur geringe Unterschiede zwischen der Experimental- und der Kontrollgruppe. Interessant sind jedoch die Details: Bei der *Experimentalgruppe* (angepasste Lesetexte) erreichen die mittleren Leserinnen und Leser den grössten Zuwachs, während dies in der *Kontrollgruppe* (schwieriger Lesetext für alle) die starken Leserinnen und Leser sind.

Für die schwachen Leserinnen und Leser macht es offenbar beim Aufbau von konzeptionellem Verständnis von Geschichte – im Unterschied zum Aufbau von Begriffswissen – keinen Unterschied, ob sie

einen schwierigen oder einen leichten Sachtext lesen: Die schwachen Leserinnen und Leser weisen hier kaum einen Lernzuwachs auf.

Die oben beschriebenen Ergebnisse akzentuieren sich noch, wenn man die Sprachgruppen getrennt analysiert. Bei den deutschsprachigen Kindern zeigen in der Experimentalgruppe (mit angepassten Texten) die mittleren Leserinnen und Leser einen signifikanten Leistungszuwachs zwischen den Prä- und Posttests, während in der Kontrollgruppe die lesestarken Schülerinnen und Schüler den grössten Zuwachs verzeichnen. Bei den rätoromanischsprachigen Leserinnen und Lesern verzeichnen in beiden Untersuchungsgruppen die lesestarken Schülerinnen und Schüler den grössten Leistungszuwachs. Dieser markante Unterschied zwischen den Sprachgruppen ist zurzeit nicht erklärbar.⁴

4 | Wir verfolgen in der Klärung dieser Frage zur Zeit drei Möglichkeiten:

1. *Stichprobeneffekt*: Wir untersuchen, ob wir in der rätoromanischsprachigen Teilstichprobe besondere Spracheffekte (etwa deutschsprachige Kinder, die rätoromanisch eingeschult wurden) haben. Erste Abklärungen zeigen, dass diese Erklärung eine kleine Wahrscheinlichkeit aufweist.
2. *Textseitige Effekte*: Wir prüfen noch einmal, ob die Textschwierigkeiten zwischen Rätoromanisch (Sursilvan, Vallader, Rumantsch Grischun) und Deutsch eventuell nicht identisch gewesen waren. Zudem ist zu beachten, dass in den Interviews die rätoromanischsprachigen Kinder des

Öfteren einen deutschen Fachbegriff (wie „Burggraben“) kannten, der rätoromanische Begriff ihnen aber nicht bekannt war. Dieser Effekt scheint uns der wahrscheinlichste zu sein.

3. *Ratingeffekt*: Die Ratingvorlage des Konzeptwechsels und die Ankerbeispiele waren in Deutsch und beruhten auf deutschen Beispielen. Die Interviews der rätoromanischen Kinder wurden also aufgrund deutschsprachiger Ratings gewertet. Da aber die rätoromanischsprechenden Codierinnen und Codierer sehr gut Deutsch sprechen, scheint uns dieser Effekt eher wenig wahrscheinlich.

Tabelle 4

Mittelwertunterschiede zwischen Prä- und Posttests hinsichtlich der Konzeptebenenveränderungen (Wert 1 = eine Konzeptebene von insgesamt sechs) (LN = Leseniveau)

Veränderungen ⁵	Alle Niveaus	LN tief	LN mittel	LN hoch
Experimentalgruppe	0.14	0.03	0.22	0.17
Kontrollgruppe	0.15	0.08	0.10	0.22

Tabelle 5

Mittelwertunterschiede zwischen Prä- und Posttests hinsichtlich der Konzeptebenenveränderungen (Wert 1 = eine Konzeptebene von insgesamt sechs) nach Sprachgruppen (LN = Leseniveau)

Veränderungen	Deutschsprachige Kinder				Rätoromanischsprachige Kinder			
	Alle Kinder	LN tief	LN mittel	LN hoch	Alle Kinder	LN tief	LN mittel	LN hoch
Experimentalgruppe	0.18	0.04	0.50	0.09	0.10	0.02	-0.02	0.27
Kontrollgruppe	0.16	-0.01	0.08	0.27	0.14	0.11	0.09	0.17

5 | Die Werte spiegeln die relative Veränderung zu den definierten Niveaus (A, B1, B2, C1, C2, D1, D2). Ein Differenzwert von 1 bedeutet demnach die Zunahme eines ganzen Niveauschrittes (z.B. von C2 zu D1). Die Skala geht von 0 bis 6.

4

Interpretation der Ergebnisse und deren Bedeutsamkeit für die Produktion und Verwendung von Sachtexten

Gemäss unserem Setting haben binnendifferenzierte Texte offensichtlich das Potential, Begriffswissen bei allen Leseniveaus zu vertiefen und zu schärfen. Insbesondere profitierten die schwachen Leserinnen und Leser der Experimentalgruppe verglichen mit den schwachen Leserinnen und Lesern der Kontrollgruppe beim Begriffslernen überdurchschnittlich davon, dass sie zu einem Thema einen ihrem Leseniveau angepassten (d. h. leichteren) Sachtext bearbeiten konnten.

Kritisch ist zu bemerken, dass sich beim Vergleich der Experimental- und Kontrollgruppe kein signifikanter Unterschied im Zuwachs an Begriffswissen nachweisen lässt. Alle Leseniveaus der Experimentalgruppe profitieren bezüglich des Begriffswissens in vergleichbarem Ausmass. Dies könnte vor allem daran liegen, dass die Anzahl Versuchspersonen im tiefen Leseniveau zum Nachweis eines statistisch bedeutsamen Unterschieds zwischen Experimental- und Kontrollgruppe nicht genügend gross war.

Bezüglich des konzeptuellen Verständnisses lässt sich kein bedeutender Unterschied zwischen der Experimental- und der Kontrollgruppe feststellen, die Unterschiede entsprechen den Erwartungen (siehe Tabelle 4). Somit kann die Hypothese 1, nämlich

dass binnendifferenzierte Lehr- und Lernmaterialien die Verstehensleistungen von Inhalten (Kategorien, Zusammenhänge, Konzepte) für starke und schwache Leserinnen und Leser steigern, nicht bestätigt werden. Innerhalb der Experimentalgruppe sind die Unterschiede im konzeptuellen Verständnis zwischen den mittleren und starken Leserinnen und Lesern klein. Innerhalb dieser kleinen Prä-/ Posttest-Unterschiede steigt das konzeptuelle Verständnis bei den mittleren und starken Leserinnen und Lesern am meisten.

Die 2. Hypothese kann in Bezug auf das Begriffswissen bestätigt werden: Die schwächeren Leserinnen und Leser der Experimentalgruppe zeigen mit auf ihr Leseniveau angepassten Texten deutlich mehr Wissenszuwachs als diejenigen der Kontrollgruppe. Für den Erwerb konzeptuellen Verständnisses von Geschichte bieten dem Leseniveau angepasste Sachtexte den schwächeren Leserinnen und Lesern der Experimentalgruppe aber nach unseren Befunden nicht genügend Grundlagen. Unsere zweite Hypothese kann somit lediglich bezüglich des Begriffswissens bestätigt werden.

Die schwachen Effekte bezüglich Konzeptlevelveränderung mögen damit zu tun haben, dass die Intervention mit einem Lesetext, der in 20 bis 30 Minuten gele-

sen werden soll, vermutlich zu wenig Anreiz für einen Konzeptwechsel bedeutet. Es ist wahrscheinlich, dass ein Konzeptwechsel erst vertieft möglich wird, wenn neben der Lektüre einer Sacherzählung auch noch Fragen zum Text oder eine Verarbeitungsrunde angeschlossen werden. Das entspricht den veröffentlichten Befunden, die aufzeigen können, dass Präkonzepte relativ stabil sind (Möller, 1999) und nachhaltigere Interventionen erfordern. Um diese Hypothese zu verifizieren, müsste (wie dies in den erwähnten naturwissenschaftlichen Untersuchungen durchgeführt wurde) eine unterrichtliche Intervention (Verarbeitung und Vertiefung) die Binnendifferenzierung der Lektüre des Sachtextes begleiten.

Unsere Untersuchung testete ausschliesslich die Wirkung eines angepassten Sachtextes ohne die Wirkung von unterrichtlicher Verarbeitung und Vertiefung einzubeziehen. Trotzdem können als Trend Aussagen für die Produktion von Unterrichtsmaterialien sowie für die Durchführung von Unterricht getroffen werden.

Einerseits lohnt es sich offenbar, für schwache Leserinnen und Leser bei komplexen Sachthemen vereinfachte Sachtexte anzubieten. Diese ermöglichen den nicht routinierten Lesenden grundlegendes Begriffswissen und basale Begriffsnetze aufzubauen, die als Grundlage für unterrichtliche Verarbeitung und Vertiefung genutzt werden können. Schwierige Sachtexte, wie sie gemäss allgemeinem Befund üblicherweise in Lehrmitteln anzutreffen sind, scheinen für gute bis sehr gute Leserinnen und Leser keine allzu grosse Hürde darzustellen. Hier ist vermutlich binnendifferen-

zierender Unterricht wirksamer als eine textseitige Intervention. Für schwache Leserinnen und Leser scheinen schwierige Texte jedoch eine ernstzunehmende Hürde zu sein. Dem Erwerb von Begriffsnetzen durch vereinfachte Sachtexte muss vermutlich ein verarbeitender und vertiefender Unterricht folgen, um entsprechende (historische) Konzepte aufbauen zu können. Auch der Lehrplan 21 (D-EDK, 2014) in seiner Fassung vom November 2014 zeigt die Wichtigkeit der Verbindung von Sache und Sprache.

Unsere Daten werden noch weiter untersucht, insbesondere hinsichtlich der Effekte der unterschiedlichen Textniveaus im Rätoromanischen. Ebenfalls weiterzufolgen sind Untersuchungen von Lernarrangements, die die Wirkung binnendifferenzierender Sachtexte in Kombination mit angepassten Vertiefungs- und Verarbeitungsangeboten für die unterschiedlichen Leistungsniveaus testen.

Texts tematics differenziads a l'intern sco med per sustegnair l'acquisiziun dal savair e da concepts da domenas specificas

Executive Summary

—
Andreas Imhof, Eva Engeli, Markus Kübler

1

Fundament teoretic e dumondas

Il project „Texts tematics differenziads a l'intern sco med per sustegnair l'acquisiziun dal savair e da concepts da domenas specificas“ sa basa sin la constataziun che meds d'instruir e d'emprender (ed ils texts tematics cuntegnids en quels) statan anc adina en il center da l'instrucziun (Oelkers, 2010, p. 39; Kahlert, 2010), ma na satisfan betg a las pretensiuns per in'instrucziun individualisanta (Moser Opitz, 2010). Questa constataziun stat en connex cun las exigenzas ad ina scola inclusiva ed a las consequenzas organisatoricas da la differenziaziun interna. La diagnosa che meds d'instrucziun èn per regla bler memia difficils per in nivel da lectura mesaun, vegn sustegnida da diversas retschertgas en differentas domenas (tranter auter Lässig, 2010; Leisen, 2006; von Borries, 2010). I para plausibel che surtut uffants da linguas estras èn pertutgads da quest fatg (Moser, 2001) e ch'i dat in basegn da producir texts tematics adattads al nivel cun la supposiziun giustifitgada che materialias d'instruir e d'emprender differenziadas a l'intern pon augmentar la chapientscha da cuntegns (terminologia, categorias, associaziuns, concepts).

Da questas dumondas deducin nus duas ipotesas:

Ipotesa 1: Materialias d'instruir e d'emprender differenziadas a l'intern pon augmentar la prestaziun da la chapientscha da cuntegns (terminologia, categorias, associaziuns, concepts) tant tar lecturas e lecturs fermes sco era tar flaiVELS.

Ipotesa 2: La differenziaziun interna è surtut utila per lecturas e lecturs flaiVELS.

In *indicatur* per la correctadad da las ipotesas fiss l'augment da l'emprendi, sch'ins cumpareglia gruppas che emprendan cun texts tematics differenziads a l'intern cun gruppas che n'emprendan betg cun texts differenziads a l'intern. Cun l'augment da l'emprendi è manegiada l'acquisiziun da concepts istorics. Latiers ans basain nus sin disposiziuns constructivisticas: tenor quellas ston ils emprendiders construir sez lur savair sin basa da lur presavida (p.ex. Duit & Widodo, 2004). Quai signiftga che l'ambient pedagogic sto tegnair quint dal savair preexistent dals scolars e da las scolaras, surtut era da lur concepziuns che divergeschan dal punct da vista scientific. Il nun-respectar da questas concepziuns vegn considerà sco motiv principal per difficultads d'emprender (p.ex. Bransford, Brown & Cocking 2000; Duit, 1995; Treagust, Duit & Niewsandt,

2000). La modellaziun da las cumpetenzas da Hardy et al. (2010) parta d'ina descripziun generala da la midada conceptuala (conceptual change, Möller, 2015; Sodian, 2008; Carey & Spelke, 1994; Carey, 2000; Carey, 2011; Möller, 2012; Anohina et al., 2009) da concepts naivs sur concepziuns intermediaras fin a concepziuns scientificamain valablas. A moda sumeglianta vegn era descritta l'acquisiziun da concepts istorics en l'emprender istoric en il project *Pensar istoric (Historisches Denken)* (Bietenhader & Kübler, 2012). Sa basond sin las prelavurs dad Engelen et al. (2002) e Hardy et al. (2010) e las constataziuns scientificas dal temp paleolitic (Höhneisen, 1998; Parzinger, 2014; Osterwalder & Robert, 1990), ha il team da project formulà ina descripziun da concepts istorics per il temp da crap e quai malgrà la difficultad che la scienza da l'istorgia è ina domena nundefinida ("ill-defined") (Günther-Arndt, 2006).

Il savair fundà da la scienza da legibladad ha fin ussa strusch gi in'influenza sin la produenziun da meds d'instrucziun (Leitner, 2008, p. 708; Kahlert, 2010, p. 49s.). La chapibilitad d'in text en il senn da logica dal text, vocabulari spezialisà, relaziun cun l'ambient da viver n'è betg independenta dal connex cun il cuntegn. Anzi, ella è fitg specifica a la domena e perquai era mo transferibla dad ina domena a tschella cun grondas resalvas (Lindauer & Schmelletin, 2011; Sodian, 2008). Sin basa da la scienza davart la chapibilitad da texts e la chapientscha da texts (Wellenreuther, 2008; Schmolzer-Eibinger & Egger, 2012; Akhondi, 2010) avain nus sviluprà in model euristic davart la difficultad da texts e sin basa da quest model in raquint tematic davart il

temp paleolitic cun trais nivels da lectura, mintgamai per tudestg e per rumantsch. Sco facturs per la difficultad dal text vegn considerata la lunghezza dal text, la sintaxa ed il vocabulari. La basa per formular il raquint tematic è stà il text tudestg. Perquai che texts rumantschs cuntegnan autras difficultads (p.ex. plets cumponids), èn il model euristic ed il raquint tematic vegnids adattads correspondentamain. Per mantegnair il fluss da l'istorgia èn vegnidas integradas fanestras da text cun ulteriuras infurmaziuns tematicas sin basa dal model euristic.

1.1 Situaziun da partenza linguistica

La Svizra è in pajais quadriling cun in sistem da scola federalistic. Il rumantsch è ina lingua minoritara en il chantun Grischun, ils meds d'instrucziun èn per il pli translats dal tudestg (cun excepziun dals meds d'instrucziun da lingua).

Illustraziun 1

Situaziun da las linguas en Svizra (Funtauna: Uffizi federal da statistica)

Ripartizione delle lingue ufficiali in Svizzera (2000)



Ufficio federale di statistica, www.bfs.admin.ch; censimento 2000

En il rumantsch datti tschintg linguas da scrittira creschidas istoricamain, numnadas idioms. Auter che tar il diever usità da la noziun *idiom* en la linguistica, na vegn ella betg duvrada en quest cas per in dialect, ma per las linguas regionalas standardisadas. En pli datti dapi l'onn 1982 ina lingua da scrittira cuminaivla, il rumantsch grischun ch'è renconuschì uffizialmain sin nivel federal e chantunal (GR).

Las scolaras ed ils scolar participads al studi derivan da trais regiuns idiomaticas differentas. Da translatar ils texts en ils singuls idioms è stada in'incumbensa complexa, perquai ch'ils idioms locals dovran savens plets fitg differentes per la medema chausa. Sco exempel saja numnà qua il pled izun (Heidelbeere). Gia il pled

da basa tudestg Beere è cumplettaimain different en ils idioms vallader (pommaraida) e sursilvan (cocla). Il pled izun (Heidelbeere) ch'è identifitgabel en il tudestg senz'auter sco cumpositum dal pled "Beere", na vegn ni en vallader (uzuns) ni en sursilvan (izuns) furmà dad in cumpositum da "Beere". Sco en quest exempel datti en las variantas scrittis dal rumantsch blers plets che na pon betg vegnir deducids da terms generals sco en il tudestg.

2

Design metodic

2.1

Prova statistica e raquint tematic

Per intercurir las ipotesas han ins duvrà in design quasi-experimental en sis classes dal chantun Schaffusa e nov classes dal chantun Grischun cun totalmain 221 uffants (108 mattas e 113 mats) dal quart stgalim primar.

Ins ha pudì utilisar 204 unitads da datas. In'unitad da datas consista da las datotecas audio da las duas intervistas (pretest, posttest), in disegn e duas fotografias da la tecnica da reconstrucziun (pretest, posttest). 17 unitads da datas han stuì vegnir exclusas, perquai ch'ellas n'eran betg complettas u avevan mancanzas tecnicas.

Per la retschertga en il chantun Schaffusa è vegnidas tschernidas scolas cun ina cumpart mesauna fin auta dad uffants che n'han betg tudestg – pia la lingua d'instrucziun - sco lingua materna. En il chantun Grischun han ins tschernì scolas en la regiun Surselva ed en l'Engiadina Bassa/Val Müstair. In criteri da seleccziun spezial n'è betg stà necessari, gia che tut las scolas en questas regiuns instrueschan da l'entschatta da scola fin a la quarta classa mo rumantsch.¹

Per il tractament è vegnids stgaffids texts tematics en trais nivels da difficultad differentes en tudestg e per il chantun

Grischun èn quels vegnids transponids ed adattads en ils idioms rumantschs sursilvan e vallader resp. en la lingua da standard rumantsch grischun per la Val Müstair.

1 | En autras scolas en la regiun rumantscha datti classas bilinguas a partir da l'emprima classa primara. En general datti en las regiuns rumantschas ina cumpart mesauna fin auta dad uffants che han in'otra lingua materna che la lingua d'instrucziun.

2.2

Andament da la retschertga ed enquista tar ils uffants

La raccolta da datas è vegnida fatga senza gronds intervals da temp tranter il pretest, il tractament ed il posttest ed è stada organisada suandantamain:

Tabella 1

Schema da l'andament da la retschertga

Gruppa experimental		
Pretest	Intervenziun 25-30 min leger il text tematic	Posttest
Intervista davart il dissegn e la metoda da reconstrucziun	Lecturas e lecturs fermes legian il text tematic da nivel aut	Intervista davart il text e la metoda da reconstrucziun
	Lecturas e lecturs mesauns legian il text tematic da nivel mesaun	
	Lecturas e lecturs flaivels legian il text tematic da nivel bass	
Gruppa da controlla		
Pretest	Interveziun	Posttest
Intervista davart il dissegn e la metoda da reconstrucziun	Text tematic da nivel aut	Intervista davart il text e la metoda da reconstrucziun

A l'entschatta han ils uffants stuì far in dissegn davart il tema temp da crap. La finamira da quest'incumbensa era dad activar la presavida. Il pretest è sa cumponì da la pre-intervista suandanta e da la metoda da reconstrucziun. En l'intervista èn ils uffants l'emprim vegnids interrogads davart lur dissegn e suenter han els stuì ordinar pensond ad auta vusch (cf. Fromann, 2005) cartinas da maletgs e plets davart il tema temp da crap tenor las trais cate-

gorias "tutgà tar il temp da crap", "na tutgà betg tar il temp da crap" e "noziun betg clera u nunenconuschenta" (cf. tabella 2) e construir cun quellas ina structura (metoda da reconstrucziun). En in segund pass han ils uffants stuì metter las cartinas en ina relaziun ina cun tschella e sche pussaivel motivar lur structura. En il posttest han ins interrogà l'emprim ils uffants davart l'istorgia en il text. Suenter han els puspè stuì posiziunar cartinas da

maletg e plets cun explitgar, perquai ch'els han fatg eventuals midaments en confrunt cun l'emprima structura.

Per garantir ina retschertga unifitgada en ils chantuns Schaffusa e Grischn han ins fatg in mussavia per las intervistas cun indicaziuns precisas davart il manar l'intervista. Las intervistadras ed ils intervistaders èn vegnids scolads l'emprim per lur incumbensa.

2.3

Metodas d'evaluaziun

a) Evaluaziun da las intervistas cun tecnica da reconstrucziun (posiziunament da las cartinas)

Ils uffants han pudì metter ils singuls items en trais champs. Or d'in punct da vista istoric ha mintga decisiun pudì vegnir validada sco correcta u fallada. Uschia datti per mintga item sis pussaivladads da categorisaziun (cf. tabella 2).

Tabella 2

Schema d'evaluaziun ed interpretaziun psico-cognitiva sin basa da la posiziun dals items. Ils champs alvs vegnan validads sco corrects, ils colorads pli stgir sco fallads. Ils champs colorads pli cler vegnan tut tenor la scala validads autramain (cf. sutvart).

	Tutgà tar il temp da crap	Na tutgà betg tar il temp da crap	Jau na sai betg
Evaluaziun istorica dals items	Items corrects (dumber = 22)	"Tutgada": savair correct resp. element da savida lià correctamain en la memoria	"Sbagliada": savair fallà resp. element da savida lià al fauss lieu en la memoria
	Items fallads (dumber = 6)	"Fauss alarm": concept fallà, element da savida lià al fauss lieu en la memoria	"Refusa correcta": savair correct resp. element da savida correct betg lià en la memoria
			„Malsegirtad faussa": "nunsavida" fallada, element da savida betg anc lià resp. element da savida enconuschent e malsegirtad davart l'appartegnientscha
			"Malsegirtad correcta": "nunsavida" correcta, element da savida betg anc lià, resp. element da savida enconuschent e malsegirtad davart l'appartegnientscha

Per mintga scolara e per mintga scolar poschia vegnir calculada ina valur, la quala reflectescha sia prestaziun. Ina valur pli auta stat per ina meglra prestaziun:

Scala per la tecnica da reconstrucziun = tutgada + refusa correcta = 28 – sbagliada – fauss alarm – malsegirtad faussa – malsegirtad correcta

La scala presuppona ch'ils items piazzads sut "Jau na sai betg" n'èn betg enconuschents a las scolaras ed als scolars ed els vegnan quintads sco sbagls, gia che las probandas ed ils probands na san betg inditgar l'appartegnientscha explicita-main.² Per las valurs da la scala è vegnida calculada suenter in'analisa da varianza da dus facturs cun ina repetiziun da la mesiraziun (variablas independentas: gruppa experimental, nivel da lectura) per testar las ipotesas dal studi.

Cun il posiziunament da las cartinas po l'organisaziun cognitiva da la savida da las probandas ed ils probands vegnir giuditgà. Questa analisa sa limitescha dentant a la simpla attribuziun dals items tar il tema istoric temp paleolitic. Las concepziuns da las scolaras ed ils scolars davart il temp paleolitic na pon betg vegnir giuditgadas cun analizar il posiziunament

da las cartinas.³ Dus puncts pon vegnir evaluads a basa da la tecnica da reconstrucziun: sco emprim sche in term è enconuschent als emprendiders e sco segund, sche els mettan quel a moda cognitiva en relaziun cun il cuntegn dal temp paleolitic. Cun l'analisa dal posiziunament da las cartinas vegn l'emprender perquai mesirà sin ses nivel da basa, q.v.d. la furmaziun da colliaziuns associativas cun singuls cuntegns da la memoria.

b) Evaluaziun da las intervistas a maun da la matrix dal conceptual change

Las evaluaziuns da las intervistas èn vegnidas fatgas a maun da datotecas audio cun inditgar la validaziun directamain en il nivel correspondent dal raster. La definiziun dals quatter nivels è vegnida fatga en in'emprima fasa a moda euristica resp. specifica a la domena ed è vegnida perfeziunada pass per pass tras duas retschertgas da pilot. Or da quellas han ins operazionalisà exempels da referenza.

3 | Ina retschertga cuntinuanta che po testar pli precis la chapientscha dal cuntegn da las scolaras ed ils scolars è vegnida considerata en il rom dal studi, ma n'è betg vegnida realisada pervi dals custs. L'idea da basa fiss da betg mo analizar la categorisaziun (tutgà tar, na tutgà betg tar), ma da resguardar surtut il posiziunament dals terms in sper l'auter.

2 | In'ulteriura scala che resguarda ils items posiziunads sut „Jau na sai betg“ è era vegnida testada. Gia che ils resultats da las duas scalas sa differenzieschan mo minimamain, vegn mo descritta la scala "tecnica da reconstrucziun".

Tabella 3

Model che mussa ils nivels da concept da la savida istorica (e raster d'evaluaziun cun ils quatter stgalims da concept definids (suenter Engelen et al., 2002 e Moser et al., 2007)

D2	Concepts dal pensar istoric en il senn d'ina integraziun cognitiva da la reconstrucziun e da la deconstrucziun Nivel conceptual 2: relaziun tar oz, cumparegliaziun oz – pli baud
D1	Concepts preliminars dal pensar istoric en il senn dad access conceptuels a l'istorgia, classificaziun temporal Nivel conceptual 1: Concept cumplet tar terms concrets, concept approximativ tar terms abstracts
C2	Concepts resp. concepziuns dals scolars amplifitgabels (associaziuns etc. da natira complexa) Associaziuns 2: categorias, noziuns generalas, noziuns abstractas, construcziuns da frusas cun conjuncziuns (damai, perquai che, schebain che etc.)
C1	Concepts resp. concepziuns dals scolars amplifitgabels (associaziuns etc. da natira simpla) Associaziuns 1: categorias, noziuns generalas, noziuns abstractas, frusas declarativas finalas (per...); funcziuns cun verb adequat
B2	Concepts resp. concepziuns che na pon betg vegnir attribuids a la scienza istorica (enumeraziuns differenziadas, teorias da mintgadi, analogias cun il preschent, supposiziuns) Terms specifics e concrets: verbs differenziads (duvrar, abitar); denominaziun da > 2 terms; supposiziun concreta
B1	Elements da savida resp. concepziuns betg liads correctamain a la scienza istorica (enumeraziuns simplas, teorias da mintgadi, analogias cun il preschent, supposiziuns) Terms generalas e concrets; verb avair; supposiziuns vagas
A	Nagin concept istoric correct Declaraziuns generalas or da l'ambient da viver dals uffants

Las datas rimnadas en las intervistas servan a mesurar ils midaments sin il nivel conceptual. Decleraziuns sin ils nivels conceptuels pli bass eran pli frequentas che decleraziuns sin ils nivels conceptuels pli auts. Per dar a decleraziuns dals nivels conceptuels pli auts dapli paisa, han ins validà quellas cun in index e dà a quellas ina valur pli auta. Sche in uffant ha per exempel fatg en l'intervista mo decleraziuns sin il nivel conceptual D2, survegn el ina valur pli auta ch'in uffant che fa tuttina bleras de-

cleraziuns sin il nivel conceptual B1.

c) Interrater-reliability

Per quantifitgar la *interrater-reliability* èn vegnidas tschernidas quatter intervistas. Quellas èn vegnidas evaluadas separadamain ed ils resultats èn vegnids cumparegliads. Igl ha dà ina congruenza da 0.804, quai ch'ins po taxar sco bun. Questa metoda da rating serva sco basa per l'evaluaziun da las datas (cf. chapitel 2.3 a).

3

Resultats

3.1 Midaments da structuradas cognitivas tras texts tematics

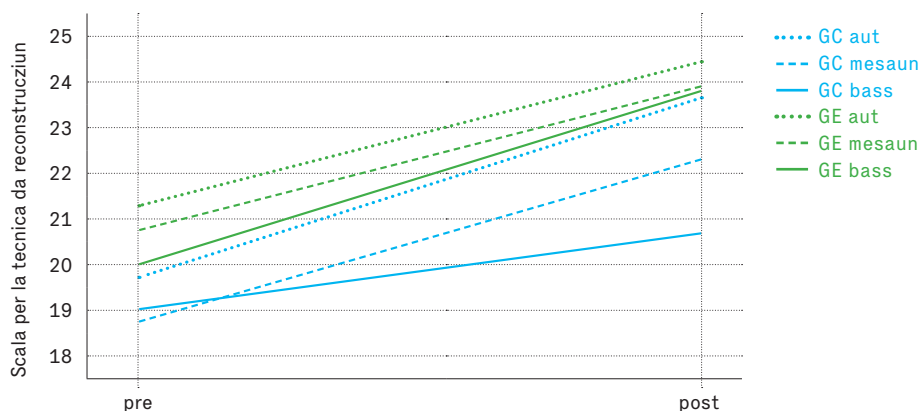
La cumparegliaziun tranter il posiziunament da las cartinas dal pre- e posttest mussa in meglierament da la prestaziun da las enconuschientschas terminologicas e quai tant per la gruppa experimentalada sco era per la gruppa da controlla (cf. illustraziun 2). Perquai che bunamain tut las scolaras e tut ils scolars han cuntanschì in meglier resultat cun la tecnica da reconstrucziun e perquai che quest augment da la prestaziun è marcant, resultan

gronds effects tar la cumparegliaziun tranter ils pre- ed ils posttests da las duas gruppas: gruppa experimentalada $d=1.38$, gruppa da controlla $d=1.07$. Ina gronda part da l'augment da la prestaziun po vegnir declerà cun il fatg ch'ils items ch'èn vegnids posiziunads tar la decleraziun "Jau na sai betg" èn pir vegnids integrads correctamain in il posttest. Cun separar tut las gruppas tenor nivel da lectura, pon ins era constatar in augment da prestaziun tar tuttas.

Per las enconuschientschas terminologicas sa laschan demussar effects principals signifitgants per la gruppa experi-

Illustraziun 2

Midament da las enconuschientschas terminologicas tar la gruppa experimentalada (GE) e la gruppa da controlla (GC)



mentalada, ($F(1/205)=13.104$, $p<0.001$), il nivel da lectura ($F(2/205)=3.656$, $p=0.027$) ed il mument dal test ($F(1/205)=330.242$, $p<0.001$). Era sa lascha demussar in'interacziun tripla signifitganta da la gruppa experimentalada, dal nivel da lectura e dal mument dal test ($F(2/205)=5.133$, $p=0.007$).

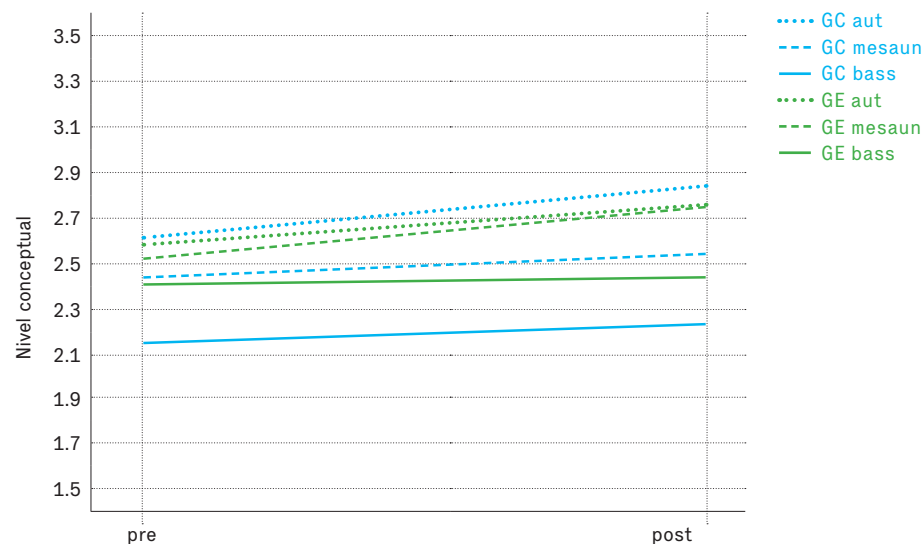
Sco differenza principala tranter la gruppa experimentalada e la gruppa da controlla sa fa valair l'augment da prestaziun tar las lecturas ed ils lecturs flaivels: tar in test da leger adattà (nivel bass) da la gruppa experimentalada cuntanschan las lecturas ed ils lecturs flaivels in augment da prestaziun pli grond che ils mesauns u fermes. Tar la gruppa da controlla è l'augment da prestaziun da las lecturas ed ils lecturs flaivels marcantamain pli pitschen che tar las autras duas gruppas da prestaziun.

3.2 Midaments en il savair conceptual tras texts tematics

Per il savair conceptual sa laschan demussar effects principals signifitgants per il nivel da lectura ($F(2/192)=14.856$, $p<0.001$) ed il mument dal test ($F(1/205)=22.425$, $p<0.001$). Percunter na cuntanscha nagina da las interacziuns pussaivlas il nivel significativ necessari. L'illustraziun 3 sutvart mussa en furma grafica il svilup tranter il pre- ed il posttest cumparegliand la gruppa experimentalada e la gruppa da controlla.

Illustraziun 3

Midament dal savair conceptual tar la gruppa experimentalada (GE) e la gruppa da controlla (GC)



L'augment da las valurs media en la tabella 4 mussa ch'ils uffants examinads han en total in (lev) augment da lur chapientscha istorica. Evident è era ch'ils nivels da prestaziun da lectura (aut, mesaun bass) sa mussan cleramain. La tabella 4 descriva tut ils uffants; la tabella 5 differenziescha tranter las gruppas linguisticas. L'effect da l'augment è dentant plitost flavel e na conferma betg nossa ipotesa da l'entschatta. I dat mo pitschnas differenzas tranter la gruppa experimental e la gruppa da controlla. Interessant èn però il detagls: Tar la *gruppa experimental* (texts da lectura adattads) cuntanschan las lecturas ed ils lecturs mesauns il pli grond augment, entant che quai vala en la *gruppa da controlla* (texts da lectura difficils per tuts) per las lecturas ed ils lecturs fermes.

Per las lecturas ed il lecturs flavels na fai apparentamain nagina differenza tar la construcziun da la chapientscha concep-

tuala d'istorgia, sch'els legian in text tematic difficil u facil; las lecturas ed ils lecturs flavels n'augmentan betg l'empredi. Tar la construcziun da las enconuschientschas terminologicas sa mussa in auter maletg. Ils resultats descritti survart vegnan anc accentuads, sch'ins analysescha las gruppas linguisticas separadamain. Dals uffants da lingua tudestga mussan en la gruppa experimental (cun texts adattads) las lecturas ed ils lecturs mesauns in augment da prestaziun signifitgant tranter ils pre- ed ils posttests. En la gruppa da controlla percenter han las lecturas ed ils lecturs fermes l'augment il pli grond. Tar ils uffants da lingua rumantscha mussan en tut las duas gruppas las lecturas ed ils lecturs fermes il pli grond augment da prestaziun. Questa differenza marcanta tranter las gruppas linguisticas n'è al mument betg explitgabl.⁴

4 | Per sclerir questa dumonda suandain nus al mument trais pussaivladads:

1. *Effect da la prova statistica*: nus controllain, sche nus avain en la prova statistica rumantscha effects da lingua particulars (p.ex. uffants da lingua tudestga ch'èn vegnids scolarisads per rumantsch). Emprimas clerificaziuns mussan, che questa explicaziun ha ina probabilitad pitschna.
2. *Effects dal text*: Nus controllain anc ina giada, sche las difficultads dal text tranter il rumantsch (sursilvan, vallader, rumantsch grischun) e tudestg n'erain eventualmain betg identicas. Latiers ston ins resguardar che ils uffants enconu-

schevan en las intervistas adina puspè in term tecnic tudestg (sco "Burggraben"), ma ch'els n'enconuschevan betg il term rumantsch correspondent. Quest effect ans para il pli probabel.

3. *Effect dal rating*: Il model da rating da la midada conceptuala ed ils exempels da referenza eran tudestgs ed èn vegnids construids a basa dad exempels tudestgs. Las intervistas cun ils uffants rumantschs èn pia vegnids validats tenor ratings da lingua tudestga. Gia che las codadras ed ils codaders rumantschs san tuts fitg bain tudestg, para quest effect fitg pauc probabel.

Tabella 4

Differenzas da las valurs media tranter ils pre- ed ils posttests en regard als midaments dals nivels conceptuels (valor 1 = in nivel conceptual da totalmain 6) (NL = nivel da lectura)

Midaments ⁵	Tut ils nivels	NL bass	NL mesaun	NL aut
Gruppa experimental	0.14	0.03	0.22	0.17
Gruppa da controlla	0.15	0.08	0.10	0.22

Tabella 5

Differenzas da las valurs media tranter ils pre- ed ils posttests en regard als midaments dals nivels conceptuels (valor 1 = in nivel conceptual da 6) tenor gruppa linguistica (NL = nivel da lectura)

Midaments	Uffants da lingua tudestga				Uffants da lingua rumantscha			
	Tut ils uffants	NL bass	NL mesaun	NL aut	Tut ils uffants	NL bass	NL mesaun	NL aut
Gruppa experimental	0.18	0.04	0.50	0.09	0.10	0.02	-0.02	0.27
Gruppa da controlla	0.16	-0.01	0.08	0.27	0.14	0.11	0.09	0.17

5 | Las valurs reflecteschan il midament relativ per ils nivels definids (A, B1, B2, C1, C2, D1, D2). Ina valur da differenza dad 1 signifitgescha lasuenter l'augment dad in entir nivel (p. ex. da C2 a D1). La scala va da 0 a 6.

4

Interpretaziun dals resultats e lur significaziun per la producziun ed il diever da texts tematics

Tenor noss setting han texts differenziads a l'intern apparentamain il potenzial dad approfondar e precisar las enconuschientschas terminologicas tar tut ils nivels da lectura. En cumparegliaziun cun las lecturas ed ils lecturs flaivels da la grupp da controlla, profitan las lecturas ed ils lecturs flaivels da la grupp experimental cun emperder terminologia surproporzionalmain, sch'els pon tractar in text tematic adattà a lur nivel da lectura (vul dir in text pli simpel).

Igl è da remartgar criticamain ch'ins n'ha betg pudì demussar in augment signifitgant da las enconuschientschas terminologicas cumparegliond la grupp experimental cun la grupp da controlla. Tut ils nivels da lectura da la grupp experimental profitan concernent las enconuschientschas terminologicas en ina dimensiun cumparegliabla. Quai pudess esser in effect dal fatg ch'ins aveva ina grupp da probands memia pitschna cun in nivel da lectura bass per demussar ina differenza statisticamain signifitganta tranter la grupp experimental e la grupp da controlla.

Concernent la chapientscha conceptuala, na pon ins betg constatar ina differenza marcanta tranter la grupp experimental e la grupp da controlla; las differenzas correspundan a las spetgas

(cf. tabella 4). Uschia na po l'ipotesa 1 betg vegnir cumprovada: quella ipotesa ha supponì che materialias d'instruir e d'emperder differenziadas a l'intern augmentan las prestaziuns da chapir cuntegns (categorias, associaziuns, concepts) da lecturas e lecturs fermes e flaivels. A l'intern da la grupp experimental èn las differenzas da la chapientscha conceptuala pitschnas tranter lecturas e lecturs mesauns e fermes. Entaifer questas pitschnas differenzas tranter il pre- ed il posttest crescha la chapientscha conceptuala tar lecturas e lecturs mesauns e fermes il pli fitg.

L'ipotesa 2 po vegnir confermada concernent las enconuschientschas terminologicas: lecturas e lecturs pli flaivels da la grupp experimental mussan en confrunt cun quels da la grupp da controlla cleramain in pli grond augment da savida, sch'els survegnan texts adattads a lur nivel da leger. Per l'acquisiziun da la chapientscha conceptuala d'istorgia na porschan, tenor noss resultats, texts tematics adattads al nivel da lectura a las lecturas ed als lecturs pli flaivels da la grupp experimental però betg in fundament bun avunda. Nossa segunda ipotesa po cun quai mo vegnir confermada concernent las enconuschientschas terminologicas.

Ils effects flaivels concernent la midada dal nivel conceptual pon esser liads al fatg che l'intervenziun cun in text da lectura, che duai vegnir legì entaifer 20 fin 30 minutas, offra probablmain memia pouca stimulaziun per ina midada conceptuala. Igl è probabel ch'ina midada conceptuala profunda capita pir, sch'ins sustegna la lectura d'in raquint tematic cun dumondas davart il text u ina runda d'elavuraziun. Quai corresponda als resultats publitgads che mussan preconcepts relativamain stabils (Möller, 1999) che dovran intervenziuns pli persistentas. Per verifitgar questa ipotesa stuess ins far (sco che quai è vegnì fatg en las retschertgas da las scienzas natiralas numnadas) ina intervenziun en l'instrucziun (elavuraziun ed approfondaziun) che sustegna la differenziaziun interna da la lectura dal text tematic. Nossa retschertga ha mo testà l'effect dad in text tematic adattà senza integrar l'effect da l'elavuraziun ed approfondaziun en l'instrucziun. Tuttina pon ins far constataziuns per la producziun da materialias d'instrucziun ed era per l'organisaziun da l'instrucziun. D'ina vart vali apparentamain la paina da porscher a lecturas e lecturs flaivels tar temas specificis texts tematics simplifitgads. Quests pussibiliteschan als lecturs betg rutinads da stgaffir enconuschientschas terminologicas e raits da terminologia da basa che pon vegnir duvrads sco fundament per l'elavuraziun ed approfondaziun da l'instrucziun. Texts tematics difficils, ch'ins inscuntra tenor nossa expertisa savens en meds d'instrucziun, na paran betg da far gronds problems a buns fin fitg buns lecturs. Qua è probablmain l'instrucziun

differenziada a l'intern pli effectiva che in'intervenziun sur il text. Per lecturas e lecturs flaivels paran texts difficils però da preschentar in obstachel veritabel. A l'acquisiziun da raits terminologicas cun texts tematics simplifitgads sto probablmain suandar in'instrucziun che elavurescha ed approfondescha il legì per pudair construir ils concepts istorics correspundents. Er il plan d'instrucziun 21 (D-EDK, 2014) en sia versiun dal november 2014 mussa l'impurtanza da la relaziun tranter ils temas e la lingua.

Nossas datas vegnan anc analisadas vinavant, surtut a regard ils effects dals differents nivels da text en il rumantsch. Plinavant ston ins far vinavant perscrutaziuns davart ils arranschaments d'emperder, che testeschan l'effect da texts tematics differenziads a l'intern en cumbinaziun cun purschidas d'approfondaziun ed elavuraziun per differents nivels da prestaziun.

La différenciation interne de textes factuels comme moyen de promouvoir l'acquisition de savoirs et de concepts relevant de domaines spécifiques

Notes de synthèse

—

Andreas Imhof, Eva Engeli, Markus Kübler

1

Fondement théorique et hypothèse

Le projet *Différenciation interne de textes factuels comme moyen de promouvoir l'acquisition de savoirs et de concepts relevant de domaines spécifiques* est fondé sur le constat que les outils pédagogiques (et les textes factuels qu'ils contiennent) constituent encore le cœur de l'enseignement (Oelkers, 2010, p. 39 ; Kahlert, 2010) mais ne permettent plus de répondre aux exigences d'un enseignement individualisé (Moser Opitz, 2010). Ce propos doit être compris à la lumière des exigences d'une école inclusive et des conséquences d'ordre organisationnel résultant de la différenciation interne à la classe. Plusieurs études dans différents domaines (notamment Lässig, 2010 ; Leisen, 2006 ; von Borries, 2010) confirment que les outils pédagogiques sont en général beaucoup trop difficiles pour des lectrices et lecteurs de niveau moyen. Il est plausible que des enfants de langue étrangère soient particulièrement concernés par ce phénomène (Moser, 2010) et qu'il soit nécessaire d'adapter les textes factuels à leur niveau si l'on se fonde sur la présomption que le matériel pédagogique différencié est en mesure d'améliorer la capacité de compréhension des contenus (notions, catégories, associations, concepts).

A partir de ces constats, nous formulons deux hypothèses :

1ère hypothèse : le matériel pédagogique différencié est en mesure d'améliorer la capacité de compréhension des contenus (notions, catégories, associations, concepts), que le niveau des lectrices et lecteurs soit élevé ou faible.

2e hypothèse : la différenciation interne profite avant tout aux lectrices et lecteurs de niveau faible.

Un *indicateur* de l'exactitude de ces hypothèses serait une augmentation des acquis pédagogiques plus importante au sein de groupes d'apprentissage bénéficiant de textes factuels différenciés qu'au sein de groupes d'apprentissage n'en bénéficiant pas. Par augmentation des acquis pédagogiques, nous entendons l'acquisition de concepts historiques. Ce faisant, nous partons des principes constructivistes selon lesquels les apprenant-e-s doivent élaborer eux-mêmes de nouveaux savoirs sur la base de leurs acquis (p. ex. Duit & Widodo, 2004). Ceci signifie que l'environnement pédagogique doit tenir compte du savoir préexistant des élèves, et en particulier des conceptions divergeant du point de vue scientifique. Ne pas tenir compte de ces conceptions est consi-

déré comme une des causes principales des difficultés d'apprentissage (p. ex. Bransford, Brown & Cocking 2000 ; Duit, 1995 ; Treagust, Duit & Niewsandt, 2000). La modélisation des compétences de Hardy et al. (2010) prend comme point de départ une description générale du changement conceptuel (*conceptual change*, Möller, 2015 ; Sodian, 2008 ; Carey & Spelke, 1994 ; Carey, 2000 ; Carey, 2011 ; Möller, 2012 ; Anohina et al., 2009) depuis les concepts naïfs jusqu'aux conceptions scientifiquement validées en passant par les conceptions intermédiaires. D'une façon comparable, l'acquisition de concepts spécifiques à l'apprentissage historique est aussi décrite dans le projet *Pensée historique* (Bietenhader & Kübler, 2012). Se fondant sur les travaux préliminaires d'Engelen et al. (2002) et de Hardy et al. (2010) et sur les découvertes scientifiques concernant le Paléolithique (Höhneisen, 1998 ; Parzinger, 2014 ; Osterwalder & Robert, 1990), l'équipe de projet a décrit une série de concepts historiques de l'âge de pierre, nonobstant le fait que l'Histoire constitue un domaine mal défini (« ill-defined ») (Günther-Arndt, 2006).

Les vastes connaissances dans le domaine de la recherche sur la lisibilité n'ont eu jusqu'ici que peu d'impact sur la production d'outils pédagogiques (Leitner, 2008, p. 708 ; Kahlert, 2010, p. 49s.). La compréhensibilité d'un texte au sens de la logique de texte, du vocabulaire factuel et du rapport au monde n'est pas sans lien avec le rapport au contenu, mais relève au contraire essentiellement du domaine spécifique. C'est pourquoi elle ne peut être transférée d'une spécialité à l'autre qu'à certaines conditions (Lindauer & Schmellettin,

2011 ; Sodian, 2008). Sur la base des recherches portant sur la compréhensibilité et la compréhension de texte (Wellenreuther, 2008 ; Schmolzer-Eibinger & Egger, 2012 ; Akhondi, 2010), nous avons développé un modèle heuristique relatif à la difficulté des textes et, sur la base de ce modèle, rédigé un récit factuel sur le thème du Paléolithique avec trois niveaux de lecture, en allemand et en romanche respectivement. Les facteurs pris en compte pour la difficulté des textes sont la longueur, le niveau de syntaxe et le niveau de vocabulaire. Pour la formulation du récit factuel, la version allemande a servi de base. La difficulté du texte étant parfois de nature différente en romanche (p. ex. les mots composés), le modèle heuristique et le récit factuel ont été adaptés en conséquence. Afin de préserver la fluidité du récit, des fenêtres de texte contenant des informations factuelles supplémentaires ont été intégrées sur la base du modèle heuristique.

1.1

Situation linguistique initiale

La Suisse est un pays quadrilingue possédant un système scolaire fédéral. Le romanche est une langue minoritaire du canton des Grisons où les outils pédagogiques utilisés sont en principe traduits de l'allemand (à l'exception des manuels d'apprentissage de langue romanche).

Figure 1
Situation linguistique en Suisse (source : Office fédéral de la statistique)



Il existe historiquement cinq différentes langues écrites en romanche. Elles sont appelées idiomes. Contrairement à l'acception habituelle en linguistique du terme *idiome*, il est ici utilisé pour qualifier les langues écrites régionales standardisées et non un dialecte. Par ailleurs, il existe depuis 1982 une variante écrite commune, le « rumantsch grischun » (romanche grison), reconnue officiellement par la Confédération et le canton des Grisons.

Les élèves romanchophones participant à l'étude sont issus de trois régions aux idiomes différents. La traduction des textes dans chaque idiome s'est révélée complexe, notamment du fait que les idiomes locaux emploient des mots souvent très différents pour qualifier la même chose. Nous mentionnerons ici le terme

« Heidelbeere » (myrtille) à titre d'exemple. Le terme de base allemand *Beere* (baie) est déjà totalement différent dans les deux idiomes vallader (pommaraida) et sursilvan (coclas). Le mot *Heidelbeere* (myrtille) en allemand est à l'évidence une différenciation du mot *Beere* (baie) et ne se laisse traduire ni en vallader (uzuns) ni en sursilvan (izuns) en partant du radical *Beere*. Tout comme dans cet exemple, il existe dans chaque variante écrite du romanche bon nombre de mots qui ne peuvent dériver d'un terme générique comme en allemand.

2

Modèle méthodologique

2.1 Echantillon et récit factuel

Afin d'étudier les hypothèses, un modèle quasi expérimental a été mis en place et appliqué dans six classes du canton de Schaffhouse et neuf classes du canton des Grisons comptant en tout 221 enfants (108 filles et 113 garçons) de quatrième primaire.

204 dossiers ont pu être évalués, ceux-ci étant constitués des enregistrements audio des deux interviews (pré-test, post-test), d'un dessin et de deux photographies de la technique de reconstruction SLT (structure-laying technique) (pré-test, post-test). 17 dossiers ont dû être éliminés car ils étaient incomplets ou présentaient des défauts techniques.

Pour l'étude dans le canton de Schaffhouse, le choix s'est porté sur des établissements scolaires présentant une proportion moyenne à élevée d'enfants dont la langue maternelle n'est pas l'allemand, donc une langue distincte de la langue de l'école. Dans le canton des Grisons, le choix s'est porté sur des établissements scolaires des régions de la Surselva (sursilvan), de Basse-Engadine (vallader) et de la Val Müstair (rumantsch grischun). Aucun critère de sélection particulier n'était requis, à l'exception du fait que tous les établissements scolaires de ces

régions pratiquent un enseignement intégralement en romanche depuis le début de la scolarisation jusqu'en quatrième année primaire.¹

Lors du *treatment*, les textes factuels ont été rédigés en allemand dans trois niveaux de difficulté différents et traduits et adaptés pour le canton des Grisons dans les idiomes sursilvan et vallader, resp. dans la langue standardisée rumantsch grischun pour la Val Müstair.

1 | Ceci contrairement à d'autres communes scolaires dans la région linguistique romanche qui pratiquent un enseignement bilingue depuis la première année. Dans toutes les régions romanches en général, on enregistre une proportion moyenne à élevée d'enfants dont la langue maternelle ne correspond pas à la langue de l'école.

2.2 Déroulement de l'étude et interrogation des enfants

La collecte des données a été effectuée en peu de temps à l'aide d'un pré-test, du *treatment* et d'un post-test selon le schéma suivant :

Tableau 1
Schéma du déroulement de l'étude

Groupe expérimental		
Pré-test	Intervention 25-30 min. lecture du texte factuel	Post-test
Interview sur le dessin et la méthode de reconstruction (structure-laying technique SLT)	Lectrices et lecteurs de niveau élevé Niveau élevé du texte factuel	Interview sur le texte factuel et la SLT
	Lectrices et lecteurs de niveau moyen Niveau moyen du texte factuel	
	Lectrices et lecteurs de niveau faible Niveau faible du texte factuel	
Groupe de contrôle		
Pré-test	Intervention	Post-test
Interview sur le dessin et la méthode de reconstruction	Niveau élevé du texte factuel	Interview sur le texte factuel et la SLT

Pour commencer, les enfants devaient faire un dessin sur le thème l'âge de pierre. Le but de cette tâche était d'activer le savoir pré-existant. Le pré-test comportait la pré-interview subséquente et la méthode de reconstruction. Lors de l'interview, les enfants ont d'abord été interrogés sur leur dessin, puis ils recevaient des cartes d'images et de mots sur le thème de l'âge de pierre qu'ils devaient, en réfléchissant

à haute voix, classer en trois catégories « Fait partie de l'âge de pierre », « Ne fait pas partie de l'âge de pierre » et « Notion inconnue ou peu claire » (cf. tableau 2) avant de les placer dans un certain ordre (méthode de reconstruction SLT). Dans un deuxième temps, les enfants devaient établir des liens entre les cartes, si possible en les expliquant, et placer les cartes de manière correspondante. Au cours du post-test, les

enfants ont d'abord été interrogés sur l'histoire racontée dans le texte factuel, puis ils plaçaient à nouveau les cartes d'images et de mots dans un certain ordre et étaient priés d'expliquer si possible les éventuelles modifications par rapport au premier positionnement des cartes.

Afin de garantir l'uniformité des études à Schaffhouse et dans les Grisons, un guide détaillé a été établi pour mener les interviews. Les personnes menant les interviews ont été formées à leur tâche au préalable.

2.3 Méthodes d'évaluation

a) Evaluation des interviews incluant la technique de reconstruction (positionnement des cartes)

Chaque item pouvait être placé dans trois champs différents, correspondant à une réponse juste ou fautive d'un point de vue historique. Il en résulte pour chaque item six catégorisations possibles (cf. tableau 2).

Tableau 2
Schéma d'évaluation et interprétation psycho-cognitive sur la base de la position des items. Les champs blancs sont considérés comme corrects, les colorés comme incorrects. Les champs légèrement colorés sont évalués différemment en fonction du barème (cf. plus bas).

Evaluation de l'item d'un point de vue historique	Fait partie de l'âge de pierre	Ne fait pas partie de l'âge de pierre	Ne sait pas
	Items corrects (nombre=22)	« Réussi » : savoir correct, c.-à-d. élément de savoir correctement associé dans la mémoire	« Manqué » : savoir incorrect, c.-à-d. élément de savoir incorrectement associé dans la mémoire
Items incorrects (nombre=6)	« Fausse alerte » : concept erroné, élément de savoir associé incorrectement dans la mémoire	« Rejet correct » : savoir correct, c.-à-d. élément de savoir correctement non associé	« Incertitude correcte » : « méconnaissance » correcte, élément de savoir non encore associé, c.-à-d. élément de savoir connu et incertitude quant à son appartenance

Ainsi, chaque élève peut obtenir une évaluation qui reflète sa performance, les valeurs plus élevées représentant une meilleure performance :

Barème technique de reconstruction
 = Réussi + Rejet correct = 28
 – Manqué – Fausse alerte
 – Incertitude, correcte – Incertitude, incorrecte

Le barème suppose que les items rejetés dans « Ne sait pas » ne sont pas connus des élèves et sont par conséquent considérés comme des « fautes », les personnes participant à l'étude n'étant pas en mesure d'explicitier leur appartenance.² Pour établir les valeurs du barème, une analyse bi-factorielle de la variance avec des mesures répétées a été effectuée (variables indépendantes : groupe expérimental, niveau de lecture) afin de tester les hypothèses de l'étude.

Grâce au positionnement des cartes, il est possible d'évaluer l'organisation cognitive du savoir des personnes participant à l'étude. Cette analyse se limite toutefois au simple classement des items concernant l'âge de pierre. La compréhension conceptuelle des élèves sur l'âge de pierre ne peut être évaluée par le biais de l'analyse du

positionnement des cartes.³ Deux éléments peuvent être évalués sur la base de la SLT : premièrement si une notion est connue des apprenant-e-s, et deuxièmement si elles ou ils l'associent sur le plan cognitif à l'âge de pierre. L'analyse du positionnement des cartes permet ainsi de mesurer le niveau basal de l'apprentissage, c.-à-d. la constitution d'associations avec des contenus individuels dans la mémoire.

b) Analyse des interviews au moyen de la matrice de conceptual change

Les analyses des interviews ont été réalisées à partir des données audio en consignnant l'évaluation des déclarations directement dans la grille au niveau correspondant. La définition des quatre niveaux a été réalisée au cours d'une première phase heuristique ou relevant de domaines spécifiques, puis affinée pas à pas au moyen de deux études pilote et enfin opérationnalisée à l'aide d'exemples de référence.

3 | Un examen détaillé permettant de tester plus précisément la compréhension du contenu par les élèves a été envisagé dans le cadre de cette étude mais n'a pas été poursuivi en raison de son coût. Fondamentalement, l'idée était d'analyser non seulement la catégorisation (appartient à ; n'appartient pas à) mais aussi et avant tout le positionnement des notions entre elles.

2 | Un autre barème prenant en compte les items rejetés dans « Ne sait pas » a également été testé. Comme les résultats des deux barèmes ne diffèrent que légèrement, seul le barème « Technique de reconstruction » est décrit.

Tableau 3
 Modélisation des niveaux conceptuels du savoir historique (et grille d'évaluation contenant les quatre degrés conceptuels définis (selon Engelen et al., 2002, et Moser et al., 2007)

D2	Concepts de pensée historique au sens d'intégration cognitive de la reconstruction et de la déconstruction Niveau conceptuel 2 : rapport au présent, comparaison présent-passé
D1	Concepts préexistants de pensée historique au sens d'accès conceptuels à l'histoire, classements chronologiques Niveau conceptuel 1 : concept intégral sur des notions concrètes, concept approximatif sur des notions abstraites
C2	Concepts évolutifs, resp. compréhension conceptuelle des élèves (associations etc. complexes) Associations 2 : catégories, termes génériques, abstractions ; constructions de phrases avec des conjonctions (comme, parce que, encore que etc.)
C1	Concepts évolutifs, resp. compréhension conceptuelle des élèves (associations etc. simples) Associations 1 : catégories, termes génériques, abstractions ; énonciations finales (afin de...); fonction avec verbe approprié
B2	Concepts non fondés d'un point de vue scientifique, resp. compréhension conceptuelle des élèves (énumérations différenciées, théories du quotidien, analogies au présent, suppositions) Notions spécifiques et faits concrets ; verbes différenciés (employer, habiter) ; mention de > 2 notions ; supposition concrète
B1	Éléments de savoir incorrectement associés au domaine scientifique, resp. compréhension conceptuelle des élèves (énumérations simples, théories du quotidien, analogies au présent, suppositions) Notions générales et faits concrets ; verbe avoir ; suppositions vagues
A	Aucun concept historique correct Déclarations générales issues du monde des enfants

Les données recueillies dans les interviews ont servi à mesurer les variations du niveau conceptuel. Des déclarations d'un niveau conceptuel plus faible sont apparues plus fréquemment que des déclarations d'un niveau plus élevé. Afin de donner plus de poids aux déclarations de niveau conceptuel plus élevé, celles-ci ont été pondérées à l'aide d'un indice afin d'obtenir une valeur plus élevée. Par exemple, lorsqu'un enfant ne fait que des déclarations de niveau conceptuel D2, il obtient un score plus élevé qu'un enfant qui

fait autant de déclarations de niveau conceptuel B1.

c) Interrater reliability

Afin de quantifier la *interrater reliability*, quatre interviews ont été sélectionnées, évaluées séparément et comparées entre elles. Un taux de conformité de 0.804 a été atteint, ce qui peut être considéré comme bon. Cette méthode de *rating* est aussi à la base de l'évaluation des données (cf. chap. 2.3 a).

3

Résultats

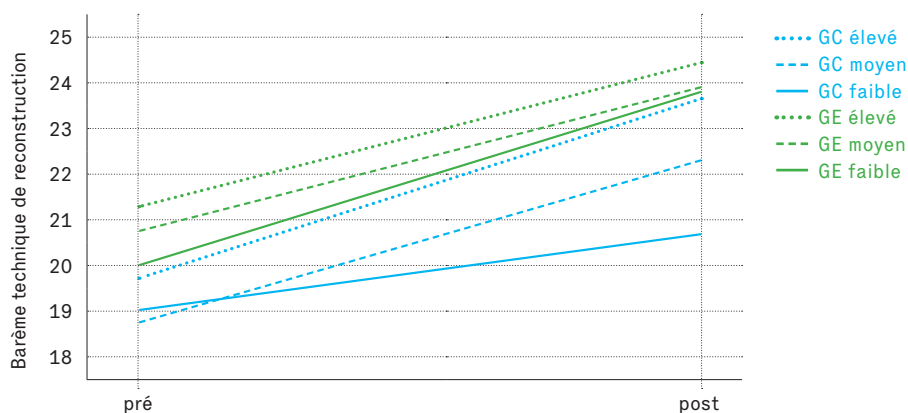
3.1 Variation des structures cognitives liées aux textes factuels

La comparaison des pré-tests et des post-tests relatifs aux positionnements de cartes (SLT) révèle une amélioration des performances quant à la compréhension des notions au sein du groupe expérimental autant qu'au sein du groupe de contrôle (cf. figure 2). Du fait que pratiquement tous les élèves obtiennent une amélioration de leur score en technique de reconstruction, et que cette augmentation de performance est relativement prononcée, il résulte de la

comparaison des pré-tests et des post-tests des deux groupes des effets très significatifs : groupe expérimental $d=1.30$, groupe de contrôle $d=1.07$. Une forte part des améliorations de performance peut être expliquée par le fait que les items faisant l'objet de la déclaration « Ne sait pas » n'ont été intégrés correctement dans le positionnement des cartes qu'au cours du post-test. Pris isolément selon le niveau de lecture, tous les groupes présentent également une amélioration des performances.

Sur le plan de la compréhension des notions, des effets principaux significatifs

Figure 2
Variation de la compréhension des notions du groupe expérimental (GE) et du groupe de contrôle (GC)



sur le groupe expérimental ($F(1/205)=13.104$, $p<0.001$), le niveau de lecture ($F(2/205)=3.656$, $p=0.027$) et le moment du test ($F(1/205)=330.242$, $p<0.001$) sont démontrés. De même, une triple interaction significative est démontrée entre le groupe expérimental, le niveau de lecture et le moment du test ($F(2/205)=5.133$, $p=0.007$).

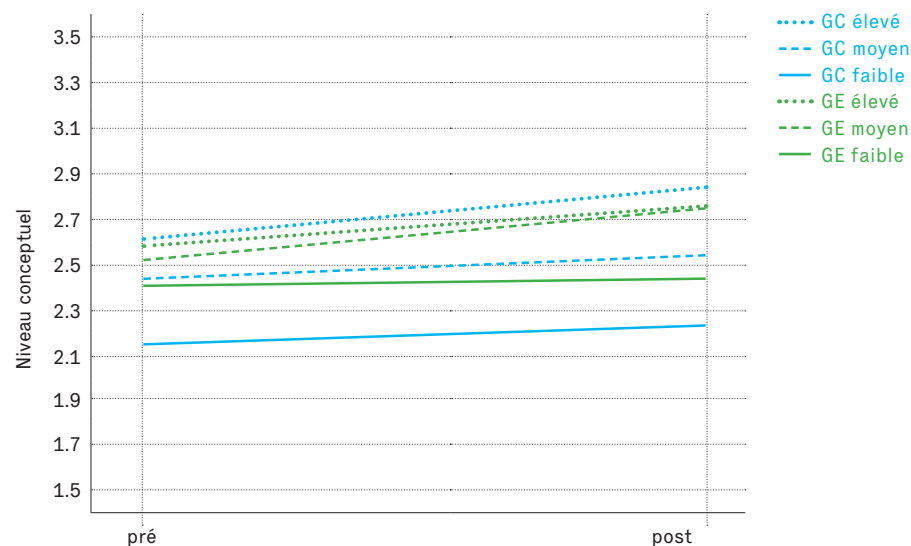
L'amélioration des performances des lectrices et lecteurs de niveau faible constitue la différence principale entre le groupe expérimental et le groupe de contrôle : en ce qui concerne le texte de lecture (de niveau faible) du groupe expérimental, les lectrices et lecteurs de niveau faible montrent une hausse des performances plus importante que ceux de niveau moyen et élevé. Dans le groupe de contrôle, la hausse des performances des lectrices et lecteurs de niveau

faible est nettement moins importante que celle des deux autres groupes.

3.2 Variations de la compréhension conceptuelle liée aux textes factuels

Sur le plan de la compréhension conceptuelle, des effets principaux significatifs sont démontrés sur le niveau de lecture ($F(2/192)=14.856$, $p<0.001$) et le moment du test ($F(1/205)=22.425$, $p<0.001$). En revanche, aucune des interactions potentielles n'atteint le niveau significatif nécessaire. La figure 3 ci-dessous présente sous forme graphique l'évolution entre le pré-test et le post-test du groupe expérimental et du groupe de contrôle comparativement.

Figure 3
Variation de la compréhension conceptuelle du groupe expérimental (GE) et du groupe de contrôle (GC)



L'augmentation des valeurs moyennes selon le tableau 4 indique que les enfants ayant participé à l'étude ont dans l'ensemble (légèrement) amélioré leur compréhension historique. Il est également évident que les niveaux de performance de lecture (élevé, moyen, faible) se dessinent clairement. Le tableau 4 prend en compte tous les enfants; le tableau 5 fait la distinction entre les groupes linguistiques. L'effet de l'augmentation est néanmoins relativement faible et ne confirme pas notre hypothèse de départ. Il n'existe que de faibles différences entre le groupe expérimental et le groupe de contrôle. Il convient toutefois de relever des détails intéressants: au sein du *groupe expérimental* (textes de lecture adaptés), les lectrices et lecteurs de niveau moyen obtiennent la plus forte augmentation, alors qu'au sein du *groupe de contrôle* (textes de lecture plus difficiles pour tous), ce sont les lectrices et lecteurs de niveau élevé.

Pour les lectrices et lecteurs de niveau faible, l'édification de la compréhension conceptuelle historique – par opposition à l'édification de la compréhension des

notions – n'est manifestement pas différente si elles ou ils lisent un texte factuel difficile ou simple: les lectrices et lecteurs de niveau faible ne présentent ici presque pas d'augmentation des acquis pédagogiques.

Les résultats ci-dessus s'accroissent encore lorsqu'on analyse les groupes linguistiques séparément. Parmi les enfants germanophones dans le groupe expérimental (avec textes adaptés), les lectrices et lecteurs de niveau moyen présentent une amélioration des performances significative entre les pré-tests et les post-tests, alors qu'au sein du groupe de contrôle, ce sont les élèves les plus forts en lecture qui présentent l'amélioration la plus importante. Parmi les lectrices et lecteurs romanchophones, ce sont les élèves les plus forts en lecture dans les deux groupes examinés qui présentent l'amélioration des performances la plus importante. Cette différence prononcée entre les groupes linguistiques ne s'explique pas à ce jour.⁴

souvent un terme spécifique allemand (comme «Burggraben»), alors que le terme romanche leur était inconnu. C'est l'effet qui nous semble le plus probable.

3. *Effet du rating*: le modèle de *rating* du changement conceptuel et les exemples de référence étaient en allemand et se basaient sur des exemples allemands. Les interviews des enfants romanches ont ainsi été évaluées sur la base de *ratings* germanophones. Toutefois, du fait que les personnes romanchophones qui encodèrent ces notations parlaient parfaitement l'allemand, cet effet nous semble plutôt peu probable.

4 | Trois hypothèses peuvent à notre sens expliquer ce phénomène:

1. *Effet de l'échantillon*: nous examinons si nous avons affaire à des effets langagiers particuliers au sein de l'échantillon partiel romanchophone (soit des enfants de langue allemande scolarisés en romanche). Selon les premiers éléments, il semblerait que cette explication soit peu probable.
2. *Effet du texte*: nous vérifions une fois encore si les difficultés entre les textes en romanche (sursilvan, vallader, rumantsch grischun) et en allemand n'étaient éventuellement pas identiques. De plus il faut noter que, au cours des interviews, les enfants romanchophones connaissaient

Tableau 4

Valeurs moyennes différentielles entre pré-tests et post-tests du point de vue de la variation du niveau conceptuel (valeur 1 = un niveau conceptuel sur 6 en tout) (NL = niveau de lecture)

Variations ⁵	Tous les niveaux	NL faible	NL moyen	NL élevé
Groupe expérimental	0.14	0.03	0.22	0.17
Groupe de contrôle	0.15	0.08	0.10	0.22

Tableau 5

Valeurs moyennes différentielles entre pré-tests et post-tests du point de vue de la variation du niveau conceptuel (valeur 1 = un niveau conceptuel sur 6 en tout) par groupe linguistique (NL = niveau de lecture)

Variations	Enfants germanophones				Enfants romanchophones			
	Tous les enfants	NL faible	NL moyen	NL élevé	Tous les enfants	NL faible	NL moyen	NL élevé
Groupe expérimental	0.18	0.04	0.50	0.09	0.10	0.02	-0.02	0.27
Groupe de contrôle	0.16	-0.01	0.08	0.27	0.14	0.11	0.09	0.17

5 | Les valeurs reflètent la variation relative des niveaux définis (A, B1, B2, C1, C2, D1, D2). Une valeur différentielle de 1 signifie par conséquent le passage au niveau supérieur (p. ex. de C2 à D1). Le barème va de 0 à 6.

4

Interprétation des résultats et de leur signification pour la production et l'utilisation de textes factuels

Selon notre *setting*, les textes différenciés internes ont manifestement le potentiel d'approfondir et de renforcer la compréhension des notions pour tous les niveaux de lecture. Ce sont en particulier les lectrices et lecteurs de niveau faible du groupe expérimental qui améliorent le plus leur niveau d'apprentissage des notions, en comparaison avec les lectrices et lecteurs de niveau faible du groupe de contrôle, lorsqu'elles et ils sont en mesure de travailler sur un thème avec un texte factuel adapté à leur niveau de lecture (c.-à-d. plus facile).

Il convient néanmoins de noter que la comparaison entre le groupe expérimental et le groupe de contrôle n'a pas présenté une différence significative de l'amélioration de la compréhension des notions. Tous les niveaux de lecture du groupe expérimental améliorent leur niveau de compréhension des notions dans une mesure comparable. Il est possible que ce soit dû au fait que le nombre de personnes participant à l'étude de niveau de lecture faible n'ait pas été suffisant pour valider une différence statistique significative entre le groupe expérimental et le groupe de contrôle.

En ce qui concerne la compréhension conceptuelle, il n'y a pas de différence significative entre le groupe expérimental et le groupe de contrôle, les différences correspondant aux attentes (cf. tableau 4). En conséquence, il n'est pas possible de confirmer l'hypothèse 1, à savoir que le matériel pédagogique différencié améliore les performances de compréhension de contenu (catégories, associations, concepts) des lectrices et lecteurs de niveau élevé ou faible. Au sein du groupe expérimental, les différences de compréhension conceptuelle entre les lectrices et lecteurs de niveau moyen à élevé sont minimes. C'est la compréhension conceptuelle des lectrices et lecteurs de niveau moyen à élevé qui augmente le plus dans l'intervalle de ces petites différences entre pré-test et post-test.

La seconde hypothèse relative à la compréhension des notions peut être confirmée : les lectrices et lecteurs du groupe expérimental avec des textes adaptés à leur niveau de lecture présentent une augmentation du savoir clairement supérieure à ceux du groupe de contrôle. D'après nos résultats, les textes factuels adaptés au niveau de lecture

n'offrent toutefois pas suffisamment de bases aux lectrices et lecteurs du groupe expérimental de niveau faible pour acquérir une compréhension conceptuelle historique. Notre seconde hypothèse ne peut ainsi être validée que pour ce qui touche à la compréhension des notions.

Les faibles effets sur la variation du niveau conceptuel sont vraisemblablement liés au fait qu'une intervention sur un texte à lire en 20 ou 30 minutes ne constitue probablement pas une stimulation suffisante pour obtenir un changement conceptuel. Il est probable qu'un changement conceptuel profond ne soit possible que lorsque la lecture d'un récit factuel est suivie de questions sur ce texte ou d'un exercice d'assimilation. Ceci correspond aux résultats publiés qui démontrent que les concepts préexistants sont relativement stables (Möller, 1999) et demandent des interventions plus durables. Afin de vérifier cette hypothèse, il faudrait (comme cela a été réalisé dans les études scientifiques mentionnées) une intervention pédagogique (assimilation et approfondissement) pour accompagner la différenciation interne de la lecture du texte factuel.

Notre étude a exclusivement testé l'impact d'un texte factuel adapté, sans inclure l'impact de l'assimilation et de l'approfondissement en classe. Malgré tout, une tendance se dégage pour la production de matériel pédagogique et pour la conduite de l'enseignement.

D'une part, lorsqu'il s'agit de thèmes factuels complexes, il vaut manifestement la peine de proposer des textes factuels simplifiés aux lectrices et lecteurs de ni-

veau faible. Ceux-ci permettent aux personnes qui lisent peu de comprendre les notions fondamentales et de constituer un réseau de notions basales qui pourront ensuite être utilisées comme socle pour l'assimilation et l'approfondissement en classe. Les textes factuels difficiles, tels que sont en général perçus les outils pédagogiques, semblent ne pas représenter d'obstacle majeur pour les lectrices et lecteurs de très bon niveau. Dans ce cas, un enseignement différencié interne à la classe est vraisemblablement plus efficace qu'une intervention sur les textes. Pour les lectrices et lecteurs de niveau faible toutefois, il semble que les textes difficiles représentent un obstacle à prendre au sérieux. L'acquisition de réseaux de notions par le biais de textes factuels doit vraisemblablement être consécutive à un enseignement d'assimilation et d'approfondissement en classe afin d'être en mesure de constituer les concepts (historiques) appropriés. Le plan d'études 21 (Lehrplan 21, D-EDK, 2014) dans sa version de novembre 2014 souligne l'importance de la relation entre objet et langage.

Nos données feront encore l'objet d'études, en particulier en ce qui concerne les effets des différents niveaux de texte en romanche. De même, des études porteront sur l'organisation de l'apprentissage testant l'impact des textes factuels différenciés en combinaison avec des offres d'approfondissement et d'assimilation adaptées à chaque niveau de performance.

Testi differenziati quale sostegno all'acquisizione di conoscenze e concetti specifici

Executive Summary

—

Andreas Imhof, Eva Engeli, Markus Kübler

1

Base teorica e formulazione

Il progetto *Testi differenziati quale sostegno all'acquisizione di conoscenze e concetti specifici* si basa sulla constatazione che gli strumenti d'insegnamento e apprendimento (e i testi in essi contenuti) continuano a rappresentare il fulcro delle lezioni (Oelkers, 2010, p. 39; Kahlert, 2010), ma non soddisfano i requisiti di un insegnamento individualizzato (Moser Opitz, 2010). La discussione ha origine dalle esigenze di una scuola inclusiva e dalle conseguenze organizzative della differenziazione interna. La conclusione secondo la quale gli strumenti didattici di regola sono troppo complessi per un livello di lettura medio è supportata da numerose ricerche in diversi settori (tra cui Lässig, 2010; Leisen, 2006; von Borries, 2010). Pare plausibile che ciò interessi in primis i bambini alloglotti (Moser, 2001) e che sussista necessità di produrre testi adeguati a un livello medio, in considerazione del fatto che materiali d'insegnamento e apprendimento differenziati potrebbero incrementare le capacità di comprensione (termini, categorie, nessi, concetti).

Queste constatazioni si traducono in due ipotesi.

1^a ipotesi: i materiali d'insegnamento e apprendimento differenziati possono incrementare le capacità di comprensione (ter-

mini, categorie, nessi, concetti) sia tra i lettori più bravi sia tra quelli meno bravi.

2^a ipotesi: la differenziazione interna è utile soprattutto per i lettori meno bravi.

Un *indicatore* della pertinenza delle ipotesi sarebbe il maggiore apprendimento, inteso come acquisizione di conoscenze tecniche, dei gruppi che dispongono di testi differenziati rispetto a quello di gruppi senza testi differenziati. Si tratterebbe di partire dal presupposto costruttivista per cui i discenti creano autonomamente nuove conoscenze a partire da quelle già assimilate (p.es. Duit & Widodo, 2004). Ciò significa che il contesto di apprendimento dovrebbe considerare le nozioni di cui gli allievi sono già in possesso, comprese le idee che non coincidono con il punto di vista scientifico. Il mancato intervento su queste nozioni è considerato una delle cause principali delle difficoltà di apprendimento (p.es. Bransford, Brown & Cocking, 2000; Duit, 1995; Treagust, Duit & Niewsandt, 2000). Nel loro modello delle competenze, Hardy et al. (2010) ritengono che vi sia un cambiamento concettuale (*conceptual change*, Möller, 2015; Sodian, 2008; Carey & Spelke, 1994; Carey, 2000; Carey, 2011; Möller, 2012; Anohina et al., 2009) da idee ingenuie a nozioni intermedie, fino a conoscenze scientificamente

valide. L'acquisizione di competenze tecniche viene descritta in modo analogo per quanto riguarda l'apprendimento della storia nel progetto *Historisches Denken* (Bietenhader & Kübler, 2012). Sulla base dei lavori preliminari di Engelen et al. (2002), di Hardy et al. (2010) e delle conoscenze scientifiche sul paleolitico (Höhneisen, 1998; Parzinger, 2014; Osterwalder & Robert, 1990), il team di progetto ha formulato una descrizione di concetti storici sull'*Età della pietra* nonostante le difficoltà determinate da un settore *ill-defined* come quello delle scienze storiche (Günther-Arndt, 2006).

Le conoscenze tratte dalle ricerche sulla leggibilità non hanno finora praticamente influenzato la produzione di strumenti didattici (Leitner, 2008, p. 708; Kahlert, 2010, p. 49 segg.). La comprensibilità di un testo – sul piano logico, lessicale e relativo al “mondo della vita” – non è indipendente dal contenuto, anzi, è fortemente definita dal settore specifico, e di conseguenza solo limitatamente trasferibile da un ambito scientifico all'altro (Lindauer & Schmelletin, 2011; Sodian, 2008). Sulla base delle ricerche sulla comprensibilità e la comprensione testuali (Wellenreuther, 2008; Schmölzer-Eibinger & Egger, 2012; Akhondi, 2010), abbiamo sviluppato un modello euristico, in funzione del quale abbiamo poi redatto tre versioni di livello diverso di un racconto sul paleolitico, sia in tedesco sia in retoromancio. Le versioni si differenziavano per la lunghezza, la sintassi e il lessico. La base per la formulazione del racconto era la versione tedesca. Dato che il genere di difficoltà non corrisponde sempre nelle due lingue (p.es. il problema

delle parole composte), il modello euristico e il testo sono stati adattati per il roman- cio. Per mantenere il flusso della storia, sono state inserite finestre di testo con informazioni aggiuntive sulla scorta del modello euristico.

1.1

Situazione linguistica di partenza

La Svizzera è un paese quadrilingue con un sistema scolastico federalista. Il retoromancio è una lingua minoritaria nel Cantone dei Grigioni, dove il materiale didattico (con l'eccezione di quello per l'apprendimento linguistico) è per lo più tradotto dal tedesco.

Figura 1
Situazione linguistica in Svizzera (fonte: Ufficio federale di statistica)

Ripartizione delle lingue ufficiali in Svizzera (2000)



Ufficio federale di statistica, www.bfs.admin.ch; censimento 2000

Il retoromancio è suddiviso in cinque *idiotimi*, vere e proprie lingue standardizzate che hanno sviluppato una tradizione scritta. Dal 1982, ne esiste una variante comune, il Rumantsch Grischun, riconosciuta ufficialmente a livello federale e cantonale (GR).

Gli allievi di lingua retoromancia partecipanti allo studio provenivano da tre regioni con idiomi differenti. La traduzione dei testi nei singoli idiomi è stata complicata dal fatto che le varianti regionali prevedono sovente vocaboli molto diversi per

indicare la stessa cosa, basti pensare alla parola tedesca *Heidelbeere* (mirtillo): la voce radicale *Beere* (bacca) si traduce in Vallader con *pommaraida* e in Sursilvano con *coclas*, mentre *Heidelbeere* si trasforma in *uzuns* in Vallader e in *izuns* in Sursilvano. Esempi analoghi sono numerosi in tutte le varianti del retoromancio.

2

Metodologia

2.1 Prova a campione e racconto

Per l'esame delle ipotesi si è fatto ricorso a una struttura quasi sperimentale, attuata in sei classi del Cantone Sciaffusa e in nove classi del Cantone dei Grigioni, per un totale di 221 allievi (108 femmine e 113 maschi) di quarta elementare.

Sono stati analizzati 204 set di dati. Un set comprendeva due file audio delle interviste (pre-test, post-test), un disegno e due fotografie della tecnica di ricostruzione (pre-test, post-test). Diciassette set di dati sono stati esclusi perché incompleti e presentanti carenze tecniche

Nel Cantone Sciaffusa, sono state selezionate scuole con un numero medio-alto di bambini la cui lingua madre non è il tedesco. Nel Cantone dei Grigioni, sono invece stati scelti istituti delle regioni Surselva e Bassa Engadina/Val Monastero. Non è stato necessario applicare un particolare criterio di selezione, se non il fatto che l'insegnamento si tenesse esclusivamente in romancio dalla prima alla quarta elementare.¹

Per l'elaborazione, sono stati redatti testi in tedesco di tre livelli di difficoltà, successivamente tradotti e adattati negli idiomi retoromanci Sursilvano e Vallader (risp. nella variante standard Rumantsch Grischun per la Val Monastero).

1 | In alcune scuole dell'area romancia, è previsto un insegnamento bilingue sin dalla prima elementare. In generale, in queste regioni si registra un numero medio-alto di bambini la cui lingua madre non corrisponde a quella della scuola.

2.2

Svolgimento e sondaggio tra i bambini

Il rilevamento dei dati prevedeva un pre-test, un'elaborazione e un post-test, secondo lo schema seguente.

Tabella 1

Schema dello svolgimento dello studio

Gruppo sperimentale		
Pre-test	Intervento 25-30 min., lettura del testo	Post-test
Colloquio sul disegno e sul metodo di ricostruzione	Buoni lettori Testo di livello elevato	Colloquio sul testo e sul metodo di ricostruzione
	Lettori medi Testo di livello medio	
	Lettori deboli Testo di livello basso	
Gruppo di controllo		
Pre-test	Intervento	Post-test
Colloquio sul disegno e sul metodo di ricostruzione	Testo di livello elevato	Colloquio sul testo e sul metodo di ricostruzione

All'inizio, i bambini dovevano realizzare un disegno sul tema dell'Età della pietra, il cui scopo era di attivare le conoscenze già assimilate. Il pre-test comprendeva un pre-colloquio e il metodo di ricostruzione. Gli allievi venivano dapprima interpellati sul loro disegno e, in seguito, pensando ad alta voce (cfr. Fromann, 2005) dovevano assegnare schede con immagini e parole sul tema dell'Età della pietra alle categorie "Appartiene all'Età della pietra", "Non appartiene all'Età della pietra" e "Termine sconosciuto o non chiaro" (vedi tabella 2). In una seconda fase, i bambini

erano chiamati a porre le schede in relazione tra loro, disponendole di conseguenza e motivando la decisione. Nel post-test, dopo essere stati dapprima interpellati sul racconto, gli allievi dovevano riposizionare le schede, apportando eventuali modifiche all'ordine scelto in precedenza e motivando tali rettifiche.

Per garantire l'unitarietà dello studio nel Cantone Sciaffusa e nei Grigioni, sono state redatte linee guida con istruzioni precise per la conduzione dei colloqui. È inoltre stata organizzata una formazione per gli intervistatori.

2.3

Metodi di valutazione

a) Valutazione dei colloqui con tecnica di ricostruzione (posizionamento delle schede)

I bambini dovevano assegnare i singoli elementi a una di tre categorie. Dal punto di

vista scientifico, ogni decisione poteva essere ritenuta corretta o errata. Ne derivano sei possibili categorizzazioni (vedi tabella 2).

Tabella 2

Schema di valutazione e interpretazione psicologico-cognitiva in base alla posizione degli elementi. I campi bianchi sono stati ritenuti corretti, quelli di colore scuro errati. I campi di colore chiaro sono stati valutati diversamente secondo la scala (vedi sotto).

Valutazione specialistica dell'elemento	Appartiene all'Età della pietra	Non appartiene all'Età della pietra	Non so
	Elementi corretti (numero=22)	Giusto: nozione corretta, risp. elemento conservato correttamente nella memoria	Mancato: nozione errata, risp. elemento conservato in modo errato nella memoria
Elementi errati (numero=6)	Falso allarme: concetto errato, risp. elemento conservato in modo errato nella memoria	Rifiuto corretto: nozione corretta, risp. elemento corretto non memorizzato	Insicurezza corretto: "non conoscenza" corretta, elemento non ancora memorizzato, risp. elemento noto ma dubbi nell'assegnarlo

Per ogni allievo, è possibile calcolare un valore che rispecchi la sua prestazione. Più alto è il valore, migliore è la prestazione:

Scala per la tecnica di ricostruzione
= "giusto" + "rifiuto corretto" = 28
– "mancato" – "falso allarme"
– "insicurezza corretto" – "insicurezza errato"

La scala presuppone che gli elementi assegnati alla categoria "Non so" siano sconosciuti ai ragazzi. Essi vengono pertanto considerati errori, dato che i partecipanti allo studio non sono stati in grado di assegnarli correttamente.² Per i valori della scala, è stata in seguito effettuata un'analisi bi-fattoriale della varianza con ripetizione della misurazione (variabili indipendenti: gruppo sperimentale, livello di lettura) allo scopo di testare le ipotesi dello studio.

Il posizionamento delle schede consente di valutare l'organizzazione cognitiva delle conoscenze dei partecipanti. L'analisi si limita tuttavia alla semplice assegnazione degli elementi al tema dell'Età della pietra e non permette di esprimersi sulle idee concettuali degli allievi in merito all'argomento.³ Riassumendo, la tecnica di ricostruzione consente di scoprire se un termine è noto o no e se gli scolari sono in grado di porlo cognitivamente in relazione con il tema

2 | È stata testata anche una scala che considera pure gli elementi assegnati alla categoria "Non so" ma, essendo gli esiti in gran parte identici, non verrà approfondita in questa sede.

dell'Età della pietra. L'analisi del posizionamento delle schede rileva quindi l'apprendimento al livello di base, ossia la formazione di legami associativi tra singoli contenuti memorizzati.

b) Valutazione dei colloqui mediante la matrice del conceptual change

I colloqui sono stati analizzati sulla scorta dei file audio, annotando direttamente la valutazione delle dichiarazioni nell'apposito modulo. La definizione dei quattro livelli è stata effettuata in una prima fase a livello euristico, rispettivamente in modo specifico secondo il settore, in seguito affinata progressivamente mediante due inchieste pilota e operazionalizzata con esempi tratti dal testo.

3 | L'ipotesi di procedere a un'analisi più completa che permettesse di approfondire maggiormente la comprensione contenutistica degli allievi è stata accantonata in considerazione dell'onere che un tale studio avrebbe comportato. L'idea sarebbe stata quella di non limitarsi alla categorizzazione (appartiene, non appartiene), bensì di analizzare il posizionamento dei diversi termini.

Tabella 3
Modello per la rappresentazione dei livelli concettuali delle conoscenze storiche e schema di valutazione con i quattro livelli concettuali (Engelen et al., 2002 e Moser et al., 2007)

D2	Concetti di pensiero storico sotto forma di cognizione di ri- e decostruzione Livello concettuale 2: riferimento all'oggi, confronto oggi-passato
D1	Concetti preliminari di pensiero storico sotto forma di accessi concettuali alla storia, classificazioni temporali Livello concettuale 1: concetto completo con termini concreti, concetto approssimativo con termini astratti
C2	Concetti, risp. nozioni scolastiche ampliabili (nessi ecc. di natura complessa) Nessi 2: categorie, termini generici, termini astratti; costruzioni frasali con locuzioni congiuntive (dal momento che, anche se ecc.)
C1	Concetti, risp. nozioni scolastiche ampliabili (nessi ecc. di natura elementare) Nessi 1: categorie, termini generici, termini astratti; enunciati finali (affinché ecc.); funzione con verbo adeguato
B2	Concetti, risp. nozioni scolastiche non supportati scientificamente (enumerazioni differenziate, teorie della vita quotidiana, analogie con il presente, supposizioni) Termini specifici e concreti; verbi più ricercati; menzione > 2 termini; supposizione concreta
B1	Conoscenze, risp. nozioni scolastiche non memorizzate correttamente dal punto di vista scientifico (enumerazioni semplici, teorie della vita quotidiana, analogie con il presente, supposizioni) Termini generali e concreti; verbi elementari; supposizioni vaghe
A	Concetto storico errato Dichiarazioni generali infantili

I dati acquisiti con i colloqui sono serviti al rilevamento dei cambiamenti nei livelli concettuali. Le dichiarazioni dei livelli concettuali inferiori erano più frequenti di quelle dei livelli superiori. Per ponderare maggiormente le dichiarazioni appartenenti ai livelli concettuali più alti, è stato assegnato loro un indice, cosicché se un allievo formulava solo dichiarazioni del livello D2 otteneva un valore superiore rispetto a chi formulava un numero equivalente di dichiarazioni di livello B1.

c) Inter-rater reliability

Per quantificare l'*inter-rater reliability*, sono stati selezionati, valutati separatamente e confrontati tra loro quattro colloqui. È stata ottenuta una corrispondenza pari a 0,804, considerata buona. Questo tipo di valutazione funge anche da base per l'analisi dei dati (v. punto 2.3 a).

3

Risultati

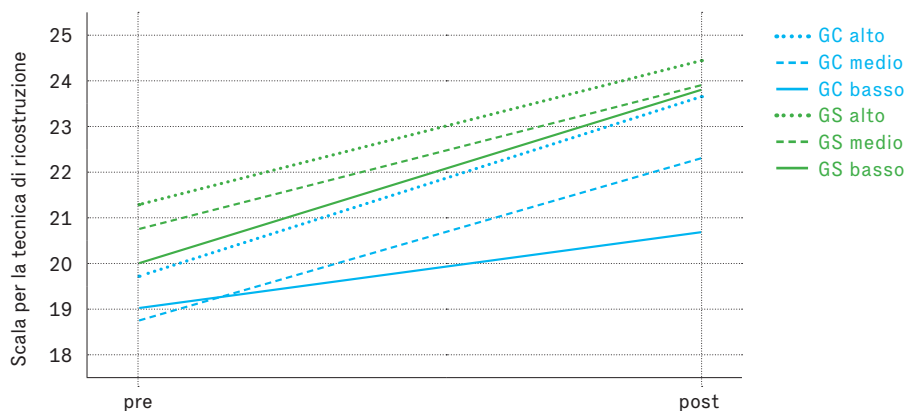
3.1 Cambiamento delle strutture cognitive con l'ausilio dei testi

Il confronto pre- e post-test del posizionamento delle schede rileva un miglioramento delle prestazioni a livello di conoscenze terminologiche sia nel gruppo sperimentale sia in quello di controllo (vedi figura 2). Dato che quasi tutti gli allievi hanno ottenuto un miglioramento notevole dello score della tecnica di ricostruzione, la portata dell'effetto è molto elevata: gruppo sperimentale $d=1,38$; gruppo di

controllo $d=1,07$. Buona parte del miglioramento è riconducibile al fatto che gli elementi assegnati alla categoria "Non so" sono stati spostati in quella corretta solo nella fase post-test. Anche considerando separatamente i gruppi secondo il livello di lettura, si osserva in tutti un incremento delle prestazioni.

Per le conoscenze terminologiche, si constatano effetti principali significativi per il gruppo sperimentale ($F(1/205)=13,104$, $p<0,001$), il livello di lettura ($F(2/205)=3,656$, $p=0,027$) e il momento del test ($F(1/205)=330,242$, $p<0,001$).

Figura 2
Cambiamento delle conoscenze terminologiche nel gruppo sperimentale (GS) e nel gruppo di controllo (GC)



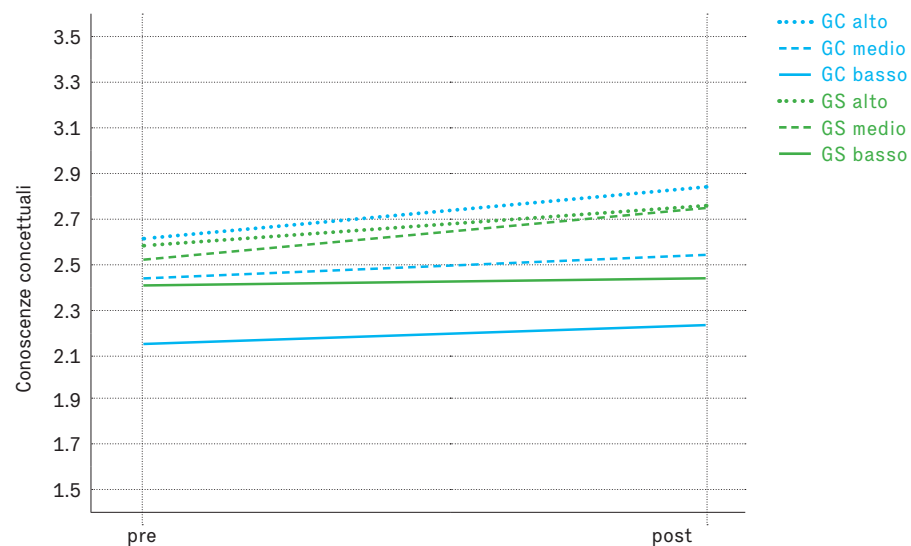
Emerge inoltre una triplice interazione significativa del gruppo sperimentale, del livello di lettura e del momento del test ($F(2/205)=5,133$, $p=0,007$).

La differenza principale tra il gruppo sperimentale e il gruppo di controllo è l'aumento delle prestazioni tra i lettori deboli: con un testo adattato (livello basso), i lettori deboli del gruppo sperimentale presentano un miglioramento superiore rispetto ai lettori medi e buoni. In seno al gruppo di controllo, invece, l'incremento delle prestazioni tra i lettori deboli è sensibilmente inferiore rispetto ai lettori medi e buoni.

3.2 Cambiamento delle conoscenze concettuali con l'ausilio dei testi

Per quanto riguarda le conoscenze concettuali, si constatano effetti principali rilevanti per il livello di lettura ($F(2/192)=14,856$, $p<0,001$) e il momento del test ($F(1/205)=22,425$, $p<0,001$). Nessuna delle possibili interazioni raggiunge invece un livello significativo. La figura 3 rappresenta graficamente l'evoluzione tra pre- e post-test nel confronto tra il gruppo sperimentale e il gruppo di controllo.

Figura 3
Cambiamento delle conoscenze concettuali nel gruppo sperimentale (GS) e nel gruppo di controllo (GC)



L'aumento dei valori medi nella tabella 4 mostra che gli allievi partecipanti hanno conseguito un (leggero) incremento delle loro nozioni storiche. È altrettanto evidente la distinzione tra i livelli di lettura (elevato, medio, basso). La tabella 4 considera tutti i ragazzi, mentre la tabella 5 separa i gruppi linguistici. L'effetto dell'aumento è tuttavia piuttosto debole e non conferma la nostra ipotesi iniziale. Sussistono solo lievi differenze tra il gruppo sperimentale e il gruppo di controllo. Sono invece interessanti i dettagli: nel *gruppo sperimentale* (testi adattati), sono i lettori medi a ottenere il miglioramento maggiore, nel *gruppo di controllo* (testo difficile per tutti) quelli bravi.

Nell'ottica dell'ampliamento delle nozioni concettuali di storia – a differenza dell'ampliamento delle conoscenze terminologiche –, per i lettori deboli non sembra fare differenza se leggono un testo difficile o facile: in nessuno dei casi si constatano miglioramenti.

I risultati descritti si accentuano se si analizzano separatamente i gruppi linguistici. Tra i bambini tedescofoni, i lettori medi nel gruppo sperimentale (con testi

adattati) presentano un aumento significativo delle prestazioni tra pre- e post-test. Nel gruppo di controllo, invece, sono i buoni lettori a evidenziare il maggiore incremento. Per quanto riguarda i ragazzi di lingua retoromancia, in entrambi i gruppi il maggiore miglioramento è ottenuto dai buoni lettori. Questa notevole differenza tra gruppi linguistici è al momento ancora inesplicabile.⁴

4 | Stiamo seguendo tre possibili teorie.

1. *Effetto della prova a campione*: verifichiamo se il campione retoromancio non sia soggetto a particolari effetti linguistici (p.es. bambini tedescofoni in scuole retoromance). Secondo i primi chiarimenti, questa teoria è poco probabile.
2. *Effetti legati al testo*: verifichiamo di nuovo la confrontabilità tra la difficoltà del testo in retoromancio (Sursilvano, Vallader, Rumantsch Grischun) e in tedesco. Va inoltre osservato che nei colloqui i bambini romanciofoni conoscevano sovente un

termine tecnico tedesco (come *Burggra-ben*), ignorando quale fosse l'equivalente retoromancio. Questo effetto ci sembra il più probabile.

3. *Effetti legati alla valutazione*: il modello di valutazione e gli esempi tratti dal testo erano in tedesco. I colloqui con i bambini romanciofoni sono quindi stati valutati sulla scorta del modello tedesco. Dato però che i codificatori romanciofoni parlavano molto bene tedesco, questo effetto ci pare poco probabile.

Tabella 4

Differenze dei valori medi tra pre- e post-test in quanto a cambiamento di livello concettuale (valore 1 = un livello su sei)

<i>Cambiamenti</i> ⁵	Tutti i livelli	Lettori deboli	Lettori medi	Buoni lettori
Gruppo sperimentale	0.14	0.03	0.22	0.17
Gruppo di controllo	0.15	0.08	0.10	0.22

Tabella 5

Differenze dei valori medi tra pre- e post-test in quanto a cambiamento di livello concettuale (valore 1 = un livello su sei) secondo il gruppo linguistico

<i>Cambiamenti</i>	Bambini tedescofoni				Bambini romanciofoni			
	Tutti i bambini	Lettori deboli	Lettori medi	Buoni lettori	Tutti i bambini	Lettori deboli	Lettori medi	Buoni lettori
Gruppo sperimentale	0.18	0.04	0.50	0.09	0.10	0.02	-0.02	0.27
Gruppo di controllo	0.16	-0.01	0.08	0.27	0.14	0.11	0.09	0.17

5 | I valori rispecchiano il cambiamento relativo dei livelli definiti (A, B1, B2, C1, C2, D1, D2). Un valore differenziale di 1 significa pertanto il passaggio a un livello superiore (p.es. da C2 a D1). La scala va da 0 a 6.

4

Interpretazione dei risultati e importanza per la produzione e l'utilizzo di testi

Secondo il contesto del nostro studio, i testi differenziati possiedono il potenziale di approfondire e migliorare le conoscenze terminologiche per i ragazzi di tutti i livelli di lettura. In particolare, i lettori deboli del gruppo sperimentale hanno beneficiato molto più dei lettori deboli del gruppo di controllo del fatto di poter contare su un testo adattato al loro livello più basso di lettura.

L'assenza di una differenza significativa a livello di aumento delle conoscenze terminologiche tra il gruppo sperimentale e il gruppo di controllo va valutata criticamente. Tutti i livelli di lettura del gruppo sperimentale hanno beneficiato in egual misura di un incremento delle conoscenze terminologiche. Ciò potrebbe essere riconducibile al fatto che il numero di partecipanti con un livello di lettura basso era insufficiente per comprovare una differenza statisticamente significativa tra il gruppo sperimentale e il gruppo di controllo.

Per quanto riguarda le conoscenze concettuali, come atteso non si constata alcuna differenza significativa tra il gruppo sperimentale e il gruppo di controllo (vedi tabella 4). L'ipotesi 1, ossia il fatto che i materiali d'insegnamento e apprendimento differenziati possono incrementare le capacità di comprensione (categorie, nessi,

concetti) sia tra i lettori più bravi sia tra quelli meno bravi, non può essere confermata. In seno al gruppo sperimentale, la differenza di comprensione concettuale tra lettori medi e buoni è lieve, ma si tratta comunque di un aumento di comprensione superiore a quello constatato nella categoria dei lettori deboli.

La seconda ipotesi può essere confermata per quanto attiene alle conoscenze terminologiche: potendo contare su testi adattati, i lettori più deboli del gruppo sperimentale dimostrano un incremento delle conoscenze chiaramente superiore a quello rilevato nel gruppo di controllo. Per l'acquisizione di nozioni di storia, invece, secondo il nostro studio i testi adattati per i lettori più deboli del gruppo sperimentale non sembrano fornire basi sufficienti. La seconda ipotesi è pertanto confermata solo a livello di conoscenze terminologiche.

Questi deboli effetti sui cambiamenti a livello concettuale potrebbero essere spiegabili con il fatto che un testo da leggere in venti-trenta minuti non rappresenta uno stimolo sufficiente. Occorrerebbe probabilmente integrare la lettura con domande sul testo oppure un'elaborazione dello stesso. Ciò corrisponde all'esito di altri studi, secondo i quali i concetti già

assimilati sono relativamente stabili (Möller, 1999) e richiedono interventi più duraturi. Per verificare tale ipotesi, un intervento didattico (elaborazione e approfondimento) dovrebbe accompagnare la differenziazione della lettura, così come è stato fatto nelle ricerche scientifiche menzionate.

Il nostro studio si limitava a testare l'effetto di un testo adattato, senza considerare quello di un'elaborazione e un approfondimento didattici. Ciò nonostante, è possibile trarre conclusioni sia per la produzione di materiale didattico sia per lo svolgimento delle lezioni.

Per affrontare temi complessi vale quindi la pena fornire ai lettori deboli testi semplificati, così da formare una base terminologica da sfruttare per l'elaborazione e l'approfondimento didattici. I testi più difficili, sul genere di quelli che si trovano di consueto negli strumenti didattici, non sembrano porre ostacoli insormontabili ai lettori buoni o molto buoni. In questi casi, una lezione differenziata è verosimilmente più efficace di un intervento complementare al testo. Per i lettori deboli, invece, i testi complessi sembrano costituire un ostacolo da non sottovalutare. L'acquisizione di basi terminologiche con l'ausilio di testi semplificati va verosimilmente integrata con un'elaborazione e un approfondimento didattici allo scopo di insediare i relativi concetti (storici). Anche il piano di studio 21 (D-EDK, 2014) nella sua versione di novembre 2014 sottolinea l'importanza di mettere in relazione lingua e contenuti della materia.

Continueremo ad analizzare i nostri dati, segnatamente nell'ottica degli effetti dei diversi livelli di testo in retoroman-

cio, e a seguire gli studi sugli accorgimenti didattici volti a testare l'efficacia dei testi differenziati in combinazione con offerte adattate di approfondimento e di elaborazione per i diversi livelli di prestazione.

Internally differentiated topical texts as a means to support learning

Executive Summary

—

Andreas Imhof, Eva Engeli, Markus Kübler

1

Theoretical basis and hypotheses

This project is based on the observation that teaching materials (and the topical texts used therein) remain the core component in educational settings (Oelkers, 2010, p. 39; Kahlert, 2010) but that they cannot meet the demands of individualised teaching practices (Moser Opitz, 2010). This issue must be understood against the backdrop of calls for an inclusive school system as well as in the context of the organisational consequences of developing internally differentiated teaching materials. Numerous studies in various disciplines have corroborated findings that standard textbooks are generally much too difficult for average readers (e.g. Lässig, 2010; Leisen, 2006; von Borries, 2010). It therefore appears plausible that this state of affairs particularly impacts children who speak a foreign language (Moser, 2001) and that there is a need to prepare texts according to level of difficulty. Such approaches are justified by the claim that differentiated teaching materials can improve learner performance in understanding content (terms, categories, relation, concepts).

We have derived two hypotheses from these observations:

First hypothesis: Internally differentiated educational material can improve learner performance in understanding content (terms, categories, relations, concepts) in both strong and weak readers.

Second hypothesis: Weak readers benefit most from internally differentiated texts.

An indicator that the hypotheses are correct would be a greater increase in knowledge among learners who were given differentiated texts compared with learners using non-differentiated texts. We understand an increase in knowledge to be the acquisition of subject-specific concepts. This understanding is based on constructivist approaches maintaining that learners construct new knowledge based on their prior knowledge (e.g. Duit & Widodo, 2004). This therefore implies that educational settings must activate and link existing knowledge in learners – even conceptions that are inconsistent with scholarly or scientific findings. The omission to pick up such inconsistent conceptions is believed to be one of the main factors in learning difficulties (e.g. Bransford, Brown & Cocking, 2000; Duit, 1995; Treagust, Duit & Niewsandt, 2000). Hardy et al.

(2010) developed their competence model on the basis of a general description of conceptual change (Möller, 2015; Sodian, 2008; Carey & Spelke, 1994; Carey, 2000; Carey, 2011; Möller, 2012; Anohina et al., 2009) that progresses from naïve conceptions to explanations of everyday life on to scientific explanations. The project *Historisches Denken* (historical thinking) (Bietenhader & Kübler, 2012) describes the acquisition of subject-specific concepts in history lessons along the same lines. Building on the work of Engelen et al. (2002) and Hardy et al. (2010) as well as on the findings from scholarship on the Palaeolithic Era (Höhneisen, 1998; Parzinger, 2014; Osterwalder & Robert, 1990), we formulated a description of historical concepts from the *Stone Age* – despite the fact that history is generally considered an ill-defined domain (Günther-Arndt, 2006).

To date, the comprehensive findings from research on text comprehensibility have had little impact on the production of teaching materials (Leitner, 2008, p. 708; Kahlert, 2010, p. 49f.). Text comprehensibility in the sense of textual coherence, terminology and relevance to real life is not independent of context; indeed, it is essentially defined by a given discipline and thus difficult to transfer from one subject to another (Lindauer & Schmelletin, 2011; Sodian, 2008). Based on research on text comprehensibility and reading comprehension (Wellenreuther, 2008; Schmolzer-Eibinger & Egger, 2012; Akhondi, 2010), we developed a heuristic model for text difficulty and, based on this model, created a non-fiction narrative on the Palaeolithic Era for three reading levels in both

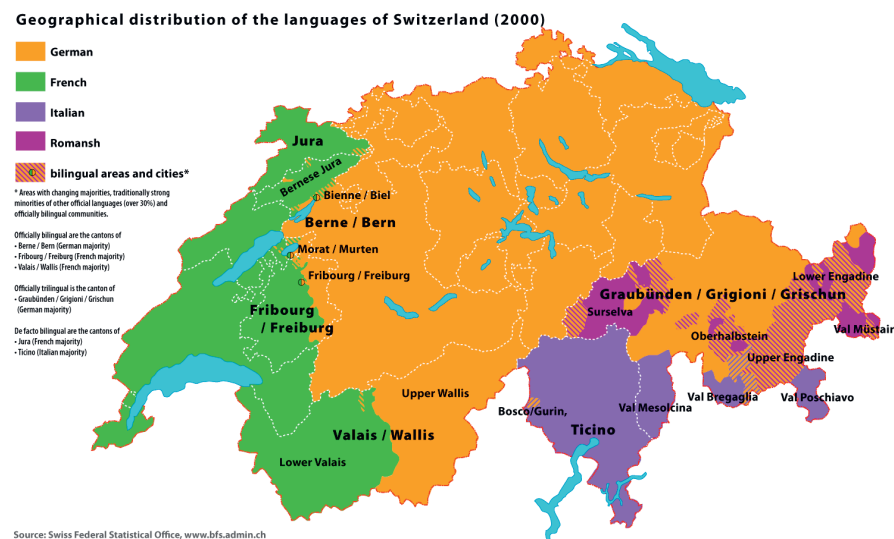
German and Rhaeto-Romanic. Factors for level of difficulty include text length, syntax and lexis. The German version served as the basis and was translated into Rhaeto-Romanic. Because difficulty in Rhaeto-Romanic is rooted in other linguistic factors (e.g. compound words), the heuristic model and the narrative were adjusted accordingly. To retain good flow in the story, text windows with additional facts were incorporated on the basis of the heuristic model.

1.1 The linguistic starting point

Four languages are spoken in Switzerland, and the school system reflects the country's federalist structure. Rhaeto-Romanic is a minority language spoken in the Canton of Graubünden, and the textbooks used in Rhaeto-Romanic school districts are mainly translated from German (with the clear exception of the language textbook).

Chart 1

Linguistic situation in Switzerland (Source: Swiss Federal Statistical Office)



There are five historically developed varieties of written Rhaeto-Romanic called idioms. In contrast to the term *idiom* as it is generally used in linguistics, dialects are not meant in this setting; instead, the Rhaeto-Romanic idioms are regional, standardised written language varieties. In addition to the historical idioms, a common written variety, Rumantsch Grischun, has existed since 1982; this variety is officially recognised at the federal and cantonal (GR) levels.

The Rhaeto-Romanic-speaking participants in the study are from three regions with different idioms. Translating the source text into the individual idioms was complex because, among other reasons, the local idioms often use very different

words to designate the same thing: a discussion of the German and various Rhaeto-Romanic words for “blueberry” serves as an example. In the Vallader and Sursilvan idioms the basic terms for “berry” – pommaraida and coclas, respectively – are already entirely unrelated. Moreover, while the German word *Heidelbeere* (like the English word “blueberry”) is easily recognised as a modification of the root word *Beere* (“berry”), the Vallader or Sursilvan terms for blueberry – uzuns and izuns, respectively – cannot be derived from their words for berry.

2

Methodological design

2.1 Participants and non-fiction narrative

To test our hypotheses, we selected a quasi-experimental design and applied it in six classes at schools in the Canton of Schaffhausen and nine classes at schools in the Canton of Graubünden; a total of 221 children (108 girls and 113 boys) in the 4th class of primary school participated.

204 datasets were analysed; each dataset comprised two audio files of the interviews (pre-test, post-test), a drawing and two photos of the structure laying technique SLT (pre-test, post-test). Seventeen datasets were rejected because they were either incomplete or they had technical defects.

Schools with an average or high number of children who do not speak German (the language of instruction) as a native language were chosen for the experiment in the Canton of Schaffhausen. In the Canton of Graubünden, schools in the regions of Surselva and Lower Engadine/Val Müstair were selected. No particular criterion was necessary other than all schools in these regions conduct lessons entirely in Rhaeto-Romanic from the start of schooling through the fourth class.¹

For the treatment, topical texts with three levels of difficulty were drafted in

German and then translated and adapted into the Rhaeto-Romanic idioms Sursilvan and Vallader (as well as into the standardised language Rumantsch Grischun for Val Müstair) for the schools in the Canton of Graubünden.

1 | This differs from other school districts in the Rhaeto-Romanic-speaking areas in which, from the first class on, there are two languages of instruction: German and Rhaeto-Romanic. In general, there is an average to high percentage of children in the Rhaeto-Romanic-speaking areas whose native language is not the language of instruction.

2.2

Procedures and interviews with children

Data was collected within the same time frame as the pre-test, the treatment and the post-test. The procedure was as follows:

Table 1
Overview of procedure

Test group		
Pre-test	Intervention 25-30 min. reading the topical text	Post-test
Interview about drawing and structure laying technique (SLT)	strong readers high difficulty	Interview about text and SLT
	average readers average difficulty	
	weak readers low difficulty	
Control group		
Pre-test	Intervention	Post-test
Interview about drawing and structure laying technique (SLT)	high difficulty	Interview about text and SLT

At the outset, the children were asked to draw a picture about the Stone Age, an exercise designed to activate prior knowledge. The pre-test consisted of the subsequent pre-interview and the structure laying technique. In the interview, the children were first asked about their drawings; afterwards they were shown flashcards with pictures and words related to the Stone Age and were instructed to think aloud (cf. Frommann, 2005) while placing the cards in one of the three categories “Belongs to the

Stone Age”, “Doesn’t belong to the Stone Age” and “I don’t know” (see Table 2), thus creating a placement order (structure laying technique SLT). In a second step, the children were asked to identify relations between the cards, to justify their decisions as well as possible and to place the cards accordingly. For the post-test, the children were first asked questions about the narrative; they then again laid flashcards with pictures and words with the requirement that they explain any differences compared

with their first decisions as well as possible.

To guarantee consistent test situations in Schaffhausen and Graubünden, we developed a guideline with exact instructions for conducting the interviews. The interviewers were given advance training.

2.3

Methods of data interpretation

a) Interpretation of the interviews with the structure laying technique (placement order)

The children were allowed to place the individual items in three fields. The solution chosen could be right or wrong with regard to the historical facts; as a result, there are six possible categories for each item (see Table 2).

Table 2
Assessment scheme and cognitive psychological interpretations based on the placement of the items. The white fields are assessed as correct, the coloured fields as incorrect. Depending on the scale, the lightly coloured fields were evaluated differently (see below.)

		Belongs to the Stone Age	Doesn't belong to the Stone Age	I don't know
Assessment of the items	Right items (Number=22)	“Hits”: element factually right or correctly linked	“Misses”: element factually wrong or incorrectly linked	“Uncertainty wrong”: false lack of knowledge, element not yet linked, or a known element combined with uncertainty about the relation
	Wrong items (Number=6)	“False alarm”: wrong concept, element linked incorrectly	“Correct disagreement”: element factually right or correctly not linked	“Uncertainty right”: element not yet linked, or a known element combined with uncertainty about the relation

With this scheme, a value reflecting performance can be determined for each pupil; the better the performance, the higher the value:

Scale structure laying technique =
Hits + Correct disagreements = 28
– Misses – False alarms
– Uncertainty right – Uncertainty wrong

The scale is designed on the understanding that the pupils are not familiar with the items placed in the “I don’t know” field. These items are deemed errors because the test persons are unable to identify a correlation explicitly.² Subsequently, to test the study’s hypotheses, a two-factor variance analysis with repeated measurements was calculated to attain the scale’s values (independent variables were test group and reading level).

The placement order allows an assessment of the test person’s cognitive organisation of his or her knowledge. This analysis is, however, limited solely to the arrangement of the items on the subject of the Stone Age. The learner’s conceptual understanding of the Stone Age cannot be assessed using this technique.³ SLT allows an assessment of two aspects: first, whether a learner is familiar with a term and, second, whether they can relate it cognitively to the Stone Age. The analysis of the placement order therefore allows a

2 | Another scale that considered the items placed in “I don’t know” was also tested. Only the scale for the structure laying technique is described here, as the results for both scales demonstrated only minor differences.

measurement of learning at a basal level, i.e. a measurement of the development of associative links of individual elements stored in a test person’s memory.

b) Interpretation of the interviews using the matrix of conceptual change

The interviews were interpreted using the audio files; the team noted the assessments of statements directly in the corresponding level of the chart. In the first phase, the definition of four conceptual levels was done according to heuristic and discipline-specific principles; it was refined step by step using two pilot studies and operationalised by means of reference examples.

3 | A follow-up study able to better test the learners’ comprehension of content was proposed during the study but not pursued further due to its complicated nature. The basic idea, however, is to not only analyse the categories (belongs to; doesn’t belong to) but mainly how the concepts were related to one another.

Table 3
Model to present the conceptual levels of historical knowledge (an assessment scale with four defined conceptual levels) (acc. to Engelen et al., 2002, and Moser et al., 2007)

D2	Concepts of historical thinking as an insight into reconstruction and deconstruction Conceptual level 2: Relationship to today, comparison today-earlier
D1	Preconceptions of historical thinking as a conceptual introduction to history, temporal classifications Conceptual level 1: Entire concept concerning concrete terms, partial concept concerning abstract terms
C2	Historical preconcepts or pupil conceptions (e.g. correlations of a complex nature) Correlations 2: Categories, generic terms, abstract terms; sentence constructions with conjunctions (e.g. since, because, although)
C1	Historical preconcepts or pupil conceptions (e.g. correlations of a simple nature) Correlations 1: Categories, generic terms, abstract terms; final clauses (in order to...); function with suitable verb
B2	Non-viable subject-specific concepts or pupil conceptions (nuanced enumerations, generic theories, present-day analogies, conjectures) Specific and concrete terms; nuanced verbs (use, live); naming > 2 terms; concrete conjectures
B1	Factually incorrectly linked knowledge or incorrect pupil conceptions (simple enumerations, generic theories, present-day analogies, conjectures) General and concrete terms; verb have; vague conjectures
A	No correct historical concept General statements from the pupils’ life experience

The data collected in the interviews was used to measure changes in conceptual level. Statements in the lower conceptual levels were more frequent than in the higher levels. Statements in a higher conceptual level were weighted using an index to give them greater significance; they thus received a higher value. For instance, if a learner only made statements within the conceptual level D2, a higher value was given compared to an equal number of statements made in the level B1.

c) Interrater reliability

To quantify the inter-rater reliability, four interviews were selected, assessed separately and subsequently compared. A concordance of 0.804 was attained, which is a good result. This type of rating also provides a basis to interpret the data (see chapter 2.3 a).

3

Findings

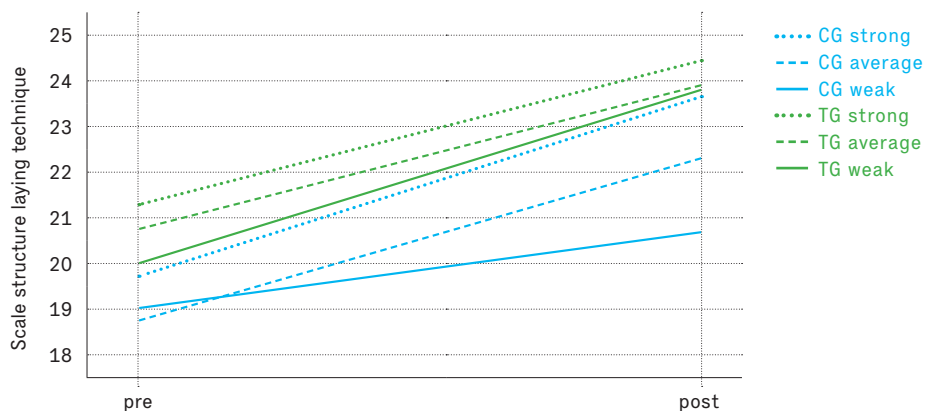
3.1 Changes in cognitive structures

The pre-test and post-test comparison of the placement orders (SLT) reveals an improvement in the learners' knowledge of historical terms, both in the test group and in the control group (cf. Chart 2). Because nearly all learners achieved better results in the structure laying order and because the improvement is relatively appreciable, the comparison of pre-test and post-test in both groups yields very high effect sizes: test group $d=1.38$, control group $d=1.07$. A large part of the

improvement is related to the items placed in the field "I don't know"; these were not correctly integrated into the placement order until the post-test was implemented. An improved performance was also registered in all groups when the analysis was isolated according to reading level.

Significant main effects were observed in the area of knowledge of historical terms for the test group ($F(1/205)=13.104$, $p<0.001$), the reading level ($F(2/205)=3.656$, $p=0.027$) and the time of the test ($F(1/205)=330.242$, $p<0.001$). A significant triple interaction between the test group, the reading level and the time of the test was also

Chart 2
Change in knowledge of historical terms in the test group (TG) and the control group (CG)



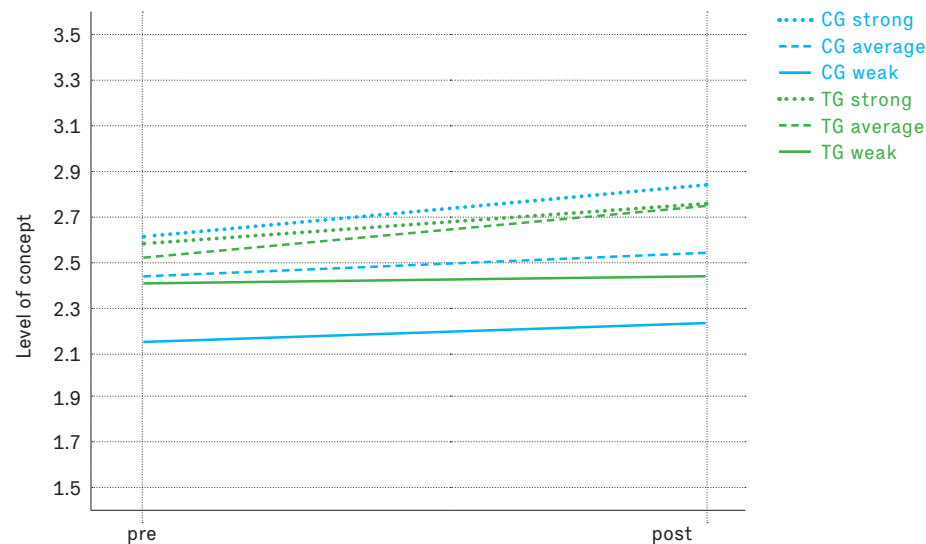
observed ($F(2/205)=5.133$, $p=0.007$).

The main difference between the test group and the control group becomes apparent in the performance improvement among weak readers. When using an adapted test (low level), weak readers in the test group demonstrate a stronger improvement than do average and strong readers. In the control group, by contrast, this improvement is significantly smaller than in the other two reading levels.

3.2 Changes in conceptual knowledge

Regarding conceptual knowledge, significant main effects can be observed for reading levels ($F(2/192)=14.856$, $p<0.001$) and for the time of the test ($F(1/205)=22.425$, $p<0.001$). Nevertheless, neither of the possible interactions achieved the required level of significance. Chart 3 (below) depicts the development between the pre-test and post-test in the test group and control group.

Chart 3
Change in conceptual knowledge in the test group (TG) and the control group (CG)



The mean increases depicted in Table 4 demonstrate that the pupils attained a (slight) improvement in their knowledge of history. Moreover, the individual reading levels are clearly recognisable: Table 4 describes all children; Table 5 differentiates between the language groups. The effect of the increase is, however, rather weak and does not confirm our hypothesis. Only negligible differences between the test group and control group are evident. The details nevertheless present an interesting picture: in the test group (with adjusted texts), the average readers attained the greatest improvement, while in the control group (difficult text for all pupils), the strong readers improved the most.

In contrast to improving their knowledge of historical terms, the development of conceptual knowledge in weak readers is apparently not influenced by a text's degree of difficulty. Indeed, weak readers attained almost no improvement in this area, even when given an adapted text.

The results described above are yet more accentuated when the language groups are analysed separately. Among the German-speaking pupils, the average readers in the test group (adjusted text) attained a significant improvement in the post-test, while the best results were achieved by the strong readers in the control group. Among the Rhaeto-Romanic speakers, it is the strong readers in both groups that achieved the greatest improvement. No explanation for the marked difference is currently available.⁴

4 | We are exploring three possibilities to explain the difference:

1. *Effect of sampling*: This possibility explores whether the sampling of the Rhaeto-Romanic-speaking pupils contains specific language-related effects (for example, German-speaking children who attend Rhaeto-Romanic-speaking schools). Initial findings indicate that this explanation is less probable.
2. *Text-inherent effects*: Here we re-evaluate whether the difficulty of the text in Rhaeto-Romanic (Sursilvan, Vallader, Rumantsch Grischun) did not correlate exactly to the difficulty of the German text. It must furthermore be noted that,

in the interviews, the Rhaeto-Romanic speakers often knew German terminology (e.g. "Burggraben") in the interviews, but they were not familiar with the Rhaeto-Romanic terms. We believe this effect is the most probable.

3. *Rating effect*: The rating template of the conceptual change and the reference examples were in German and were based on German examples. The interviews were therefore assessed on the basis of ratings in the German language. Nonetheless, we believe this explanation is less plausible because the Rhaeto-Romanic-speaking test persons are adept in German.

Table 4

Mean difference between pre-test and post-test with regard to the change in conceptual knowledge (value 1 = one conceptual level from a total of six) (RL = reading level)

<i>Changes</i> ⁵	All levels	RL weak	RL average	RL strong
Test group	0.14	0.03	0.22	0.17
Control group	0.15	0.08	0.10	0.22

Table 5

Mean difference between pre-test and post-test with regard to the change in conceptual knowledge (value 1 = one conceptual level from a total of six) according to language groups (RL = reading level)

<i>Changes</i>	German-speaking pupils				Rhaeto-Romanic-speaking pupils			
	All pupils	RL weak	RL average	RL strong	All pupils	RL weak	RL average	RL strong
Test group	0.18	0.04	0.50	0.09	0.10	0.02	-0.02	0.27
Control group	0.16	-0.01	0.08	0.27	0.14	0.11	0.09	0.17

5 | The values mirror the relative change to the defined levels (A, B1, B2, C1, C2, D1, D2). A difference of 1 therefore signifies that a learner has attained a higher conceptual level (e.g. from C2 to D1) on a scale of 0 to 6.

4

Interpretation of the findings and their significance for the production and use of topical texts

According to our setting, internally differentiated texts clearly have potential to help pupils in all three reading levels expand and refine their knowledge of historical terms. Compared with their peers in the control group, the weak readers in the test group profited particularly strongly from having the learning material on the topic adjusted (i.e. made simpler) to suit their level.

It must, however, be noted that no significant difference between the test group and control group can be observed in the improvement in knowledge of historical terms. Pupils of all abilities benefited to a comparable degree. This may be due to the circumstance that the number of test persons in the weakest reading level was too low to ascertain a statistically significant difference between the test group and the control group.

No significant difference between the test group and the control group could be established regarding conceptual knowledge, and the differences largely correspond to general expectations for learning outcomes (cf. Table 4). As such, the first hypothesis – that internally differentiated teaching materials will improve com-

prehension (terms, categories, correlations, concepts) in both strong and weak readers – cannot be corroborated. The observed differences in conceptual knowledge between the average and strong readers in the test group are small. Within these minimal improvements from pre-test to post-test, it is the conceptual knowledge of average and strong readers that increases the most.

The second hypothesis has been confirmed with regard to knowledge of historical terms. Weaker readers in the test group clearly demonstrate a greater improvement compared to their peers in the control group. Nevertheless, our findings indicate that the adapted teaching materials do not offer weaker readers a broad enough base to improve their conceptual knowledge. As such, our second hypothesis can only be confirmed with regard to knowledge of historical terms.

The weak effects regarding changes in conceptual level are possibly related to the structure of the pedagogical intervention: because the text should be read in twenty to thirty minutes, the stimuli required to effect change are probably too weak. It is likely that a concept change can only be

achieved if, in addition to reading a text, questions about the text or other forms of processing and reinforcing content are introduced. This hypothesis corresponds to published findings demonstrating that pre-conceptions are fairly stable (Möller, 1999) and that they require more sustainable interventions. To verify this hypothesis, a pedagogical intervention (reinforcing and expanding on the materials used) would have to be designed to complement the internally differentiated topical text (as was the case in the above-mentioned studies).

Our study focused solely on testing the effects of an adapted text, with no attention given the effects of processing and expanding on the topic in the course of a lesson. Nevertheless, the results allow general statements to be made regarding the preparation of teaching materials and their implementation in the classroom.

On the one hand, weak readers evidently benefit from simplified texts when learning complex subjects, as such texts enable less practiced readers to develop basic knowledge of historical terms and the fundamental ability to link these terms. This new knowledge can then be processed and expanded in subsequent teaching sequences. On the other hand, difficult topical texts – frequently found in standard teaching materials, as studies have shown – evidently do not pose a major obstacle to good to very good readers. For such learners, internally differentiated pedagogical interventions would most likely be of greater value than adapted texts. But difficult textbook materials do appear to pose a challenge to weak readers. The acquisition of the ability to

link terms by means of simplified textbook materials would most probably require additional teaching sequences to reinforce and expand the pupils' knowledge, thus allowing them to develop their conceptual knowledge (of history). The curriculum for Swiss German schools ("Lehrplan 21", designed by the German-speaking section of the Swiss Conference of Cantonal Ministers of Education) in its version from 2014 foregrounds the importance of combining subject and language.

We will continue examining the data compiled in this study, particularly with regard to the effects the different levels of difficulty had in the Rhaeto-Romanic setting. Other aspects include investigating learning sequences that test the effects of internally differentiated texts in combination with methods to reinforce and expand on a given subject that have also been adjusted to accommodate different levels of ability.

Bibliographie

Bibliografia

Bibliography

- Akhondi, M., Aziz Malayeri, F. & Abd Samad, A. (2011). How to Teach Expository Text Structure to Facilitate Reading Comprehension. In: *The Reading Teacher*, 64(5), 368-372.
- Anohina, A., Vilkelis, M. & Lukasenka, R. (2009). Incremental Improvement of the Evaluation Algorithm in the Concept Map Based Knowledge Assessment System. *International Journal of Computers, Communication & Control*, 4(1), 6-16.
- Beerenwinkel, A. & Parchmann, I. (2010). Ansätze zur Berücksichtigung von Lernervorstellungen in Lehrtexten und Schulbüchern zum kontextorientierten Lernen. *Beiträge zur Lehrerbildung*, 28(1), 62-72.
- Bietenhader, S. & Kübler, M. (2012). Historisches Denken von 4- bis 10-jährigen Kindern – Ergebnisse einer Pilotstudie. In: Giest, G., Heran-Dörr, E. & Archie, C. (Hrsg.), *Lernen und Lehren im Sachunterricht. Zum Verhältnis von Konstruktion und Instruktion. Probleme und Perspektiven des Sachunterrichts*. Band 22. Bad Heilbrunn: Klinkhardt, 151-57.
- Borries, B. von (2010). Wie wirken Schulbücher in den Köpfen der Schüler? Empirie am Beispiel des Faches Geschichte. In: Fuchs, E., Kahlert, J. & Sandfuchs, U. (Hrsg.), *Schulbuch konkret. Kontexte – Produktion – Unterricht*. Bad Heilbrunn: Klinkhardt, 102-17.
- Bransford, J.D., Brown, A.L. & Cocking, R.R. (2000) (Eds). *How People Learn: Brain, Mind, Experience, and School*. Washington, D.C.: National Academy Press.
- Bundesamt für Statistik. *Kartenarchiv Schweiz* (1990), Online unter: http://www.bfs.admin.ch/bfs/portal/de/index/regionen/thematische_karten/maps/bevoelkerung/sprachen_regionen.parsys.0002.PhotogalleryDownloadFile1.tmp/k16.01s.pdf (6.10.15)
- Carey, S. (2000). The Origin of Concepts. *Journal of Cognition and Development*, 1(1), 37-41.
- Carey, S. (2011). Précis of The Origin of Concepts. *Behavioral and Brain Sciences*, 34(3), 113-124.
- Carey, S. & Spelke, E. (1994). Domain-specific Knowledge and Conceptual Change. In: Hirschfeld, L.A. & Gelman, S.A. (1994/1998), *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York: Cambridge University Press, 169-200.
- D-EDK (2014). *Lehrplan 21. Sprachregionaler Lehrplan für die deutsch- und mehrsprachigen Kantone*. Luzern. Online unter: http://lehrplan.ch/sites/default/files/lp21_rahmeninformation_%202014-11-06.pdf (7.10.15)

- Duit, R. (1995). Zur Rolle der konstruktivistischen Sichtweise in der natur-wissenschaftsdidaktischen Lehr-Lernforschung. *Zeitschrift für Pädagogik* 41(6), 905-926.
- Duit, R. & Widodo, A. (2004). Konstruktivistische Sichtweisen vom Lehren und Lernen und die Praxis des Physikunterrichts. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 233-255.
- Engelen, A., Jonen, A. & Möller, K. (2002). Lernfortschrittdiagnosen durch Interviews – Ergebnisse einer Pilotstudie zum „Schwimmen und Sinken“ im Sachunterricht der Grundschule. In: Spreckelsen, K., Möller, K. & Hartinger, A. (Hrsg.), *Ansätze und Methoden empirischer Forschung zum Sachunterricht*. Bad Heilbrunn: Klinkhardt, 155-73.
- Fromann, U. (2005). Die Methode „Lautes Denken“. *e-teaching.org*. Online unter: <http://www.e-teaching.org/didaktik/qualitaet/usability/> (28.2.2012)
- Günther-Arndt, H. (2006). Conceptual-Change-Forschung: Eine Aufgabe für die Geschichtsdidaktik? In: Günther-Arndt, H. & Sauer, M. (Hrsg.), *Geschichtsdidaktik empirisch. Untersuchungen zum historischen Denken und Lernen*. Zeitgeschichte-Zeitverständnis, Bd. 14. Berlin: Lit, 251-78.
- Hardy, I., Kleickmann, Th., Koerber, S., Mayer, D., Möller, K., Pollmeier, J., Schwippert, K. & Sodian, B. (2010). Die Modellierung naturwissenschaftlicher Kompetenz im Grundschulalter. In: Klieme, E., Leutner, D. & Kenk, M. (Hrsg.), *Zeitschrift für Pädagogik. Beiheft Kompetenzmodellierung Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes*, 56, 115-125.
- Historisches Denken* (o.J.). Online unter: <http://www.historischesdenken.ch> (6.10.15)
- Höhneisen, M. (1998). Die Frühzeit des Menschen. In: Furger, A., Fischer, C. & Höhneisen, M. (Hrsg.), *Die ersten Jahrtausende*. Zürich: NZZ, 63ff.
- Kahlert, J. (2010). Das Schulbuch – Stiefkind der Erziehungswissenschaften. In: Fuchs, E., Kahlert, J. & Sandfuchs, U. (Hrsg.), *Schulbuch konkret. Kontexte – Produktion – Unterricht*. Bad Heilbrunn: Klinkhardt, 41-58.
- Langer, I. (1993). Verständlich informieren. Ein Beispiel empirischer Forschung. In: Fittkau, B. (Hrsg.), *Pädagogisch-psychologische Hilfen für Erziehung, Unterricht und Beratung*. Bd. 2. Paderborn: Pedersen, 378-401.

Lässig, S. (2010). Wer definiert relevantes Wissen? Schulbücher und ihr gesellschaftlicher Kontext. In: Fuchs, E., Kahlert, J. & Sandfuchs, U. (Hrsg.), *Schulbuch konkret. Kontexte – Produktion – Unterricht*. Bad Heilbrunn: Klinkhardt, 199-215.

Leisen, J. (2006). Ein Sachtext – Zehn Strategien zur Bearbeitung von Sachtexten. *Naturwissenschaften im Unterricht – Physik*, 95, 12-23.

Leitner, J. (2008). Die Verständlichkeit von Schulbuchtexten – Ergebnisse einer empirischen Analyse. *Erziehung und Unterricht*, 9-10, 708-713.

Lindauer, T. & Schmelletin, C. (2011). Die unterschätzte Bedeutung der Sprachkompetenzen für das fachliche Lernen. Neue Zürcher Zeitung, 29. Juni 2011. Online unter: http://www.nzz.ch/nachrichten/startseite/die_unterschaetzte_bedeutung_der_sprachkompetenzen_fuer_das_fachliche_lernen_1.11095073.html (23.2.2012)

Möller, K. (1999). Konstruktivistisch orientierte Lehr-Lernprozessforschung im naturwissenschaftlich-technischen Bereich des Sachunterrichts. In: Köhnlein, W., Marquardt-Mau, B. & Schreier, H. (Hrsg.), *Vielperspektives Denken im Sachunterricht*. Forschungen zum Sachunterricht, 3. Bad Heilbrunn: Klinkhardt, 125-91.

Möller, K. (2012). Konstruktion vs. Instruktion oder Konstruktion durch Instruktion? Konstruktionsfördernde Unterstützungsmassnahmen im Sachunterricht. In: Giest, G., Heran-Dörr, E. & Archie, C. (Hrsg.), *Lernen und Lehren im Sachunterricht. Zum Verhältnis von Konstruktion und Instruktion. Probleme und Perspektiven des Sachunterrichts*, Band 22. Bad Heilbrunn: Klinkhardt, 37-50.

Möller, K. (2015). Genetisches Lernen und Conceptual Change. In: Kahlert, J., Fölling-Albers, M., Götz, M., Hartinger, A. & Wittkowske, S. (Hrsg.), *Handbuch Didaktik des Sachunterrichts*. Bad Heilbrunn: Klinkhardt, 243-56.

Moser, D.V. & Wiher, P. (2007). Historisches und politisches Wissen von Jugendlichen – am Ende der obligatorischen Schulzeit. In: Gautschi, P., Moser, D.V., Reusser, K. & Wiher, P. (Hrsg.), *Geschichtsunterricht heute. Eine empirische Analyse ausgewählter Aspekte*. Bern: hep, 211-62.

Moser Opitz, E. (2010). Innere Differenzierung durch Lehrmittel: (Entwicklungs-) Möglichkeiten und Grenzen am Beispiel von Mathematiklehrmitteln. *Beiträge zur Lehrerbildung. Lehr- und Lernmedien als Träger von Inhalten und Konzepten – Bedeutung für die Lehrerbildung*, 28(1), 53-61.

Moser, U. (2001). *Für das Leben gerüstet? Die Grundkompetenzen der Jugendlichen – Kurzfassung des nationalen Berichtes PISA 2000. Bildungsmonitoring Schweiz OECD – PISA Programme for International Student Assessment*. Bern und Neuenburg: BFS/EDK.

Oelkers, J. (2010). Bildungsstandards und deren Wirkung auf die Lehrmittel. *Beiträge zur Lehrerbildung. Lehr- und Lernmedien als Träger von Inhalten und Konzepten – Bedeutung für die Lehrerbildung*, 28(1), 33-41.

Osterwalder, C. & Robert, A. (1990). *fundort schweiz. Von den Eiszeitjägern zu den ersten Bauern*. Band 1. Solothurn: aare.

Parzinger, H. (2014). *Die Kinder des Prometheus. Eine Geschichte der Menschheit vor der Erfindung der Schrift*. München: Beck.

Schmölzer-Eibinger, S. & Egger, E. (2012). *Sprache in Schulbüchern. Empfehlungen zur Sprachverwendung in Schulbüchern für SchulbuchautorInnen, GutachterInnen und Schulbuchverlage*. Hrsg. vom Bundesministerium für Unterricht und Kultur. Wien: Druckzentrum des BMUKK.

Schwaborn, A., Thillmann, H., Leopold, C., Sumfleth, E. & Leutner, D. (2010). Der Einsatz von vorgegebenen und selbst generierten Bildern als Textverstehenshilfe beim Lernen aus einem naturwissenschaftlichen Sachtext. *Zeitschrift für Pädagogische Psychologie*, 24(3-4), 221-233.

Sodian, B. (2008). Entwicklung des Denkens. In: Oerter, R. & Montada, L. (Hrsg.), *Entwicklungspsychologie*. 6. vollst. überarbeitete Auflage. Weinheim, Basel: Beltz, 436-79.

Treagust, D., Duit, R. & Nieswandt, M. (2000). Sources of Students' Difficulties in Learning Chemistry. *Educación Química* 11(2), 228-235.

Wellenreuther, M. (2008). *Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht*. Grundlagen der Schulpädagogik. Band 50. Hohengehren: Schneider.

