



Towards a common understanding of upper-secondary immersion education in Switzerland: Increasing awareness of differing conceptions of bilingualism through a pre-evaluation tool

Submitted in fulfilment of the requirements of the dissertation module of the *Master of Arts: Bilingualism and Multilingualism* at the University of Wales Trinity St David

by

Lynn Williams Leppich, 29001318

on

30th September 2026

Word count (excluding abstract, references and appendices): 14'140

DECLARATION FORM



Name of the Programme of Study: Master of Arts: Bilingualism and Multilingualism

Name of Student (Capital Letters): LYNN WILLIAMS LEPPICH

DECLARATION

I certify that the whole of this work is the result of my individual effort, and that all sources have been acknowledged.

Signed

.....

Lynn Williams Leppich

Date: 30-09-26

DECLARATION

I am satisfied that this work is the result of the student's effort.

Signed:

Dr Hywel Glyn Lewis
(Lecturer)

Date: 30-09-26

1.	Table of contents	3
2.	Introduction	5
3.	Literature review: An overview of models and context	8
3.1	Introduction	
3.2	Immersion education as an important (but vague) pedagogical issue	12
3.3	Paradigms and models shaping the contemporary immersion landscape	14
3.3.1	Additive bilingualism and Common Underlying Proficiency (CUP)	14
3.3.2	Bilingualism as functional competence	15
3.3.3	Sociocultural/ecological frameworks	15
3.3.4	Why competing (strong) models need contextual interpretation	
4.	Research design	16
4.1	Introduction	16
4.2	The rationale for moving from conceptual synthesis to pre-evaluation tool	18
4.3	In favour of a theory of change	19
5	Discussion and critical analysis	22
5.1	Resulting tensions	22
5.1.1	Tension 1: Language separation vs. flexible multilingualism	22
5.1.2	Tension 2: Competence frameworks vs. developmental considerations	24
5.1.3	Tension 3: Translanguaging vs. programme coherence	25
5.1.4	Tension 4: Ideology vs. pedagogy	26
5.2	Why the Swiss case is especially difficult to classify	26
5.3	Variation within the unstable ‘Swiss’ immersion model	28
5.4	Why immersion debates often become polarised	28
5.4.1	Why competing models need contextual interpretation	29
5.4.2	Conceptual fragmentation	30
5.4.3	A lack of evidence on emerging and fragmented conceptualization	30
5.5	Imposing contemporary models on the Swiss educational landscape	31
5.5.1	Immersion and CLIL: Overlapping traditions which cannot be equated	31
5.5.2	Partial immersion and CLIL: A practical distinction	32
5.5.3	Characterising a strong immersion/CLIL lesson	33
5.6	A four-fold categorisation of immersion scholarship in the Swiss context	34
6.	Implications and recommendations	35
6.1	Introduction	35
6.2	School-level	36
6.3	Research-level	37
6.4	Policy-level	38
6.5	Towards a common goal in immersion education: a research-practice agenda	38
6.6	Preparing for a future possible (informal) evaluation	39
6.7	Concluding words on concrete proposals	41
7.	Conclusion	41
	References	
	Appendix	

Acknowledgments

I would like to extend my heartfelt thanks to Dr Hywel Glyn Lewis for his helpful comments and questions, and for practical reminders and guidance along the way.

Abstract

Immersion education has undergone a significant conceptual transformation in recent decades. Whilst early models were grounded in dual-competence conceptions of bilingualism which emphasised language separation and issues of cognitive interdependence, subsequently there was a shift towards functional competence and instrumental multilingualism, especially in the European context. More recently, sociolinguistic approaches like translanguaging have challenged assumptions around bounded language and ideas around strict separation.

By extracting tensions from literature in the field, this dissertation synthesises concepts linked with the competing models of bilingualism underpinning immersion education, and is therefore designed as a ‘state of the art’ desktop research undertaking. It seeks to critically evaluate theoretical foundations, pedagogical decisions/limitations, and any structural weaknesses. Building on this analysis, it proposes a pre-evaluation tool intended to support teachers, leaders, students, policy makers, parents and other stakeholders as they strive towards a common understanding of their own context-specific form of immersion. This is followed by suggestions for future review, development and enhancement of the programme in question, independently of/before a formal evaluation.

Instead of advocating for a prescriptive shift away from language separation and towards a deliberate integration, I argue for more carefully aligned boundary management, which is in turn shaped by the respective sociolinguistic landscape (many and various in the Swiss context), by developmental considerations and by institutional constraints.

The overall aim of this dissertation is to put forward ideas to help sensitise teachers and leaders to potential misunderstandings or conflicts between their declared and actual implementation of immersion at *Gymnasium* (grammar school). It aims to highlight potential areas of tension as well as misalignments between policy and practice. It seeks to offer enhanced clarity in the field, to create a practical resource and to formulate fundamental guidance which can be employed both before/instead of formal evaluation and also in immersion teacher education.

Chapter 2 Introduction

Immersion at upper-secondary level is increasingly established and is in the meantime normalised in the Swiss educational landscape, to the extent that any school considering evaluating their programme is likely to require (and expect) a relatively simple answer to the basic question: ‘Does it work?’ Yet, the terms ‘immersion’, ‘bilingual course’ and ‘bilingual Matura’ are potentially used interchangeably (and therefore inconsistently), meaning that the very first question a school might benefit from asking is in fact: ‘What exactly are we actually evaluating?’ The central argument of this dissertation is about very fundamental, basic tenets of upper-secondary education, turning on the notion that before asking whether an immersion programme is effective, teachers, school leaders and other stakeholders need a sufficiently shared understanding of what immersion is intended to mean in their own specific context.

It is important to observe that this is not an argument against evaluation but instead it is the suggestion that conceptual alignment must be a precondition for meaningful evaluation. Several factors come into play here: regulatory compliance, programme description, pedagogical models, intended outcomes, lived experience and ultimate impact. This dissertation draws on a review of Swiss policy/regulatory documents for this level of schooling and Swiss scholarship in the field, against the background of broader bilingual education, CLIL and immersion literature.

The significance of this combined approach can be found in the notion of a ‘shared-understanding-before-measurement mindset’. That is, the evaluation ‘problem’ is partly a definition problem. The paradox is, of course, that more provision of immersion offerings does not necessarily mean greater conceptual clarity. As Brohy (2004) discusses, the nomenclature around bilingual education is problematic and in itself thought-provoking. ‘Immersion’ as a term, for example, can variously denote:

- the percentage of teaching in the L2;
- teaching particular subjects through the L2;
- whole-school / whole-year immersion;
- exchange/host school models;
- bilingual Matura as credential;
- pedagogical approach;
- programme identity/marketing term.

This means, essentially, the implication that two schools can equally designate their programmes as ‘immersive’ but the programmes themselves might differ in aspects including but not limited to: the proportion of L2 teaching, the specific subjects taught through the L2, teacher linguistic competence, pedagogical practices, expectations regarding student proficiency, assessment, entry/selection and intended outcomes (after e.g. Carroll et al. 2007; Durlak & DuPre, 2008). Even the related ‘CLIL’ suffers from a lack of conceptual clarity. CLIL itself has been defined so broadly that its boundaries are unclear (Dalton-Puffer, 2011). Programmes differ in their language goals, intensity, student population and the extent to which teachers explicitly offer language input and instruction.

Consequently, if an immersion or other CLIL-related bilingual education programme at upper-secondary level were to be evaluated, what would the evaluation in fact be evaluating? It is the central premise of this dissertation to explore how far a meaningful evaluation of immersion at upper-secondary level requires a prior (or at the very least) a concurrent process of conceptual clarification. During this process, teachers and school leaders alike would establish a shared – though not necessarily identical – understanding of the programme’s model, purpose(s), mechanisms and expected outcomes.

Even a cursory examination of relevant curricula documents reveals that, in the Swiss policy and institutional context, immersion is regulated but not fully conceptually anchored (Elmiger et al., 2022). There are essentially three strands to consider in this connection: federal/cantonal complexity, regulatory definitions/thresholds, and implication(s?) for evaluation, each of which will be addressed in turn in the discussion that follows.

Federal/cantonal complexity

The first aspect concerns Swiss educational governance, which can be further consolidated into the dimensions of cantonal responsibility, coordination by the EDK (the Swiss Conference of Cantonal Ministers of Education), and regulation and accreditation mechanisms at the level of the federation and of the SMK (that is, the Swiss Baccalaureate Commission), all against the backdrop of significant cultural, pedagogical and curriculum differences between language regions (EDK, 2024; SMK/SBFI, 2025). As a result, there is in effect no such single concept as ‘Swiss immersion’; instead, the offerings under this umbrella term represent a range of different educational initiatives and are subject to considerable institutional variation. This in turn means that stakeholders involved in the field – whether as leaders, teachers, students, parents or policy makers must take context into account.

Regulatory definitions/thresholds

In order to be accredited, Swiss immersion programmes at upper-secondary level are subject to a 2012 SMK regulation which specifies that the recognised immersion language is a Swiss national language or otherwise English and that the language in question also features in the respective curriculum as a *Grundlagenfach* (compulsory foundational subject) or a possible *Schwerpunktfach* (elective main subject). Furthermore, a minimum of three assessed subject areas are to be taught in the immersion language, including one humanities or social science subject. The total number of immersion lessons should equate to 800 learning hours and – somewhat confusingly but also understandably, so that the programme does indeed offer a balance of languages – there is a limit on the overall proportion of learning in the immersion language.

For administrative and accreditation purposes, the above are certainly useful criteria. However, it goes without saying that criteria of this nature do not necessarily lead to a complete pedagogical definition. Quantitative threshold data can usefully flesh out the picture, by capturing how much immersion, in which subjects and in which language but they do not establish as a matter of course how immersion is implemented, how language and content interact, what teachers understand their role(s) to be and what students experience.

The subject of the dissertation, then, remains relevant and is in fact perhaps particularly timely with regard to a current transition which is afoot. The new MAR/MAV framework curriculum (2024 version) along with new SMK guidance for the bilingual Matura effective from 1st January 2026 (SMK/SBFI, 2025) stipulates that, in order to teach immersion classes, teachers must have attained C2 language competence (CEFR), participate in continuing professional development in immersion and also have appropriate opportunities for professional exchange.

Implication(s?) for evaluation

As can be derived from the clear tension which results between, on the one hand, rigid accreditation criteria and, on the other hand, (in true Swiss democratic style) the freedom to fulfil those criteria in a range of different ways, questions of compliance are necessary but will not be enough to offer a definitive definition of Swiss upper-secondary ‘immersion education’. It is therefore essential that we further distinguish the basic dimensions established in the following angles of ‘immersion education’.

- Structural definition – facts of timetabling, subjects, languages;
- Pedagogical definition – declared and enacted teaching and learning processes;
- Outcome definition – observed language, subject, intercultural, motivational outcomes;
- Organisational definition – principles and processes around staffing, leadership, collaboration;
- Equity definition – aspects concerning access, i.e. who enters, who succeeds, who is excluded.

Bearing in mind the issues raised, here, this dissertation engages overall with the central premise that whilst accepted scholarship, policy and best practice in a field have their undeniable place and importance, context is paramount (Braun et al., 2011), writing on policy enactment at secondary school in general): In this case, ‘immersion’ does not describe a single, uniform educational intervention, meaning models and practices cannot automatically be translated from one setting to another. The present dissertation seeks then to review the current Swiss immersion landscape to locate tensions and therefore also possible sites of negotiation and is structured as follows: This context chapter establishes the setting and places some critical questions which drive the discourse in this field, before the literature review both surveys the general direction taken by the scholarship and engages critically with key arguments. The methodology section details and rationalises the approach taken: a state-of-the-art conceptual synthesis. A fifth chapter discusses findings from the review with the subsequent chapter summarising the concerns of the dissertation as a whole, before proposing recommendations for policy and practice as well as noting limitations and possible future directions.

Chapter 3 Literature review: an overview of models and context

3.1 Introduction

Plurilingualism is a matter of national pride in Switzerland. Actively fostering multilingual freedom and promoting linguistic equality are important cornerstones of Swiss life (Kuzelewska, 2016). In connection with this notion of multilingual freedom, ‘while Swiss language policy is determined on the federal level, it is still in [sic] actual matter of practice

for cantonal implementation' (Kuzelewska, 2016, p. 131). The resulting implementation is key because, as Schneider & Ingram (1993) observe, 'policies that fail to solve problems or represent interests and that confuse, deceive, or disempower citizens do not serve democracy (Schneider & Ingram, 1993, p. 345).

On the surface, this democratic and egalitarian approach to language strategy may seem ideal. However, some scholars express concerns. Berthele (2021), for example, perceives a 'selective celebration of linguistic diversity' (Berthele, 2021, p. 125). Hega (2001) also comments on the possible consequences of existing attitudes to and the resulting policies and practices in Swiss language education, especially against the backdrop of national decentralisation and direct democracy, noting the 'slowness' of policy making in general (Hega, 2001, p. 221). Hega even goes as far as to suggest that certain issues such as the introduction of formalised second language learning at primary school level 'illustrates how Switzerland's cultural and linguistic diversity has frustrated past attempts at policy coordination and continues to hamper recent reform initiatives' (Hega, 2001, p. 205). Lastly, and of greatest relevance to the present area of study, it is the imprecise and indiscriminate use specifically of the term 'immersion' which Cathomas (2011) objects to, leading him – understandably – to argue for a clearer typology. From his argumentation, it emerges that the label 'immersion' does little to help any interested party understand what is meant in a particular context.

Against this backdrop in the general Swiss upper-secondary educational landscape, bilingual immersion education at upper-secondary level might and – it can be argued – first of all *must* be understood as essentially a raft of organisational arrangements, reflecting the politically-informed federal structure of Swiss education. Whilst cantons and individual schools retain substantial responsibility for implementation, federal and inter-cantonal regulation establishes the conditions under which a Matura may carry a multilingual designation (Elmiger et al., 2022; SMK, 2012). Swiss plurilingual education has historically displayed considerable variation in the geographically and politically defined (first) target language as well as starting age, compulsory or optional status, and intensity, with the result that seemingly innocuous labels like 'bilingual', 'immersive' or 'plurilingual' can conceal significant local diversity (e.g. Brohy, 2004; Berthele, 2021).

Specifically in the Swiss educational context, immersion needs to be understood against a linguistic landscape in which multilingualism is not an exceptional educational objective, but

which is instead institutionally anchored in the curriculum itself. At overall *Gymnasium* level, as embedded in the national framework curriculum, the *Rahmenlehrplan* (RLP) as stipulated by the EDK, students are expected to study a first and second national language, as well as a third language, which may be another national language, English or a classical language and languages consequently occupy a substantial portion of curricular time (EDK). In moving foreign language learning beyond the language classroom, immersion represents a significant shift within this already multilingual landscape. The additional language of English (or, depending on geographical location, also French, German or Italian) becomes not simply students learn about as content but a language through which they also acquire disciplinary knowledge. This means that bilingual education is especially well connected with the Swiss tension between linguistic plurality and the increasingly international orientation of upper-secondary education (Williams Leppich, 2026).

This is perhaps especially in evidence at *Gymnasium* (grammar school) level, where bilingual education has become institutionally established due to the instigation of the bilingual Matura (the Swiss educational diploma which grants school leavers direct access to university study). In Zurich, for example, the introduction of German-English bilingual programmes starting in the 2001-2 academic year saw standard English teaching combined with immersion in subjects such as History and Maths, allowing language development to take place *alongside*, rather than separately from, academic learning in the disciplines (Stebler & Maag Merki, 2010).

The significance of this innovation is of course not simply that students in this particular Swiss content benefit from more input and time in English. Instead, it reflects a reconfiguration of what counts as language learning in an academically selective environment. Linguistic competence necessarily becomes embedded in the practices of studying, reasoning and communicating within a subject discipline. The subsequent spread of bilingual programmes is indicative of how this model has moved from being an experimental innovation towards being structurally important and therefore also a recognised (and valued) component of Swiss *Gymnasium* education. Most recognised *Gymnasien* now offer a bilingual programme, whilst Elmiger et al.'s (2022) national inventory documents bilingual provision across all of the grammar school, vocational and other upper-secondary sectors. In 2022, 185 programmes were identified at grammar school level alone; this expansion is significant because it means that Swiss 'immersion' is no longer best regarded as a fleeting experimental innovation on the periphery of school life.

Still, precisely because participation occurs within the (anyway) selective *Gymnasium* content (in turn raising issues around access and social justice), positive outcomes cannot automatically be interpreted as effects of immersion alone. Students enrolling on a bilingual programme may already stand out in terms of linguistic ability, motivation or educational orientation, meaning the relationship between immersion and achievement has to become as much question of educational processes (e.g. more focus on pluriliteracies, Beacco et al., 2016; Llinares et al., 2012) as one of measurable language gains (Stebler, 2010).

Current regulatory terminology is also revealing, as is the lack of national school-level data on immersion classes (SKBF, 2026). Since January 2026, the Swiss Matura Commission (SMK) has defined ‘immersive teaching’ as the teaching on particular non-language subjects (‘Sachfächer’) in a language other than the school’s standard language of instruction, unlike the OECD who use ‘immersive’ also as a designation for learning embedded in technology. The academic objectives of those subjects remain those of the standard curriculum, with the additional language (the L2) functioning as a medium through which disciplinary knowledge is acquired (e.g. Bieri, 2018; Ting, 2024). Further, the regulations also envisage increased competence in the immersion language, with C1 as an intended outcome. In conceptual terms, this is important: Swiss immersion does not normally replace subject education with language education. History or Biology taught in an L2 remain intact in terms of content matter. Language learning is produced, at least partly, through sustained participation in disciplinary activity (Lindholm-Leary, 2012).

In recent decades there has been growing international interest in immersion education and this expansion in turn has generated both several co-existing models and overlapping and fuzzy/unstable terminology. Scholarly debates have become increasingly polarised (see e.g. Berthele, 2011). These debates often overlook the importance of local socio-linguistic and policy contexts. Switzerland in particular offers up an especially rich backdrop, combining as it does institutional multilingualism with highly diverse educational settings and therefore exemplifying many of the relevant tensions in one education system. As a result, there is a need not for another effectiveness review but instead for a state-of-the-art conceptual synthesis which maps the competing frameworks, identifies areas of convergence and divergence, and proposes a more context-sensitive conceptualisation of immersion education.

The following aspects will be discussed in turn below, in sections 3.2-3.4 inclusive, as the line of argument develops through this literature review:

3.2 Immersion education as an important (but vague) pedagogical issue

3.3 Paradigms and models shaping the contemporary immersion landscape

3.4 Competing models in need of contextual interpretation

3.2 Immersion education as an important (but vague) pedagogical issue

As Baker (2017, pp. 239-240) summarises, the many advantages of bilingual education are as follows:

- higher proficiency in two languages, supporting broader communication;
- deeper intercultural understanding and sensitivity;
- enhanced biliteracy, increasing career and cultural opportunities;
- improved academic outcomes, as evidenced in various bilingual models;
- cognitive benefits for age-appropriate bilingual learners;
- increased employability in a globalised world;
- stronger personal and cultural identity.

Across Europe and beyond, immersion education has become one of the most influential approaches for promoting multilingual competence whilst maintaining academic achievement. However, despite decades of research (see e.g. Wright and Baker, 2025, for a historical survey), there remains little consensus about what counts as immersion or how it differs from CLIL (Lasagabaster & Sierra, 2010), about bilingual education or multilingual pedagogies, and about which models are most appropriate in certain contexts. As immersion education has expanded into increasingly diverse educational settings, conceptual clarity has in turn become more rather than less (Cathomas, 2011).

Though immersion education is frequently presented as an especially effective response to the demands of multilingual and internationally oriented education, precisely what is meant by the term remains surprisingly unclear. Historical traditions of bilingual education have resulted in somewhat competing conceptions of the relationship between language and content (and around how much language capacity is at all cognitively possible or desirable). One tradition will place greater emphasis on the acquisition of subject knowledge by means of extensive and sustained exposure to the additional language, the L2 (e.g. Thomas & Collier, 2002), whilst another will prefer to foreground the systematic development of linguistic competence alongside disciplinary competence.

Contemporary approaches like Content and Language Integrated Learning ('CLIL') have further complicated this distinction by positioning language and content as interconnected but not necessarily equally explicit pedagogical objectives and have traditionally been seen as belonging more to the language classroom (to lend language learning weight, depth and purpose) than to the subject classroom. These two different traditions carry with them different assumptions about what teachers should teach, what students should learn and crucially about who is responsible for language learning (Coyle & Meyer, 2021). It is therefore perhaps rather unsurprising that discussions around immersion can quickly become polarised: Apparently similar programmes may be premised upon ultimately quite different understandings of bilingualism, language learning and effective pedagogical implementation.

Distinctions of this nature can become especially pertinent when immersion is examined in context. Pedagogical models do not simply transfer unchanged from one linguistic, institutional or historical setting to another. Instead, they are interpreted, filtered and enacted by teachers working within specific curricular frameworks and constraints, school structures and professional and social cultures. This is perhaps especially true of Switzerland, where multilingualism is constitutionally and therefore also educationally embedded but where at the same time upper-secondary immersion programmes operate within the extremely varied context of the *Gymnasium*. Swiss programmes may be at least formally designated late partial immersion, but the educational reality of these programmes might be significantly closer to CLIL, with subject teachers understandably prioritising their disciplinary objectives over any linguistic goals and also varying in the extent to which they identify as (also) educators in language.

This of course raises a central question: What happens when the policy conception of bilingual education and the pedagogical conception held by teachers do not fully align? More concretely, how do teachers currently understand bilingualism, immersion and their own responsibility for language development (which is additionally anchored in the curriculum for any school subject at *Gymnasium* as 'sprachsensibler Unterricht', EDK, 2024). Instead of treating the teachers as some kind of 'problem', it would be worth approaching this discrepancy as a question of professional knowledge, institutional conditions, policy communication and educational design decisions. If teachers have only limited access to explicit linguistic concepts for analysing and supporting language development within their subjects, expecting immersion programmes to realise a more ambitious conception of bilingual education may be unrealistic. The role of school leadership is therefore also

important. Teacher readiness is unlikely to develop purely through individual responsibility but in fact requires structures that validate professional learning, collaboration and continuing teacher education.

This dissertation therefore explores how evaluative questions might be used not simply to measure the ‘success’ (or otherwise) of existing programmes, but also to make their underlying assumptions visible. Examining teacher attitudes, for example, alongside enacted pedagogy can reveal where declared programme aims align with or diverge from what actually happens in the classroom. An approach like this might in turn support more targeted teacher education (Tedick, 2020) as well as stronger institutional provision, whilst at the same time also allowing schools to (re-)consider whether and how far their current practices adequately reflect the bilingual models they claim to offer.

For the field in general, clarifying the above distinction (and all the shades on the spectrum in between) matters because it can be tempting, in immersion research, training and professional exchange, to treat programmes as ‘stable’ educational entities with effects which can be identified and evaluated essentially independently of their implementation in the local setting. A context-sensitive perspective instead asks what *kind* of immersion is in fact being enacted, by whom, under what conditions and according to which conceptualisation of bilingualism. Bringing questions of this nature together may give us a more useful way of understanding variation in immersion provision, which in turn might help us close – or at least narrow – the gap between policy declaration and educational realities.

3.3 Paradigms and models shaping the contemporary immersion landscape

This section outlines existing paradigms in bilingual education and related competing models plus resulting pedagogies to facilitate an understanding of significant concepts, designations, policies and practices Swiss immersion education is informed and shaped by. This should go some way towards explaining why conceptions of immersion might become polarised.

3.3.1 Additive bilingualism and Common Underlying Proficiency

The central idea behind this paradigm is the very foundation of immersion education, i.e. that learning through another language does not threaten or diminish first-language development; rather, the two languages support one another cognitively. Instead of seeing languages as

competing systems, then, according to this view, bilingualism is considered to enhance cognitive development when both languages are supported, hence the designation ‘additive’ rather than ‘subtractive’. Key models drawn on include Cummins’ (1979) linguistic interdependence hypothesis. Whilst this paradigm is still well received, some later scholars argue that Cummins’ model still assumes relatively separate language systems and can (therefore) underplay the fluidity of multilingual practices (Wright & Baker, 2025), a deficiency which is neatly addressed through the application of contemporary approaches like translanguaging (see 3.3.3).

3.3.2 Bilingualism as functional competence

This stance puts the idea of preparing citizens for multilingual societies firmly centre stage. Seen from this perspective, languages become employability skills, mobility resources and tools for intercultural communication, instead of primarily cognitive assets and languages are clearly being framed as economic resources for integration in Europe. Concepts around communicative and intercultural competences, plurilingual/multilingual competences, multilingual citizenship, mobility, lifelong learning and human capital are key (see e.g. Council of Europe, 2001; Bakken and Brevik, 2023; Beacco et al. 2016; Marsh, 2002;).

3.3.3 Sociocultural/ecological frameworks

The idea of possible sociological and ecological facets to language learning represents a significant conceptual shift in recent decades (Wright and Baker, 2025). The view is that of participation and mobility with the goals of intercultural competence and inclusion, identity, equity and participation (e.g. Bakken & Brevik, 2023; Fielding, 2022). According to this paradigm, multilingual speakers draw on one integrated communicative repertoire, instead of two or more separate languages (García, 2009), with the emphasis on moving from competence to practice, and from knowledge to application. The overall stance is that languages are purely social constructs and the mindset is one of empowerment, identity and belonging (Cenoz & Gorter, 2015; Gajo, 2007; Garcia and Wei, 2014; Hornberger, 2003). Critics of very pronounced sociocultural/ecological models will argue that target language exposure still matters and that translanguaging, if unchecked and dominant, may reduce opportunities for second language development, a discussion point which is becoming

increasingly prominent in immersion research (see e.g Lyster, 2007; Tedick, 2020).

3.4 Why competing (strong) models need contextual interpretation

Much evidence in support of strong bilingual models such as those captured above is solid and compelling. In one study Thomas and Collier (2002) established that students enrolled in dual language programmes were more successful than their monolingually-educated peers in both English and academic achievement tests, an outperformance margin which also endured long-term (Thomas & Collier, 2002). Similarly, Genesee et al. (2006) also offer a synthesis of research which demonstrates that strong bilingual models foster metalinguistic awareness, cognitive adaptability and enhanced academic performance across subjects. Frequently employing control groups, longitudinal data and a range of assessment tools to ensure reliability, such studies offer pleasingly robust insights into the success of strong bilingual models.

It is also clear that the major models in bilingual education are characterised by individual opportunities and challenges and as a consequence, can compound and complicate noticeable tensions in immersion education. In that connection, a later chapter, 5, will build on the groundwork established here to explore a selection of those persistent tensions, seeking above all to engage with authentic and legitimate concerns, rather than with distorted or oversimplified aspects of these.

Chapter 4 Research design

4.1 Introduction

The research methodology employed in this dissertation adopts the idea of a conceptual synthesis (see e.g. Grames et al., 2022; Yazdani & Bazadini, 2020). The research goal is not to offer a meta-analysis of immersion literature relevant to the Swiss context nor, within the scope of this paper, already to evaluate immersion programmes for effectiveness (although this clearly constitutes a desirable next step). Instead, conceptual traditions are synthesised in an attempt to identify how and where variation in understandings can occur and subsequently to

propose a pre-evaluation strategy which functions as a bridging tool between analysis and evaluation.

As the introduction has indicated and as later chapters of this dissertation will show in more detail, the terminology used in the Swiss immersion context at upper-secondary level has become fragmented and it would make sense to attempt a recalibration – or at the very least, a surfacing of any misunderstandings, assumptions and inherent challenges. To serve this purpose, the research question can be articulated as follows: *What tensions emerge in the conceptualisation of immersion education in the Swiss context across different scholarly traditions? How might a pre-tool help draw out assumptions and consciously align thinking ahead of a formal evaluation?*

Going beyond the literature review of Chapter 3, then, is further engagement with the literature in the form of the document review discussed in Chapter 5, the rationale being that perhaps not enough reference is often made to somewhat obvious sources, for comparison with implementation in educational settings, and with the overall aim not of confirming a theory, but of generating new ideas from freely available sources. As McCulloch (2017) observes: ‘[d]ocuments are a significant and often underused resource for research in education’ (McCulloch 2017, p. 220) and ‘[a]nalysis aims to generate new accounts or explanations out of this rich data, rather than test existing theories’ (Richards, 2025, p. 1). Using readily available sources also precludes the need for any major administrative or logistical arrangements and providing the content of sources and related data are reported correctly and interpreted fairly and with due care, no ethical issues should arise.

The engagement with the literature was organised around several strands and sources. Firstly, curricula and policy documents were consulted for information and also for references. These were then rounded out by a consideration of institutional-level course materials and readings. In both cases, the snowball technique (Hart, 2018; Ridley, 2012) was then used to search for further sources which could be of relevance. A highly pragmatic and selective open coding approach was adopted to capture useful themes and patterns which emerged in this continuing comparative work. Review was pragmatic and selective and coding was inductive and partly in-vivo, partly descriptive to suit the qualitative and creative nature of this dissertation. Saldaña (2016) cautions that contradictions and tensions necessarily feature in qualitative inquiry, frequently brought about by researchers seeking to minimise data but at the same time to use dense description, and therefore suggests a pragmatic and eclectic approach: ‘We do not need to reconcile the contradictions; we only need to acknowledge their multiplicity and become well versed in various methods and modes of working’ (Saldaña, 2016, p. 289). The notion of

a ‘heuristic fluidity’ (Saldaña, 2016, p. 9), that is of being flexible as and when necessary and advisable, suits the intentions underlying this research. As Saldaña (2016) also notes, ‘[q]ualitative researchers are not algorithmic automatons. If we are carefully reading and reviewing the data before and as we formally code them, we cannot help but notice a theme or two (or a pattern, trend, or concept) here and there’ (Saldaña 2016, p. 16).

The idea of this dissertation, then, is not to generate new data but rather to examine sometimes neglected existing sources and to distil an original interpretation from it which can lead to the creation of a tool with potential future application in practice.

4.2 The rationale for moving from theoretical synthesis to evaluation tool

At the intellectual centre of any evaluation of an immersion programme has to be what happens upon implementation rather than any label (Ball, 2012; Brohy, 2004). The question, then, is rather what a school’s stated immersion model becomes once teachers actually teach it. If we want to formally evaluate an immersion programme, student language or subject scores alone are not enough (Stufflebeam & Zhang, 2017, on the CIPP Evaluation Model, which is framed around the four equally important pillars of Context, Input, Process and Product). A programme may produce strong outcomes whilst still being organisationally fragile. At the same time, a well-established programme may exist institutionally without producing consistently strong classroom practice. Implementation quality therefore needs to be separated from programme outcomes (Jerrim & de Vries, 2017).

This paper proposes an observation framework based around five constructs:

- *Content integrity*: Is the intellectual level of the subject in question equivalent to that expected in non-immersion classes? This corresponds directly to the current Swiss requirement that immersive subjects pursue the same subject objectives as their monolingual counterparts (Elmiger et al., 2022; SMK, 2012).
- *Language integration*: Does the teacher deliberately identify and support the vocabulary, genres and discourse structures required for disciplinary thinking?
- *Scaffolding*: How are visual resources, modelling, reformulation, task sequencing and multilingual resources employed so as to make difficult content accessible without reducing cognitive, conceptual demand?
- *Student cognitive and linguistic production*: What student learning is in evidence, both in subject matter and language production?

- *Strategic multilingualism*: When and why does the class move between languages? The evaluative question is not just ‘Does the teacher avoid the L1?’ but whether language choices facilitate or obstruct disciplinary learning (see e.g. Bieri, 2018).

It seems important not to turn these constructs into a simplistic checklist of ‘good CLIL practice’. Rather, observation should establish patterns first: For example, do, say, three locally developed models co-exist within a single school? (a highly-language conscious CLIL, largely monolingual content teaching in English and a kind of third pragmatic hybrid form of the two, for example?). Variations of this nature can in themselves point towards further learning around policy implementation. Any related evaluation needs to examine the whole chain from policy intention (see Braun et al., 2011) to classroom enactment. Possible domains could include:

1. Does the school have a shared understanding of what immersion is supposed to (or can) mean?
2. Capacity: Staffing, expertise, training, materials, coordination
3. Classroom enactment
4. Student experience
5. Outcomes
6. Long-term sustainability

Together, these domains allow us to evaluate immersion as a system rather than as a single intervention and represent features which distinguish it from a language classroom, where use of CLIL methodologies may be sporadic and unsystematic.

Overall, it is worth bearing in mind, as Baetens Beardsmore (1993) also emphasises, that all efforts within the framework of bilingualism should be firmly focused on educational goals, with multilingualism only to be considered a useful extra that is not to put other academic objectives at risk. Further, it goes without saying that bilingual education offerings are carefully conceived, enacted with sensitivity and consistently monitored (Genesee et al., 2006). Evaluating a bilingual programme demands robust and plausible evidence, as well as – crucially – common alignment as to how its success is defined.

4.3 In favour of a theory of change

The central premise of this dissertation is to establish the variation embedded within the designation ‘immersion’ and to use related discussions around terminology as an access point

to examining how far institutional policy is translated into (enacted in) classroom practice, meaning under what conditions (Durlak & DuPre, 2008), for whom, to what end and with what consequences. This distinction clearly matters because implementation research shows that teachers do not mechanically enact policy (Spillane et al., 2002) and bring their own beliefs to any educational setting (Borg, 2006). Spillane et al. (2002) characterise implementation as a form of ‘sense-making’ whereby teaching staff interpret new policies through their existing beliefs, knowledge, routines and organisational contexts. In this regard, variation among teachers cannot be considered as ‘disobedience’ but instead as symptomatic of how policy works.

In the Swiss *Gymnasium* setting, a fundamental causal chain could be depicted as follows:

policy intentions > organisational arrangements > staff understanding and capacity >
classroom enactment > student experience and outcomes > institutional sustainability

It is clear that a chain like this firmly places the focus on purpose: By concluding observations in the area of institutional sustainability, we return to ideas around overall context-based mission or policy intentions. A design cycle of this nature can help prevent a common evaluation error which can be illustrated in the following way: Even if immersion students attain strong L2 results, this alone does not corroborate the success of a school’s immersion programme, especially in a world where young people are increasingly developing their English language skills through social media, streaming services and gaming arrangements (OECD, 2024; Williams Leppich, 2026). To contextualise: students may have brought stronger English from the outset, (and possibly resulting from this) greater academic confidence as well as more privileged backgrounds or stronger motivation upon entry to the immersion programme (e.g. Bakken & Brevik, 2023; Lorenzo et al., 2021; Prusse & Nakamura, 2010). By the same token, only modest language gains would not necessarily indicate programme failure if disciplinary achievement, confidence in academic language or general participation had improved.

The policy intention end of the chain does, however, needs to be precisely defined. Current Swiss Matura guidance requires the same disciplinary objectives in immersive and standard subject teaching (SMK, 2012) whilst at the same time striving for increased competence in the immersion language, with C1 (Elmiger et al., 2022) the intended level. This means that any evaluation should refrain from considering linguistic progress to be the sole dependent variable. It is therefore crucial that it is the system as a whole which is evaluated, rather than

individual students (though this layer clearly also contributes). Basic guiding questions to frame a larger evaluation therefore include:

- What does the school (and, by way of a sub-question: ‘Who’ actually is the school, i.e. which stakeholders are involved in conceptualising its purpose and articulating its mission?) believe its immersion policy is intended to achieve?
- How do teachers understand, interpret and implement that policy?
- What experiences and outcomes are associated with that context-specific teacher implementation?
- What organisational conditions might go some way towards explaining any variation [in implementation and outcomes]?

The proposed evaluation question catalogue (see Appendix) is built upon a framework of the six domains outlined above, all of which interact with each another, and it explicitly seeks to distinguish implementation quality from programme outcomes. The first three domains address the issue of how teachers implement policy. Classroom enactment, especially, should examine whether teaching genuinely integrates disciplinary and linguistic learning. Dalton-Puffer’s (2011) research can here helpfully direct attention towards classroom discourse: Who speaks, what kinds of academic language students are required or encouraged to produce and how disciplinary knowledge is constructed in an additional language.

Domain 6 draws on the idea common in evaluation literature that successful reform should be examined for depth, sustainability, spread and shift in ownership. For example, a bilingual programme which survives only because two exceptionally committed teachers sustain it may produce excellent lessons but remain fragile at the organisational level. By contrast, though one programme has spread through a school, it may not necessarily have produced deep or lasting changes in classroom practice.

Chapter 5 Discussion and critical analysis

5.1 Resulting tensions

The literature on bilingual education examined within the context of the initial and extended literature reviews of Chapters 3 and 4 delineates and explores a range of tensions which can result from differing visions, expectations, structural arrangements and understandings in the area (Wright & Baker, 2025), some of which are examined in depth below by virtue of their relevance to Swiss upper-secondary immersion education.

5.1.1 Tension 1: Language separation vs. flexible multilingualism

Essentially, at Swiss Gymnasium level, ‘immersion’ usually denotes a late partial immersion programme which sees (only) selected subjects being taught through an additional language. Actual classroom practice consequently overlaps strongly with European CLIL, although immersion and CLIL are not theoretically identical terms. The scholarship anchors this distinction, for example Brohy (2004) and Elmiger et al. (2022) for the specifically Swiss institutional context, Stebler and Maag Merki (2010) for the evaluation of Swiss Gymnasium-based immersion and Bieri (2018) for recent upper-secondary CLIL-classroom research.

Inside the classroom, immersion can be distinguished by its curricular insistence (EDK, 2024; SMK, 2012) on content, which ultimately remains primary, as it is this which is graded in the Matura. Disagreement and/or misalignment in the conceptualisation and delivery of this kind of learning can result from differing stances on the balance of mother tongue (L1) and second language (L2) in a certain context. With respect to the issue of whether immersion should maximise target-language exposure or whether it should instead aim to leverage learners’ full linguistic repertoires, those in favour of separation may argue that by doing so they:

- create clear programme identity;
- ensure sufficient L2 input;
- avoid dominant language takeover;
- (and therefore, where relevant, also) protect minority languages.

At the same time, valid arguments against separation can also be put forward, including the fact that the resulting inflexibility of natural language use:

- is not reflective of authentic multilingual practice;
- neglects learners' linguistic resources;
- may disadvantage multilingual students;

and that

- translanguaging can support deeper conceptual learning.

Successful immersion makes language visible. Calling the typical classroom CLIL-like does not mean that every Swiss immersion teacher follows an identical methodology. There is no single universally imposed lesson format. Instead, a recurring problem in the research is how teachers can preserve the intellectual challenge of the academic subject in question whilst make the linguistic means required for participation sufficiently accessible. Brohy (2004) therefore emphasises not merely subject and language outcomes but the importance of scaffolding, learners' developing interlanguage and the actual plurilingual practices through which understanding is constructed.

Alongside practices like translanguaging, one influential international framework is Coyle's (2002) 4Cs: Content, Communication, Cognition and Culture. It asks teachers to plan not simply what disciplinary content students should know, but what thinking they must engage in, what language makes that thinking possible and how learning relates to wider cultural or intercultural perspectives (see also Borg, 2006). The framework remains highly influential in contemporary CLIL scholarship. It is best regarded, however, as a pedagogical framework available to immersion teachers rather than as the single prescribed Swiss immersion model.

Recent Swiss research reinforced this distinction. Bieri's (2019) study of English medium Biology at a Swiss upper-secondary school examines precisely how disciplinary language operates in CLIL lessons. Her analysis draws attention to technical vocabulary, multilingual resources, multimodality and translanguaging (that is, the strategic mobilisation of more than one linguistic resources while learners are negotiating demanding subject-specific meaning.) This matters because 'immersion' is sometimes mistakenly imagined as the mechanical enforcement of English-only or French-only interaction. Contemporary research gives a more complicated picture: purposeful use of another language can sometimes scaffold disciplinary understanding rather than undermine immersion.

The most defensible pedagogical interpretation is therefore content-led integration. The subject curriculum supplies the conceptual spine whilst the immersion language is the

principal medium of disciplinary activity, and effective immersion teachers provide explicit linguistic scaffolding when students need it (Bieri, 2019). This differs from an ordinary foreign language lesson organised around grammar and vocabulary as well as from an idealised sink or swim model where students are simply exposed to another language and expected to cope.

5.1.2 Tension 2: Competence frameworks vs. developmental considerations

Whilst historically, bilingual education paradigms centred on language separation, much contemporary scholarship represents an expanded picture of multilingualism, which features an integrated repertoire instead of a syllabus of finite and purely disciplinary discrete competences (Hess & Williams Leppich, manuscript submitted for publication). One especially significant feature of strong models is their pedagogical approach.

Translanguaging, for example, is increasingly acknowledged not only as a valid classroom strategy in bilingual classrooms, but also as one with potential for impact (García & Lin, 2017). The kind of holistic approach to multilingual education is one supported by Cenoz and Gorter (2011), who, like García and Wei (2014) advocate for multilinguals being encouraged to take advantage of their complete linguistic repertoire in learning situations in the interests of deeper understanding and authenticity and, in so doing, calling into question the premise that the languages are bounded constructs.

Similarly, Cummins' (2007, 2017) interdependence hypothesis proposes that knowledge acquired in one language can foster learning in another, a conclusion that clearly casts doubt on the assumption that it is the strict separation of languages which assists academic success. Empirical studies based in recent and contemporary content and foreign language classrooms consolidate the notion that multilingual practices can support comprehension and participation alike (Illman & Pietilä, 2018) and also illustrate the pragmatic approach open to teachers concurrently navigating multilingualism and institutional limitations (Pitkänen-Huhta & Mäntylä, 2021). At Swiss upper-secondary level, however, more work needs to be done in the Gymnasium classroom, so as to establish how far notions of this nature and the related practices have taken root.

5.1.3 Tension 3: Translanguaging vs. programme coherence

This especially timely debate in the discourse in bilingual education is especially pronounced, perhaps, in multilingual Switzerland. Proponents of translanguaging and related cross-linguistic and inclusive classroom practices argue for its contribution to deeper meaning making, authentic multilingual practice and refreshingly flexible pedagogy (e.g. Meyer & Coyle, 2021). Representing the counter-argument, scholars and practitioners alike raise concerns related to reduced L2 exposure, inconsistent implementation, teacher uncertainty around linguistic boundaries and challenges for assessment design, implementation and evaluation.

In the Swiss context, we can observe a research trend towards translanguaging which is currently somewhat in conflict with teacher beliefs. It is clear from more general research in teacher cognition that the interplay of beliefs, knowledge and contextual challenges can result in complex demands on educators (Borg, 2006). This already difficult situation can be potentially compounded by teacher insecurities around authenticity and guaranteeing students sufficient encounters with the target language (Hess & Williams Leppich, manuscript submitted). A teacher may, for example, claim to believe in the value of multilingual practices in the classroom whilst still operating according to an essentially monolingual model. Erling et al. (2026) recently demonstrated in this connection that this tension is not limited to practicing teachers: Even teachers in training may view translanguaging and similar practices positively and still express reservations around actually implementing these in lessons.

At this juncture, it must be noted that the bilingual Matura framework stipulates successful language development in the additional language as an indicator of level-appropriate immersion education at upper-secondary level (SMK, 2012). As a result, teachers may deem high levels of sustained language use in the L2 essential as regards programme fidelity and legitimacy. The belief-practice tensions (Borg, 2006) in evidence in teacher attitudes to contemporary practices like translanguaging (in theory: approval, in practice: disregard or outright rejection) may in fact be rooted in (context-specific) structural constraints and therefore symptomatic of a possible need for leaders and policy makers – ideally in collaboration with teachers, students, parents and any other interested stakeholders) to review their current policy as stated, reflected or implied in the respective curriculum.

5.1.4 Tension 4: Ideology vs. pedagogy

It would seem today that the design and delivery of much bilingual education is shaped somewhat less by research and more by ideology, Wright and Baker (2025) observe that CLIL, for example, is not just an educational or linguistic strategy; in fact, it also fosters political goals, supporting the EU vision of a global economy peopled and served by a workforce able to deliver and perform in multilingual capacities. Notions around national identity, economic competitiveness, language maintenance, English prestige, social integration and (see Bakken and Brevik 2023), social mobility and international orientation (e.g. de Meija, 2002) each therefore necessarily produce differences in immersion programme design. As a result, immersion models cannot be deemed ideologically neutral (May, 2014).

This stance, however, is complicated by the official position of multilingualism in Switzerland. Whilst English can be labelled a global lingua franca and thus also a significant facet in promoting academic mobility (Williams Leppich, 2026), it does not rank alongside German, French, Italian or Romansh as a national language. This affords English a rather undefined role – though it is not in the traditional sense a ‘minority language’, its promotion at institutional level nevertheless intersects with multilingualism on the national level and with pre-existing minority language forces. This is reflected at policy level: Whilst the bilingual Matura framework curricula (SMK, 2012; Bildungs- und Kulturedirektion Kanton Bern, 2022) attach importance to students attaining high levels of proficiency in academic English, it does not formally stipulate language separation. Nonetheless, an individual immersion programme may implicitly replicate and consolidate the kind of English dominant norms which are representative of monolingual traditions in language education, and this may be propagated if teachers cannot access or do not take advantage of continuing professional development offerings (Hess & Williams Leppich, manuscript submitted).

5.2 Why the Swiss case is especially difficult to classify

In the Swiss upper-secondary context, the dominant model is best described as late partial immersion. According to this designation, only part of the curriculum is taught through the L2, distinguishing it from total immersion, with ‘late’ referring to the age at which the intervention begins (Wright & Baker, 2025).

As will have become apparent, there are nevertheless differences to be found at the level of policy and operation. At programme level, we can speak of partial immersion; at classroom level, the pedagogy often overlaps strongly with CLIL. That distinction is useful because immersion describes the structure of the programme, whereas CLIL helps us analyse what happens in pedagogical terms within lessons. The question is therefore not deciding for an either/or label, but rather how CLIL-like the pedagogy within an immersion programme actually is. Overall, we can infer from scholarship in the field that the dominant Swiss model is late, partial immersion, which, in pedagogical terms, usually translates into CLIL (e.g. Bieri, 2018).

The model being used in Swiss upper-secondary classrooms can most usefully be characterised as late partial immersion, normally taking a subject-based, CLIL-like form. It is late because learners typically encounter this intensified form of bilingual schooling at secondary or upper-secondary age rather than from the beginning of primary education and it is partial because only selected academic subjects are taught through the immersion language, rather than the entire curriculum.

The empirical evidence is clear. Elmiger et al.'s (2022) national inventory classified bilingual programmes according to the proportion of teacher conducted bilingually or immersively. Among the Gymnasium programmes supplying the necessary information, 89% fell between 15% and 33% of total instruction while only 7% were above 33% and up to 50%. Thus, the typical Swiss Gymnasium programme was neither a minimal 'immersive island' nor Canadian-style total immersion. Instead, it involved a substantial but ultimately minority share of the curriculum being delivered through the additional language.

The new national framework curriculum makes this pattern even more explicit. The 2026 Swiss Matura Commission calls its standard home-school arrangement 'teilweise immersiver Unterricht... (Modell A)' – or partial immersive teaching, Model A (SMK/SBFI, 2025). It requires at least three Matura subjects to be taught through the immersion language, representing at least 18% of the relevant curriculum (whereby at least one must be a humanities or social science and one must be a science subject).

Earlier Zurich research illustrates how this works on the institutional level. Students continued standard English lessons whilst simultaneously studying non-language subjects like History and Maths through English (Stebler & Maag Merki, 2010). The findings of this study, in combination with publications by scholars including Bieri (2018; 2019) help us

towards a tentative conclusion that Swiss upper-secondary immersion is essentially late partial immersion at programme level and content-led CLIL at classroom level.

5.3 Variation within the unstable ‘Swiss’ immersion model

Having established an approximate overall characterisation of ‘immersion’ in the Swiss educational landscape, an important qualification is nevertheless necessary in this connection. The key question which arises is that of variation: How differently is ‘immersion’ conceptualised, planned and delivered, for example cantons, programmes and schools, and ultimately, what exactly does this unstable ‘Swiss immersion’ model entail?

One aspect for consideration is teaching provision. Acceptance criteria for immersion teaching vary from canton to canton and from school to school (Elmiger et al., 2022), which necessarily results in no single ‘immersion teacher profile’ being identifiable in the Swiss upper-secondary educational landscape. A further aspect around possible variation concerns conceptualisation vs. implementation, i.e. what a school’s policy documents proclaim in comparison – and possibly in contrast with – what the teaching team actually do.

This real variation in recruitment and complementary (resulting?) variation in classroom implementation of course supplies promising discussion points for informal or formal programme evaluation and for further development, whether on the individual professional level or on programme level. Fortunately, teachers are not only encouraged to exchange around these variations and the related discussion points and to use available resources (Hess & Williams Leppich, manuscript submitted); they are above all formally mandated to do so (EDK, 2024; SMK, 2012).

5.4 Why immersion debates often become polarised

As we have seen, the literature in the field of bilingual education frequently presents opposing and even contradictory positions, juxtaposing traditional Canadian immersion with CLIL, for example, before contrasting the two directions with translanguaging paradigms and multilingual education in general. Furthermore, these are often incorrectly presented as mutually exclusive. Examples of recurring and somewhat persistent debates which emerged

from the in-depth literature review and which were introduced above as tensions include discussions around:

- language separation (strict target-language policy vs flexible multilingual practices);
- content vs language priorities (subject content, language or both equally, if sometimes for different lengths of time and at different moments);
- cognitive vs. sociocultural perspectives

It is worth considering, at this juncture, that many of these misalignments may stem from scholars working from different theoretical assumptions rather than fundamentally incompatible evidence, but it is clear that none of this makes it any easier to work towards pre-evaluation consensus in a specific context.

5.4.1 Why competing models need contextual interpretation

As established above, context represents a non-negotiable element in immersion education (Cammarata & Tedick, 2012). ‘Immersion’ in no way represents one uniform educational initiative and consequently models and practices cannot automatically be transferred across contexts with no adjustments. This is because immersion programmes may differ according to:

- language ecology;
- minority/majority relations;
- political agenda;
- societal multilingualism;
- curriculum;
- teacher preparation.

The Swiss examples discussed in the literature review(s) illustrate this need for context-sensitivity. For example, whilst Canadian immersion assumes an English-language majority, a French-speaking minority and ‘additive’ bilingualism, Swiss immersion can variously involve German/French, Italian/German, Romansh, English (with its status as a non-official lingua franca) (Brohy, 2004; Tunger et al., 2023) and this all within highly decentralised cantonal educational systems (Elmiger et al., 2022). We can only conclude that immersion is always a response to a specific sociolinguistic ecology and as such cannot truly exist as one clearly defined pedagogical model.

5.4.2 Conceptual fragmentation

It is crucial to establish the context-sensitivity of immersion practices at upper-secondary level in Switzerland in order to reach some degree of common understanding. In this regard, the central issue is that the constant label ‘immersion’ is in use to denote different actual classroom practices (stemming in turn from differing teacher beliefs).

By the same token and at the same time, different labels are currently in use in the literature on Swiss immersion including:

- ‘immersion’ (referring to the Canadian model, dual language, partial immersion, English Medium Instruction or EMI);
- ‘bilingual education’ (denoting dual language programmes, transitional programmes, heritage language learning);
- CLIL (European content-language integration, which is sometimes virtually identical – or at least thought to be by practitioners – to immersion);
- ‘multilingual education’ (as an umbrella term for all kinds of language education practices and offerings).

This understandably results in:

- confusion amongst practitioners (and potentially also leadership);
- contradictory policy recommendations and likely also uneven implementation;
- studies in research only being weakly comparable (or only with difficulty);
- difficult evidence synthesis.

5.4.3 A lack of evidence on emerging and fragmented conceptualisations

In the field of immersion education in Switzerland there is extensive empirical evidence on outcomes, but at the same time comparatively little work has been done on how conceptual traditions have emerged, evolved, overlapped and fragmented. The state-of-the-art conceptual synthesis presented here attempts not to redress the balance or close any gaps, but at minimum to establish the status quo and possible avenues of development. This specifically means highlighting inconsistencies in use of terminology as well as sharing strategies for seeking clarity here, reconciling competing models and providing a framework for a way forward which is especially relevant for multilingual contexts like Switzerland.

With its official multilingualism, as well as cross-border influences, various different – cantonally influenced and regulated – educational tracks, strong linguistic identities, plus increasing linguistic diversity through migration, Switzerland provides a thought-provoking lens and, to a degree, a promising conceptual testing ground to examine the construct of immersion education. What this also means is that Swiss immersion policy cannot – and perhaps should not – rely on ‘imported’ models (Wright & Baker, 2025). Instead, the situation conclusively requires context-sensitive interpretation (Berthele, 2019; Brohy, 2004; Elmiger et al., 2022). Nevertheless, Swiss immersion can certainly benefit from maintaining an openness to existing models. At the same time, it is important to recognise the issues and complications that come to bear when external models are translated to the distinctly Swiss immersion landscape; see below for a fuller discussion of selected conflicts in this regard.

5.5 Imposing contemporary models on the Swiss educational landscape

5.5.1 Immersion and CLIL: Overlapping traditions which cannot be directly equated

The terminological difficulty around ‘immersion’ as a construct comes into sharper relief when Swiss practice is placed beside international scholarship. The classical concept of immersion education comes especially from Canadian research. Scholars including Genesee (1987) examined programmes which foresaw significant portions of the curriculum being taught through the learners’ second language, with the intention of promoting what is known as additive bilingualism without sacrificing achievement in academic subjects. In the canonical Canadian model, therefore, ‘immersion’ often implies substantial and sustained exposure to the additional language, sometimes beginning very early and sometimes encompassing most of the curriculum.

European Content and Language Integrated Learning (CLIL) developed in a somewhat different landscape. Coyle et al. (2002) define CLIL through its ‘dual focus’: an additional language is used in order to learn both disciplinary content and language. Later scholarship similarly characterises CLIL as an approach where content learning and additional language development are deliberately interconnected (e.g. Gajo, 2007). Prototypical European CLIL can be especially associated with settings where the target language is mainly encountered through schooling instead of functioning as the surrounding societal language (Hélot & Cavalli, 2016).

Nevertheless, this distinction is unstable. Cenoz et al. (2014) warn that CLIL has been defined so broadly that its boundaries with immersion and other forms of content-based instruction are often unclear (Camarata & Tedick, 2012). Programmes necessarily differ in their stated language goals, intensity, student population and the extent to which teachers explicitly teach linguistic forms.

This warning is perhaps especially relevant to the Swiss context. A Swiss school may call a programme ‘immersion’ whilst an applied linguist examining the same classrooms could reasonably describe much of its pedagogy as CLIL. The terms, then, operate at different analytical levels: immersion often describes the programme arrangements (subjects being delivered in another language) whilst CLIL is frequently used to describe the pedagogical integration of subject and language learning within those lessons. CLIL (or CBI – Content-Based instruction – in the U.S.) is a way of approaching mainstream bilingual education in a majority language and came about in the mid-1990s, inspired by Canadian immersion and U.S. dual-language models (Coyle et al., 2010). The approach sees a tenth to as much as a half of the curriculum delivered in another language, with some additional attention paid to the language itself. It can be conceived and enacted on a spectrum from weak to strong and the respective focus on language and on content can be privileged at different times in a teaching period (see e.g. Lyster, 2007, on the counter-balanced approach).

5.5.2 Partial immersion and CLIL: A practical distinction

In Swiss upper-secondary education, partial immersion usually describes the organisational model: only part of the curriculum is taught through an additional language. For example, a student might study Maths, Biology, Geography and History in English whilst continuing with subjects including Chemistry and Physics in the standard school language of instruction. The programme is therefore ‘partial’ because the immersion language is employed only in a certain proportion of total instruction time.

By contrast, CLIL is better understood as a pedagogical orientation. Coyle et al. (2010) emphasise that CLIL involves simultaneous attention to content learning and language development, meaning that whilst an English-medium History lesson can be organised as partial immersion at the institutional level, it can also be taught by means of CLIL principles in the classroom itself. In weak immersion, the teacher might simply deliver the same History lesson in English rather than the L1. A more explicitly CLIL-oriented lesson, by contrast,

would retain the same content objectives whilst also positively identifying the language students need in order to put forward historical explanations and encouraging its use, e.g. expressions of causality, comparison and qualification, such as contributed to, accelerated, in contrast or to a lesser extent (Lasagabaster, 2013). Language thus becomes a tool for disciplinary thinking rather than a separate grammatical syllabus.

A CLIL classroom can generate distinctive opportunities for academic language development through subject-specific interaction (Bieri, 2019, Dalton-Puffer, 2007). In the same vein, Bieri's (2018) work shows how disciplinary knowledge, technical vocabulary and multilingual resources interact in actual upper-secondary classrooms. The conceptual distinction is, however, not immersion versus CLIL. Instead, Swiss schools frequently operate partial-immersion programmes whose individual lessons may be more or less strongly CLIL-oriented. This formulation also goes towards explaining why two separate schools each operating programmes designated 'immersion' can feel very different in pedagogical terms.

5.5.3 Characterising a strong immersion/CLIL lesson

It will have become clear that a strong (contemporary) upper-secondary immersion lesson is not simply an ordinary subject lesson translated into another language. Its defining feature is in fact its deliberate coordination of disciplinary content, cognitive demand and linguistic support. Coyle et al.'s (2010) influential 4Cs framework captures this relationship as Content, Communication, Cognition and Culture, with the so-called 'Language Triptych' (Coyle & Meyer, 2021) further distinguishing amongst the *language of learning* (terminology required for subject knowledge), *language for learning* (words and sentence fragments needed in order to participate in classroom learning) and *language through learning* (which refers to language which emerges in classroom activity and interactions). Learners are tasked with acquiring subject matter, using language to construct that knowledge, demonstrating suitable forms of disciplinary thinking and participating in the cultural practices through which knowledge is communicated.

This is where scaffolding becomes central (Bieri, 2019; Lin, 2015; Ting, 2024). Scaffolding does not (necessarily) mean (over-) simplifying the underlying subject matter. Instead, linguistic and multimodal supports – such as diagrams, sentence prompts, technical vocabulary, reformulation, strategic repetition and occasional movement between languages –

make demanding disciplinary thinking accessible in another language. Dalton-Puffer's (2011) work on CLIL consequently places considerable emphasis on classroom discourse as well as on the opportunities learners receive to use language for specific purposes. In conclusion, we can infer that good 'immersion' maintains the appropriate difficulty level whilst at the same time scaffolding the language necessary to address that difficulty. This is essentially what distinguishes sophisticated CLIL pedagogy from simply using English as a vehicle of learning.

5.6 A four-fold categorisation of immersion scholarship in the Swiss context

The results of an extensive initial and follow-up literature review can be presented according to an essentially four-fold categorisation

A valuable existing evidence base for further programme exploration

Tunger et al. (2023) report on a hugely volume of substantial Swiss literature, with over half of these addressing the Swiss context specifically, as well as multiple educational levels, languages and models. One caveat raised here which is generally applicable is the difficulty in establishing comparability. It must still be noted that a proliferation of literature is not to be confused with conceptual coherence (and in fact, an abundance of publications may even complicate the landscape).

The Zurich bilingual Matura evaluation (Stebler & Maag Merki, 2010) as an important precedent

Stebler and Maag Merki's (2010) study, which was longitudinal and quasi-experimental, and which featured student surveys, teacher surveys, video analysis, as well as a focus on English attainment, History performance and former student follow-up can be considered a significant methodological exemplar. It could be suggested that their study prompts the idea that evaluation can be sophisticated even when still depending on a particular individual programme and the study itself involved multiple dimensions – language, subject learning, motivational self-regulation, classroom processes and transition beyond Gymnasium.

Broader Swiss evidence around heterogeneity

The inventory of bilingual education conducted in 2021/22 by Elmiger et al., 2022 reveals 185 Gymnasium programmes alongside variation in immersion languages, variation in

exchange and residence models, variation in duration/intensity, variation between Gymnasium-vocational and other upper-secondary models. English is overall especially prominent in some contexts and the national languages are especially significant in the bilingual cantons (e.g. Fribourg). Lastly, different institutions have different rationales. All of this substantial evidence points corroborates the overall claim of this dissertation that there no single instance of ‘immersion’ (and that this is insignificant, providing that stakeholders do not act as if there were). Rather than a universal model, we must, as Cathomas (2011) also suggests, therefore be concerned with establishing and interacting with some kind of suitable typology.

Broader international literature (for triangulation), also in related areas

There is clearly a need not only for access to broader international literature, but perhaps above all to an awareness-raising in this connection, as well as relevant literature being embedded in appropriate teacher education offerings. For example, in the area of policy implementation, there would be scope for greater – explicit – teacher exchange on what immersion methodologies, routines and practices exactly they are performing. Teachers do not, of course, simply implement policy; they interpret it through their existing beliefs, knowledge and routines (Borg 2006; Dalton-Puffer, 2011; Williams Leppich 2015). Variation between teachers is therefore not necessarily evidence of failure – it is evidence about how policy is being understood. We need to know not only whether teachers follow policy but also what they deem the policy requires of them. That distinction between intended, interpreted and enacted policy is central to the framework at the heart of this paper.

Chapter 6 Implications and recommendations

6.1 Introduction

With the revised Gymnasium national framework curriculum in place since 2024 (EDK, 2024) and the new bilingual/multilingual Matura guidance newly in effect since earlier this year (SMK/SBFI, 2025), the landscape is clearly evolving and any findings in Swiss upper-secondary immersion education need to be interpreted in the context of a dynamic research environment. Any future studies need to specify which framework they are situated within

and comparisons across data will need to be undertaken with caution and carefully contextualised.

For the moment, it will be instructive to offer some suggestions for the further development of a context-specific immersion offering. The suggestions are the result of the extensive literature review(s) which form the centrepiece of this dissertation, as well as over two decades' work as form teacher of students enrolled on a programme in a school which pioneered immersion, and in addition, nearly five years offering an immersion education preparation module at a teacher training institution.

6.2 School-level

Establish a shared professional vocabulary.

This could be done within the context of an interdisciplinary workshop for teachers. Examples of classroom practice could be shared so as to jumpstart thinking around immersion and provide a forum to allow teachers' implicit definitions and assumptions to emerge. Question prompts might include the following:

- What makes this particular lesson immersive?
- When are we teaching language and when are we teaching through the language?
- What language support can be deemed compatible with immersion?
- What do we expect students to be able to do in the immersion language? etc.

Draw up a mission statement

Establishing a mission statement for an educational institution can give it vision and direction; dedicating a part of that statement – or creating an entirely new one – on a school's immersion programme can potentially be a valuable exercise in distilling the essence of the offering. Existing statements can be used as models or templates and verbs carefully chosen (e.g. according to Bloom's taxonomy). Work on this kind of document could usefully precede a formal external evaluation, with both process and product supplying valuable material, in the sense of an exercise in formative school development.

Build evaluation literacy

In order for teachers and school leaders to be maximally involved in any evaluation processes, the issues under investigation should ideally come from them. Using the

evaluation questions (see Appendix) to prompt discussion will hopefully result in the identification of possible areas of focus, after which they might collaboratively agree on a(n) (action) research question, construct, data source and range of interpretation along with appropriate success indicators. In the interests of delivering useful data for national-level evaluations (see Chapter 7), the team will ideally follow this course of action, thereby generating valuable data, rather than retrofitting their area of context-specific evaluation to existing data.

6.3 Research-level

A move towards common constructs?

Researchers could aim to align somewhat to foster comparability and completion/completeness, that is to move towards common constructs without ‘enforcing’ a particular model. Ideas for shared variables include:

intensity of immersion instruction, distribution of subjects, entry criteria, teacher linguistic competence, immersion methodology preparation, pedagogical scaffolding, student language use, language attainment, subject attainment, student experience, long-term progression

A possible research infrastructure might foresee the following:

- a Swiss immersion programme typology
- a common minimum dataset
- shared reporting mechanisms

Mixed-methods and longitudinal research

In this educational context, it will be worth combining quantitative data (such as language test results, content subject achievement, participation statistics and reports on progress) with qualitative data (e.g. from lesson observations, interviews, student or teacher focus groups, teacher action research, document analysis, and other longitudinal data, meaning there could be comparisons in before/after, programme trajectories, implementation fidelity and different models).

More targeted attention needs overall to be paid to implementation. In place of variations on the over-simplified and at the same time imprecise ‘Does immersion work?’, potential

research could consider a context-specific take on immersion success, e.g. ‘Under what organisational and pedagogical conditions does a certain form of immersion produce particular outcomes for a defined set of students?’ This would move research endeavours towards a more carefully contextualised or ‘realist’ evaluation.

6.4 Policy-level

Preserve cantonal diversity

Data collection might usefully be more streamlined and standardised so as to improve comparability and generalisability, but this should not result in pressure to adopt, retain or overhaul a particular model.

Improve national visibility of provision

The literature reveals that there is an absence of national school-level data on immersion classes (e.g. SMK, 2012). It would make sense to gather consistent programme-level data, as well as details regarding immersion languages, proportion of the curriculum occupied by the respective languages, the subjects studied, student numbers, student entry criteria, teacher qualifications, as well as any exchange arrangements.

Optionally: Link accreditation to developmental quality

Current SMK guidance already provides a useful direction here by way of its requirements around C2 competence, immersion methodology training and education for teachers and the importance of teacher exchange (SMK/SBFI, 2025). Policy could here meaningfully support any of the following: professional learning, networks, exchange around practice, shared resources and research-practice partnerships. Overall, however, any accreditation criteria must remain easily distinguishable from pedagogical quality criteria.

6.5 Towards a common goal in immersion education: a proposed research-practice agenda

In summary, the above possibilities could be presented and structured by means of a research-practice agenda, which should of course be understood as an iterative cycle rather than a linear sequence:

- Short term – school level (e.g. finding a common definition, outlining programme theory, teacher professional learning communities);
- Medium term – cantonal level (shared descriptors, school-to-school professional development, comparable monitoring design, implementation and evaluation);
- Longer term – national/research level (longitudinal mixed-method studies, common outcome measures, analysis of equity and social justice policies, overall cross-language-region comparisons).

This structure deliberately mirrors how Swiss policy enactment works in other contexts, typically filtering down from national to school via cantonal level. This would, however, allow for grass-roots involvement.

6.6 Preparing for a future possible (informal) evaluation

The question prompts which feature in this dissertation have been conceived and articulate to trigger reflection and exchange amongst stakeholders in upper-secondary immersion programmes in the Swiss context, especially amongst teachers and leaders. It is very possible that, by engaging with questions like these, the perceived need to evaluate a programme sooner rather than later evaporates.

Nevertheless, should a school wish to prepare for a future evaluation, even informal, small scale, low-stakes, self-designed and implemented, and internal only, it would be worth bearing the following in mind:

Evaluation requires a theory of a programme, that is, programme features should be described explicitly rather than inferred from terminology which – as Cenoz (2015) suggests may lead us to believe that differences are more substantial than they are and therefore to perceive them as obstacles to a common understanding. This notion of the theory of a programme can neatly be expressed as a conceptual chain, taking a school from programme features to pedagogical processes, learner experience, short-term outcomes to longer term outcomes (Braun et al. 2011). For immersion this might concretely be:

age (e.g. Johnstone, 2002) > staffing / timetabling / language allocation > teacher practices > language and content interaction > student participation > subject learning and L2 development > further study / mobility / intercultural competences / identity.

Without consensus as to the causal chain, any evaluation is at risk of measuring only what is easiest to measure. An explicit programme theory (Weiss, 1997) or other logic model can help.

'Immersion' as an umbrella term

A possible typology:

- a) Intensity (e.g. % of teaching in the L2, number of lessons, number of subjects, duration)
- b) Distribution (e.g. continuous vs intermittent, selected subjects vs. broad curriculum, exchange-based vs classroom-based)
- c) Pedagogy (e.g. L2 newly as medium not object, explicit language support, scaffolding, translanguaging, disciplinary literacy, interaction, assessment)
- d) Student population (e.g. selective entry y/n, self-selection, language prerequisites, prior attainment, socio-economic/access patterns)
- e) Outcomes (e.g. language proficiency, disciplinary knowledge, academic literacy, intercultural learning, motivation, confidence and/identity, progression)

The key here is to make these dimensions visible *before* judging programme's success or otherwise (see e.g. Yu, 2024).

A common understanding does not mean forced uniformity

There is an important nuance to be found here, in that a shared conceptual vocabulary does not necessarily result in identical implementation, regardless of context. Schools can choose to retain local models and strive thereby for comparability without homogenisation, transparency without bureaucracy, and evaluability without reductionism. The overall aim, then, is not a single definition of 'Swiss immersion' but a sufficiently common conceptual language to facilitate description, discussion and evaluation of different forms of immersion in comprehensible terms.

Why teachers matter

Teachers are not merely delivery mechanisms. Instead, they interpret curriculum, language policy, disciplinary norms and student needs. Immersion can impact in complex ways upon teacher practice, introducing elements to be negotiated around language/content balancing

(Lyster, 2007; Lyster & Mori, 2006), scaffolding, and adapted questioning, assessment and classroom interaction. Current SMK guidance provides a useful point of reference here, foregrounding C2 competence, immersion methodology development and professional exchange (SMK, 2012).

Why school leaders matter

Leadership decisions shape multiple areas of school life which also impact upon immersion: staffing, timetabling (amount and priority), recruitment, professional learning, collaboration, communication with students and parents, data collection and programme sustainability. For this reason, evaluation should not be conceptualised as student outcomes (Weiss, 1997) but should rather also take into account organisational conditions (Spillane, 2006 – on distributed leadership).

6.7 Concluding words on concrete proposals

Formal evaluation is undeniably valuable so as to work towards ensuring success. However, an evaluation commissioned without conceptual alignment risks answering an ambiguous question precisely, measuring outcomes without knowing intended mechanisms, comparing unlike programmes and favouring easily measurable outcomes.

The immediate task for many Swiss upper-secondary schools may not (anymore) be to decide what ‘immersion’ is in the general conception of the word, but rather to articulate what their own special immersion model intends to accomplish, how it is expected to accomplish it and what evidence would represent meaningful ‘success’. A common understanding is therefore essential and needs to be shared (but revisable), specific (but not prescriptive), locally grounded (but internationally informed) and sufficiently explicit as to support evaluation.

Chapter 7 Conclusion

This dissertation argues that the term ‘immersion’ as a designation for bilingual education programmes is elusive and attempts to label may even ultimately complicate matters. It also suggests that – perhaps precisely *because* of these difficulties in defining the term – the long-standing question of whether ‘immersion’ is in general successful has become less productive than that of how immersion should be described, designed and represented in specific

sociolinguistic contexts. Existing models (including additive bilingualism, CLIL and translanguaging) each illuminate important dimensions of multilingual education but at the same time also reflect distinct theoretical assumptions and policy goals. Their apparent incompatibility often comes about from being applied beyond the contexts they were developed for. A context-sensitive (pre)evaluation tool could usefully shift attention from choosing between competing pedagogies to raising awareness in decision-making which aligns educational practices with local linguistic contexts, institutional constraints and learner needs (see e.g. Tarnanen and Palviainen, 2018).

In this regard, teacher education and professional development are key (Brohy, 2004; Cammarata & Tedick, 2012; Hess & Williams Leppich submitted; Illman & Pietilä, 2018; Tedick, 2020). In order to teach effectively, future immersion teachers need learn to an understanding of the immersion model their school operates according to and clear perceptions on their own classroom work. It is clear that immersion teachers require specialised skills for integrating content teaching with language development (in the general sense of CLIL pedagogy) and that these skills need to be taught, learnt, internalised and reflected upon.

This dissertation therefore proposes a series of evaluation questions which are adaptable to all kinds of stakeholders involved in a Swiss immersion programme setting. The questions are designed to be possible not compulsory and to offer options and variety, thereby prompting and structuring teacher thinking without overwhelming or alarming practitioners (or, by the same token, their leaders or research partners). Tedick and Mathieu (2025) use the rubric in their study as a mediational tool in the same way that Williams Leppich (2015) proposes course materials as discussion starter; the questions here (see Appendix) are also designed in that same tradition, to prompt reflection and articulation, rather than – necessarily – to form the foundation of a professional/formal/institutional level evaluation. The rubric supplies teachers (and potentially also pre-service immersion teachers) with a shared language so they can notice elements of classroom practice, exchange on these and collaboratively set goals for programme (and self-?) development and enhancement.

This matters because effective (immersion) teaching turns on teachers noticing and thinking explicitly about factors including how and where language and content interact (e.g. Borg, 2006; Gajo, 2007; Tedick & Mathieu, 2025) instead of only assuming that delivering content in another language will automatically contribute to linguistic skills development.

Ultimately, the question is not simply what ‘immersion’ in general is successful, but rather what kind of immersion a school actually creates, values and propagates. Policy must be seen as locally mediated rather than purely implemented (Braun et al., 2011) so evaluating an immersion programme therefore moves beyond the general and binary question of success or failure towards an analysis of implementation, conditions and consequences. What matters is the relationship between intended policy, classroom enactment and student experience. A strong programme can be said to be one where policy, teacher capacity, classroom practice and student outcomes are reasonably coherent and carefully aligned. The framework presented here is therefore designed not simply to judge the programme to understand where and why it is working differently across the school.

Success is multi-dimensional and a successful programme needs to be sustainable (Brohy, 2004). In these terms, if a programme depends on a small number of highly committed and talented individual teachers, it may be effective but fragile. The key question is whether immersion has become part of institutional practice rather than remaining the project of a few individual teachers. In order to achieve this sustainability, I would argue, it is necessary for teachers to find some degree of alignment in their context-specific immersion work.

In setting out an argument for finding that shared, specific understanding of ‘immersion’ in a defined Swiss context, this dissertation seeks to make evaluating such offerings easier to effect. With a shared understanding from the outset (having benefited from the PLC learning and discussion on the trajectory towards this understanding), any evaluation can become potentially less demanding (and therefore more appealing e.g. to school leaders). The process can become more focused and purposeful with increased transparency and enhanced local meaning and applicability. Ideally, it will also make an evaluation more comparable, thus fostering school-to-school comparability.

This conceptual clarity offers a highly productive basis for subsequent evaluations. At classroom teacher level, stakeholders will have at their disposal a (new) shared professional language as well as an enhanced ability to challenge both their own or others’ assumptions, whilst at leadership level a clearer basis for planning and strategy should emerge. In the fields of research and policy there will also be gains: Researchers stand to benefit from more precise constructs and improved interpretation of outcomes, as well as enhanced comparability, whilst policy makers will be able to draw on robust evidence distinguishing initial programme design from its resulting implementation (Durlak & DuPre, 2008). It is

nevertheless important to note that in spite of these beneficial developments, conceptual alignment is no guarantee of programme quality. Equally, evaluation – though it might consequently be postponed, reduced in scope or focused only on some areas – remains essential. Simply achieving conceptual agreement among context-bound stakeholders will not suffice. Nonetheless, conceptual clarification is a crucial first step, though not to be considered only administrative preparation for an evaluation; instead it is in and of itself a part of educational quality development and enhancement.

As already discussed, simply asking ‘Does immersion work?’ is an unproductive question as it lacks (context-)specificity. Asking instead the questions laid out in Chapter 4 means a school can make explicit their beliefs and wishes around immersion. The Swiss context offers especially fertile ground for fundamental discussion of relevant issues, due to strong national policy tradition in multilingualism, significant programme diversity and a constantly evolving regulatory framework, as well as the substantial (Tunger et al. 2023 identified approximately 1200 items, more than half of which related to Switzerland) research base selectively discussed in earlier chapters.

This dissertation turns on the essential question of what upper-secondary Swiss immersion is and how the way can be paved to evaluate it in specific contexts. The literature review of chapter 3 complicated the landscape by establishing several key points. ‘Immersion’ is not a uniform intervention and Swiss provision of immersion offerings varies substantially. Variation can be found at programme level, but significantly also even at the terminology level. Purposes for students enrolling on programmes differ (Elmiger et al., 2022) and despite significant research interest, the evidence base remains fragmented and heterogeneous (SKBF, 2026).

Nevertheless, useful conclusions can still be drawn. Firstly, immersion may be heterogeneous but it is established (and vice versa). We see evidence of a general interest in and expansion of bilingual provision (Tunger et al., 2023), with the current evidence proving that immersion is concentrated at grammar schools and there is notable variation in programme models (Elmiger et al., 2022) as well as provision in multiple languages. Taking these factors into account, it becomes impossible to genuinely speak of ‘the Swiss immersion model’. Instead, in order to bring into relief these various dimensions, we might be better placed to refer to a Swiss *landscape of immersion provision*, this designation nicely capturing the geographical and political elements in the discourse as well as signifying an ever-changing ecology of

combined content and language education. (As we have seen, policy around this is permanently in flux, with the current co-existence of the 2012 regulations and the new 2026 multilingual Matura guidance during the transition to the revised MAR/MAV national framework curriculum.)

Secondly, it has also been demonstrated that there is a raft of scholarship in the field but that this is fragmented in the same manner as the terminology and programme designs are. Existing studies show that immersion outcomes can be interrogated along different lines, including those of language, subject content learning, motivation, classroom practice and future educational paths. Stebler and Maag Merki's (2010) work in particular demonstrates the range of possible research methodologies. Elmiger et al's (2022) inventory shares a substantial bibliography, which provided useful starting points for exploration. Whilst Stebler and Maag Merki's (2010) publication reports on a wide-ranging exploration of one cantonal context, Elmiger et al.'s (2022) work sweeps the overall Swiss immersion landscape, addressing multiple contexts and dealing with issues around methodological and conceptual heterogeneity as well as the resulting difficulty in comparing and/or generalising findings; the difference between the two beautifully illustrates the tension between general and context-specific – and the undeniable need for the two to complement one another. will not suffice.

Thirdly, policy can usefully provide boundaries but it cannot represent a complete theory of practice. That is, the regulatory documentation delivers information around languages, the number of possible subjects studied in the upper-secondary immersion classroom, the minimum number of hours in immersion required for the designation 'immersion' and – to a degree – teacher qualifications (though as we have seen, in practice, this can vary from canton to canton). Policy does not prescribe pedagogical implementation or school-level purpose, nor does it make stipulations around student experience, teaching quality or access issues in the name of equity. In summary, then, regulatory frameworks and policy can define the conditions under which a programme is recognised as 'immersion' but – for reasons of much-needed context-sensitivity – cannot ultimately define on their own what good immersion looks like in every classroom.

References

- Baetens Beardsmore, H. (1993) 'The European School Model', in: Baetens Beardsmore, H. (ed.) *European Models of Bilingual Education*. Clevedon: Multilingual Matters.
- Baker, C. and Wright, W.E. (2017) *Foundations of bilingual education and bilingualism*. 6th edn. Bristol: Multilingual Matters.
- Bakken, J. and Brevik, L.M. (2023) 'Challenging the notion of CLIL elitism: A study of secondary school students' motivation for choosing CLIL in Norway', *TESOL Quarterly*, 57(4), pp. 1091–1114. DOI: 10.1002/tesq.3173.
- Ball, S.J., Maguire, M. and Braun, A. (2012) *How Schools Do Policy: Policy Enactments in Secondary Schools*. Abingdon: Routledge.
- Beacco, J.-C., Byram, M., Cavalli, M., Coste, D., Egli Cuenat, M., Goullier, F. and Panthier, J. (2016) *Guide for the development and implementation of curricula for plurilingual and intercultural education*. Strasbourg: Council of Europe.
- Berthele, R. (2021) 'The selective celebration of linguistic diversity: evidence from the Swiss language policy discourse', *Journal of Multilingual and Multicultural Development*, 42(2), pp. 125–136. DOI: 10.1080/01434632.2020.1715989.
- Bieri, A.S. (2018) 'Translanguaging practices in CLIL and non-CLIL biology lessons in Switzerland', *EuroAmerican Journal of Applied Linguistics and Languages*, 5(2), pp. 91–109. DOI: 10.21283/2376905X.9.142
- Bieri, A.S. (2019) 'Technicality and translanguaging in CLIL biology lessons', *Bulletin Suisse de linguistique appliquée*, 110, pp. 109–124.
- Bildungs- und Kulturdirektion Kanton Bern (2022.) *Zweisprachige Maturität am Gymnasium*. Available at: <https://www.bkd.be.ch/de/start/themen/bildung-im-kanton-bern/mittelschulen/gymnasium/zweisprachige-maturitaet.html> (Accessed: 23 September 2026).
- Borg, S. (2006) *Teacher cognition and language education: Research and practice*. London, England: Continuum.
- Braun, A., Ball, S., Maguire, M. and Hoskins, K. (2011) 'Taking context seriously: Towards explaining policy enactments in the secondary school', *Discourse: Studies in the Cultural Politics of Education*, 32(4), pp. 585–596. DOI: 10.1080/01596306.2011.601555.

Brohy, C. (2004) 'Plurilingual education in Switzerland. From school development to everyday school life', *Swiss Journal of Educational Research*, 26(3), pp. 465–476, DOI: 10.24452/sjer.26.3.4690.

Byrom, T. (2016) 'Constructing the Literature Section', in I. Palaiologou, D. Needham & T. Male (eds.), *Doing Research in Education: Theory and Practice* (pp. 118–135). London: SAGE Publications.

Cammarata, L. and Tedick, D.J. (2012) 'Balancing Content and Language in Instruction: The Experience of Immersion Teachers', *The Modern Language Journal*, 96(2), 251–269. DOI: 10.1111/j.1540-4781.2012.01330.x.

Carroll, C., Patterson, M., Wood, S., Booth, S., Rick, J. and Balain, S. (2007) 'A conceptual framework for implementation fidelity', *Implementation Science*, 2, article 40. DOI: 101186/1748-5908-2-40.

Cathomas, R. (2011) 'Immersion – ein Modebegriff. Eine kritische Begriffs(er)klärung anhand einer Typologie zwei- und mehrsprachiger Schulmodelle', in Todisco, V. and Trezzini, M. (eds). *Mythos Babel: Mehrsprachigkeitsdidaktik zwischen Schein, Sein und Wollen*. Zürich: Verlag Pestalozzianum, pp. 156–168.

Cenoz, J. (2015) 'Content-based instruction and CLIL: the same or different?', *Language, Culture and Curriculum*, 28(1), pp. 8–24. DOI: 10.1080/07908318.2014.1000922.

Cenoz, J. and Gorter, D. (2014) 'Critical Analysis of CLIL: Taking stock and looking forward', *Applied Linguistics*, 35(3), pp. 243–262.

Cenoz, J. and Gorter, D. (2011) 'A Holistic Approach to Multilingual Education: Introduction', *The Modern Language Journal*, 95(3), pp.339–343. DOI: 10.1111/j.1540-4781.2011.01204.x.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Coyle, D., Hood, P. and Marsh, D. (2010) *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.

Coyle, D. and Meyer, O. (2021) *Beyond CLIL: Pluriliteracies Teaching for Deeper Learning*. Cambridge: Cambridge University Press. DOI: 10.1017.9781108914505.

Cummins, J. (1979) 'Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters', *Working Papers on Bilingualism*, 19, pp.121–129.

Cummins, J. (2007) 'Rethinking monolingual instructional strategies in multilingual classrooms', *Canadian Journal of Applied Linguistics*, 10(2), pp. 221–240.

Cummins, J. (2017) 'Teaching for transfer in multilingual school contexts', in O. García,

A.M.Y. Lin and C.S. May (eds.), *Bilingual and Multilingual Education*. Cham: Springer, pp. 103–115. DOI: 10.1007/978-3-319-02258-1_8

Dalton-Puffer, C. (2007) Discourse in Content and Language Integrated Learning (CLIL) classrooms. Amsterdam: John Benjamins. DOI: 10.1075/llt.20.

Dalton-Puffer, C. (2011) ‘Content and language integrated learning: from practice to principles’, *Annual Review of Applied Linguistics*, 31, pp. 182–204

De Mejía, A.M. (2002) Power, prestige and bilingualism: International perspectives on elite bilingual education. Clevedon: Multilingual Matters.

Durlak, J.A. and DuPre, E.P. (2008) ‘Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation’, *American Journal of Community Psychology*, 41(3), pp. 328–350. DOI: 10.1007/s10464-008-9165-0.

EDK (2024) *Rahmenlehrplan für die gymnasialen Maturitätsschulen*. Bern: Schweizerische Konferenz der kantonalen Erziehungsdirektorinnen und -direktoren (EDK)

Elmiger, D., Siegenthaler, A., and Tunger, V. (2022) *Inventar des zweisprachigen Unterrichts in der Schweiz: Gesamtschau 2021/2022*. Freiburg: Institut für Mehrsprachigkeit.

Fielding, R. (2022) ‘Future Directions for Multilingualism, Identity and Interculturality in Education’, in Fielding, R. (ed.) *Multilingualism, Identity and Interculturality in Education*. Singapore: Springer, pp. 199–210. DOI: 10.1007/978-981-19-5848-9_10.

García, O. (2009) *Bilingual education in the 21st century: A global perspective*. Oxford: Wiley-Blackwell.

García, O., and Wei, L. (2014) Translanguaging: Language, bilingualism and education. Palgrave Macmillan. DOI: 10.1057/9781137385765.

García, O., Johnson, S.I. and Seltzer, K. (2017) *The Translanguaging Classroom: Leveraging Student Bilingualism for Learning*. Philadelphia, PA: Caslon.

Gajo, L. (2007) ‘Linguistic knowledge and subject knowledge: How does bilingualism contribute to subject development?’, *International Journal of Bilingual Education and Bilingualism*, 20(5), pp. 563–581. DOI: 10.2167.beb460.0.

Gajo, L., Grobet, A. and Steffen, G. (2015) ‘Saisir le niveau méso-interactionnel dans l’interaction didactique : autour des notions de saturation et de méso-alternance?’, *Bulletin suisse de linguistique appliquée*, 101, pp. 27–49.

Genesee, F., Lindholm-Leary, K., Saunders, W. and Christian, D. (2006) *Educating English language learners: A synthesis of research evidence*. Cambridge: Cambridge University Press.

Grames, E.M., Schwartz, D. and Elphick, C.S. (2022) ‘A systematic method for hypothesis synthesis and conceptual model development’, *Methods in Ecology and Evolution*, pp. 2078–

2087. DOI: 10.1111/2041-210X.13940.

Hart, C. (2018) *Doing a Literature Review: Releasing the Research Imagination*. 2nd edn. London: SAGE Publications.

Hega, G. M. (2001) 'Regional Identity, Language and Education Policy in Switzerland', *Compare: A Journal of Comparative and International Education*, 31(2), pp. 205–227. DOI: 10.1080/03057920120053210.

Hélot, C. & Cavalli, M. (2016) 'Bilingual education in Europe: Dominant languages', in García, O., Lin, A. and May, S. (eds.) *Bilingual and Multilingual Education*. Cham: Springer, pp. 471–488.

Hess, D.B. and Williams Leppich, L. (2026) 'Multilingualism as an asset in immersion teaching: Maximising its integration in teacher education'. *Manuscript submitted for publication*.

Hornberger, N.H. (2003) *Continua of biliteracy: An ecological framework for educational policy, research, and practice in multilingual settings*. Bristol: Multilingual Matters.

Illman, V. & Pietilä, P. (2018) Multilingualism as a resource in the foreign language classroom. *ELT Journal*, 72(3), pp. 237–248. DOI: 10.1093/elt/ccx073.

Jerrim, J. & de Vries, R. (2017) 'The limitations of quantitative social science for informing public policy' *Evidence and Policy*, 13(1), pp. 117–133. DOI: 10.1332/174426415X14431000856662.

Johnstone, R. (2002) 'Addressing the 'age factor': Some implications for languages policy'. Strasbourg: Council of Europe.

Kuźelewska, E. (2016) 'Language policy in Switzerland', *Studies in Logic, Grammar and Rhetoric*, 45(1), pp. 125–140. DOI: 10.1515/slgr-2016-0020.

Lasagabaster, D. (2013) 'The use of the L1 in CLIL classes: The teachers' perspective', *Latin American Journal of Content & Language Integrated Learning*, 6(2), pp. 1–21. DOI: [10.5294/3148](https://doi.org/10.5294/3148)

Lasagabaster, D. and Sierra, J.M. (2010) 'Immersion and CLIL in English: More differences than similarities', *ELT Journal*, 64(4), pp. 367–375. DOI: 10.1093/elt/ccp082.

Lin, A.M.Y. (2015) Conceptualising the potential role of L1 in CLIL. *Language, Culture and Curriculum*, 28(1), pp. 74–89. DOI: [10.1080/07908318.2014.1000926](https://doi.org/10.1080/07908318.2014.1000926)

Lindholm-Leary, K.J. (2012) Success and challenges in dual language education. *Theory into Practice*, 51(4), pp. 256–262. DOI: 10.1080/00405841.2012.726053.

Llinares, A., Morton, T., and Whittaker, R. (2021) *The roles of language in CLIL*. Cambridge: Cambridge University Press.

Lorenzo, F., Granados, A. and Rico, N. (2021) Equity in bilingual education: Socioeconomic

status and content and language integrated learning in monolingual Southern Europe. *Applied Linguistics*, 42(3), pp. 393–413. DOI: 10.1093/applin/amaa037.

Lyster, R. (2007) *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: John Benjamins.

Lyster, R. and Mori, J. (2006) 'Interactional feedback and instructional counterbalance', *Studies in Second Language Acquisition*, 28(2), pp. 269–300. DOI: 10.1017/S02722263106060128.

Marsh, D. (2002) *CLIL/EMILE – The European dimension: Actions, trends and foresight potential*. Brussels: European Commission.

May, S. (ed.) (2014) *The multilingual turn: Implications for SLA, TESOL and bilingual education*. Abingdon: Routledge.

McCulloch, G. (2017) 'Documentary methods', in Coe, R. Waring, M. Hedges, L.V. and Arthur, J. (eds.) *Research Methods and Methodologies in Education*. 2nd ed. London: SAGE Publications, pp. 215–222

Nikula, T., Dafouz, E., Moore, P. and Smit, U. (eds) (2016) *Conceptualising integration in CLIL and Multilingual Education*. Bristol: Multilingual Matters.

OECD (2024) *How 15-Year Olds Learn English: Case studies from Finland, Greece, Israel, the Netherlands and Portugal*. Paris: OECD Publishing. DOI: 10.1787/a3fcacd5-en.

Patton, M. Q. (2008) *Utilization focused evaluation*. 4th edn. Thousand Oaks, CA: SAGE Publications.

Pitkänen-Huhta, A. and Mäntylä, K. (2021) Teachers negotiating multilingualism in the EFL classroom. *European Journal of Applied Linguistics*, 9(2), pp. 283–306. DOI: 10.1515/eujal-2018-0020.

Prusse, M. and Nakamura, Y. (2010) 'Englischkompetenzen in Immersions- und Kontrollklassen', in Stebler, R. and Maag Merki, K. (eds) *Zweisprachig lernen: Prozesse und Wirkungen eines immersiven Ausbildungsganges an Gymnasien*. Münster: Waxmann. pp. 131–151.

Ricento, T. (2000) 'Historical and theoretical perspectives in language policy and planning', *Journal of Sociolinguistics*, 4(2), pp. 196–213. DOI: 10.1111/1467-9481.00111.

Ridley, D. (2012) *The Literature Review: A Step-by-Step Guide for Students*. 2nd edn. London: SAGE Publications.

Saldaña, P. (2016) *The Coding Manual for Qualitative Researchers*. London: SAGE Publications.

Schweizerische Maturitätskommission SMK. (2012) Reglement der Schweizerischen Maturitätskommission (SMK) für die Anerkennung kantonaler zweisprachiger Maturitäten. Reglement_SMK_Anerkennung_zweisprachige_Maturitaet_16.03.2012_dt.pdf

SMK/SBFI [Schweizerische Maturitätskommission](2025) *Richtlinie für die Durchführung von*

mehrsprachigen Maturitätslehrgängen. Bern. Staatssekretariat für Bildung, Forschung und Innovation (SBFI).

Spillane, J.P., Reiser, B.J. and Reimer, T. (2002) 'Policy Implementation and Cognition: Reframing and Refocusing Implementation Research', *Review of Educational Research*, 72(3), pp. 387–431.

Stebler, R. (2010) 'Kontext und Forschungsstand', in Stebler, R. and Maag Merki, K. (eds) *Zweisprachig lernen: Prozesse und Wirkungen eines immersiven Ausbildungsganges an Gymnasien*. Münster: Waxmann, pp. 13–23.

Stufflebeam, D.L. and Zhang, G. (2017) *The CIPP evaluation model: How to evaluate for improvement and accountability*. New York: Guilford Press.

Tarnanen, M. and Palviainen, Å. (2018) 'Finnish teachers as policy agents in a changing society', *Language and Education*, 32(5), pp. 428–443. DOI: 10.1080/09500782.2018.1490747

Tedick, D.J. (2020) 'Foreword', in: Bower, K., Coyle, D., Cross, R. and Chambers, G.N. (eds) *Curriculum integrated language teaching: CLIL in practice*. Cambridge University Press, pp. xi–xvi. DOI: 10.1017/9781108687867.001

Tedick, D.J. and Mathieu, C. (2025) 'Using a rubric as a mediational tool to assess pre-service immersion teacher development', *Journal of Immersion and Content-Based Language Education*. DOI: 10.1075/jicb-24015.ted.

Thomas, W.P. and Collier, V.P. (2002) *A national study of school effectiveness for language minority students' long-term academic achievement*. Center for Research on Education, Diversity & Excellence.

Ting, Y.L.T. (2024) 'Designing materials for building subject-specific literacies: Secondary-level chemistry', in Hüttner, J. and Dalton-Puffer, C. (eds) *Building disciplinary literacies in content and language integrated learning*. Abingdon: Routledge, pp. 164–194.

Tunger, V., Elmiger, D. and Siegenthaler, A. (2023) *Begleitband: Zweisprachiger Unterricht in der Schweiz. Projekte, Akteurinnen und Akteure, Kontexte, Chronologie*. Freiburg: Institut für Mehrsprachigkeit.

Weiss, C.H. (1997) 'Theory-based evaluation: Past, present, and future', *New Directions for Evaluation*, 1997(76), pp. 41–55. DOI: 10.1002/ev.1086

Weiss, C. (2005) 'Evaluation decisions: is anybody there? Does anybody care?' In *Evaluation Research Methods Volume iv* (ed.) Stern, E., pp. 286–299

Williams Leppich, L. (2015) 'Reaching the teachers through the coursebook', in Weil, M. and Vanotti, M. (eds) *Weiterbildung und Mehrsprachigkeit / Formation continue et plurilinguisme / Further education and plurilingualism*. Bern: hep-Verlag, pp. 161–175.

Williams Leppich, L. (2026) 'Ich Find's Mega Nice von Ihm! Glocal Voices: How Adolescents in Switzerland are Reimagining English', in N. Wright & D. Xerri (eds) *The*

Rise of English as a Glocal Language: Reimagining its Role in an Era of Deglobalisation. Cham: Springer, pp. 227–248.

Wright, W.E. and Baker, C. (2025) *Foundations of bilingual education and bilingualism*. 8th edn. Bristol: Multilingual Matters.

Yazdani, S. and Bayazidi, S. (2020) Undertaking a concept synthesis: a stepwise approach. *International Journal of Quantitative and Qualitative Research Methods*, 8(2), pp. 1–13. Available at: <https://ejournals.org/ijqqr/vol-8-issue-2-may-2020/undertaking-a-concept-synthesis-a-stepwise-approach/> [Accessed 29 September 2026]

Yu, K. (2024) ‘Bilingual program effectiveness: An evaluation of meta-analytic methods and findings’, *International Journal of Bilingual Education and Bilingualism*, 27(6), pp.744–771. DOI: 10.1080/13670050.2023.2232522

Evaluation questions

Key guiding question: Does the theoretical framework of the programme make it possible to evaluate it? If not, what clarification work needs to be done in the teacher-leader team in preparation?

Overall topics for discussion

1. What is the model trying to achieve?
2. Is the model appropriate to its Swiss *Gymnasium* context?
3. Are its language practices in alignment with its linguistic and educational objectives?
4. Does it enable rather than impede subject learning?
5. Do teachers have the conditions and competences required to implement it?
6. Is the enacted programme consistent enough with the intended programme to remain coherent?
7. Who does the programme work for, under what conditions and for what outcomes?
8. Which features should policy standardise and which should remain context-dependent?

1 Programme aims and coherence

- i) What is the Swiss *Gymnasium* immersion model intended to achieve: subject learning, additional language development, preparation for tertiary education, internationalisation or a combination of these?
- ii) Are the aims in i) explicit and mutually compatible?
- iii) Are programme practices consistent with the aims in i)?
- iv) Does the model have a sufficiently clear identity to distinguish it from standard foreign language teaching or CLIL?
- v) Which features are essential to programme coherence and which can legitimately vary by school, canton, subject or teacher?

2 Contextual fit

- i) To what extent is the model appropriate to the specific sociolinguistic and institutional context of the Swiss *Gymnasium*?
- ii) Which features of the model arise from specifically Swiss conditions rather than from general principles of immersion education?
- iii) How do cantonal autonomy, the local language environment and differing school contexts affect what immersion might reasonably look like?
- iv) Can the same evaluative criteria be applied across different Swiss programmes or do they need to be context-sensitive?
- v) To what extent can findings or prescriptions from immersion programmes elsewhere be legitimately transferred to the Swiss *Gymnasium* context?

3 Language learning and target language use

- i) What level and type of additional language development should the programme be expected to produce?
- ii) Is sustained use of the immersion language necessary to achieve those aims?
- iii) How much target language use is necessary for the programme to remain meaningfully ‘immersive’?
- iv) Under what circumstances does students’ use of other languages support rather than threaten or weaken immersion?
- v) Should target language use be evaluated as an end in itself, as a means to learning or both?
- vi) Is there evidence that expectations connected with target language use are appropriate to students’ developmental level?

4 Subject learning

- i) Does learning a subject in an additional language affect students’ access to subject content?
- ii) Are subject learning objectives maintained at the same level as in non-immersion classes?
- iii) Are some subjects better suited to immersion than others?
- iv) What kinds of linguistic support are necessary to prevent language difficulty from becoming subject difficulty?
- v) How should an evaluation distinguish between problems with subject knowledge and problems expressing subject knowledge in the immersion language?

5 Multilingualism and translanguaging

- i) What pedagogical functions can students’ full linguistic repertoires serve in *Gymnasium* immersion classrooms?
- ii) When is translanguaging pedagogically useful?
- iii) Are there contexts where translanguaging conflicts with any linguistic aims of the programme?
- iv) How should an evaluator distinguish strategic multilingual practice from simple abandonment of the immersion language?
- v) Does the programme allow teachers enough flexibility to respond to learner needs whilst still maintaining programme coherence?

6 Teachers and implementation

- i) What linguistic, subject-specific and pedagogical competence is required to teach effectively in a bilingual *Gymnasium* programme?
- ii) Are existing qualification requirements aligned with what bilingual teaching actually demands?

- iii) How adequately are teachers prepared for the additional pedagogical demands of teaching subject content in another language?
- iv) To what extent do teachers understand the aims of the programme in similar ways?
- v) How much variation exists between teachers in the way the model is implemented?
- vi) When teachers depart from official or recommended practice, why do they do that?
- vii) To what extent do teachers' stated beliefs correspond to their classroom practice?
- viii) What institutional constraints prevent teachers from implementing practices they deem sensible?

7 Programme legitimacy and the bilingual Matura

- i) What gives the bilingual programme legitimacy in the eyes of students, teachers, schools and external authorities?
- ii) How important is sustained use of the immersion language in that regard?
- iii) Do formal expectations concerning linguistic performance encourage pedagogically productive practices or can they also constrain useful multilingual practices?
- iv) Is there a tension between meeting institutional expectations for a bilingual Matura and responding flexibly to students' learning needs?
- v) Would greater linguistic flexibility threaten the distinctiveness of the programme or can programme coherence be guaranteed in other ways?
- vi) Are the criteria used in the accreditation or other certification of a bilingual programme aligned with what research suggests makes these programmes educationally effective?

8 Student access, selection and equity

- i) Who participates in bilingual *Gymnasium* programmes and who does not?
- ii) What linguistic or academic prerequisites are expected?
- iii) Do admission arrangements favour students who already have stronger language skills?
- iv) To what extent can positive programme outcomes be distinguished from the effects of selection?
- v) Are students with different linguistic backgrounds equally able to benefit?
- vi) Does the programme recognize multilingual competence beyond competence in the designated immersion language?
- vii) Are there students who experience the model as creating additional barriers to subject learning?

9 Student experience and developmental appropriateness

- i) How do students experience learning cognitively demanding *Gymnasium* subjects in an additional language?
- ii) Do students perceive the programme's linguistic expectations as helpful, limiting or both?
- iii) Do students' needs change over the course of the programme? How?
- iv) Should language practices change as students become more proficient? How?
- v) Does the model provide sufficient scaffolding at earlier stages whilst also allowing increasing independence later?
- vi) What interplay is there between motivation, confidence and willingness to participate and language policy?

10 Variation between schools, cantons and classrooms

- i) How much variation is there beneath the label of *Gymnasium*-level bilingual/immersion learning?
- ii) Which differences represent legitimate contextual adaptation?
- iii) Which differences – if any – undermine comparability of programme quality?
- iv) Should evaluation focus on compliance with a common model or on whether local adaptations remain aligned with common programme aims?
- v) What minimum characteristics do programmes need to share before they can meaningfully be treated as following the same model?
- vi) At what point does contextual flexibility become conceptual incoherence?

11 Outcomes and evidence

- i) What evidence should count when evaluating the model?
- ii) Should evaluation privilege language attainment, subject attainment, both or a wider range of outcomes?
- iii) Over what time period should outcomes be assessed?
- iv) What comparison is appropriate: *Gymnasium* students not in the bilingual stream, students' development over time, programme aims or externally specified standards?
- v) How can evaluation avoid attributing outcomes to immersion when they may result from student selection, teacher characteristics or school context?
- vi) Should programme success include less easily quantifiable outcomes like learner autonomy, multilingual awareness or preparation for multilingual academic settings?

12 Policy

- i) Which aspects of the model should be specified at policy level and which should remain matters of teacher or school leader judgment?
- ii) How much standardization is necessary to accommodate cantonal and school-level differences?
- iii) What evidence should trigger a change in programme policy?
- iv) Should policy prescribe classroom language use or specify outcomes and allow teachers discretion over how they are achieved?
- v) How can evaluation results be translated into policy without treating practices that work in a single specific context as universally applicable?
- vi) What mechanisms should exist for programmes to adapt in response to evaluation results?