

Appendix F

De schijnwerper op taalbad

1. Inleiding

De oproep tekst van dit project verwees expliciet naar de organisatie van taalbad “of een volwaardig alternatief dat dezelfde resultaten bereikt” als een mogelijk effectief taaltraject.

Maar wat verstaan we onder een taalbad? Het Decreet basisonderwijs (2020) hanteert de volgende definitie: “Met taalbad wordt vanaf het schooljaar 2021-2022 bedoeld: intensieve onderwijsactiviteiten die tot doel hebben de leerling door onderdompeling in de onderwijstaal deze onderwijstaal te laten verwerven in functie van een snelle integratie in de reguliere onderwijsactiviteiten. Dit kan een voltijds traject zijn. Een leerling kan gedurende het basisonderwijs maximaal één schooljaar een voltijds taalbad of voltijds gelijkwaardig alternatief volgen.” Bij de doelgroep wordt er enkel vermeld dat het gaat om “leerlingen die het Nederlands onvoldoende beheersen.”

Om de brug met internationaal onderzoek op een transparante manier te kunnen maken, lijsten we hieronder de belangrijkste eigenschappen van een taalbad als organisatievorm op. Een taalbad onderscheidt zich van alternatieve vormen van taaltrajecten door de combinatie van volgende eigenschappen:

1. Duurtijd: Een taalbad is **beperkt** in de tijd en duurt maximaal **één jaar**.
2. Intensiteit: Een taalbad wordt mogelijk **voltijds** georganiseerd.
3. Setting: Een taalbad vindt **buiten de reguliere onderwijsactiviteiten** plaats, wat inhoudt dat de leerlingen niet in de reguliere klassen opgenomen worden.
4. Doelgroep: Een taalbad wordt georganiseerd voor leerlingen met een **lage taalvaardigheid** in de onderwijstaal.
5. Doeltaal: De doeltaal is het Nederlands of de **onderwijstaal**.
6. Instructietaal: De instructietaal is het Nederlands of de **onderwijstaal**.

2. Het effect van taalbad als mogelijke organisatievorm

De vraag naar het effect van taalbad (en alternatieve organisatievormen) van taaltrajecten zit vervat in de eerste onderzoeksvraag van dit project:

OV1 – Onder welke vormen kunnen kwaliteitsvolle taaltrajecten aangeboden worden?

Deze vraag werd samen met de tweede onderzoeksvraag behandeld op basis van een specifieke vorm van een systematische literatuurstudie (zie Deel 2 voor een verantwoording en beschrijving van de gekozen onderzoeksmethode en inclusiecriteria).

In de sample van 33 reviews vonden we in geen enkele review een rechtstreekse verwijzing naar een taalbad volgens de hierboven beschreven eigenschappen. Wel wordt in één review ingezoomd op de keuze van de instructietaal: Cheung en Slavin (2012a) vergelijken het effect van tweetalig onderwijs (bilingual English-Spanish programs) versus eentalig onderwijs (English-only programs) op de (Engelse) leesvaardigheid van leerlingen in het lager onderwijs met Spaans als thuistaal en stellen vast dat leerlingen in een tweetalig programma op korte termijn (na één schooljaar) doorgaans beter leren lezen in de onderwijstaal (Engels) dan leerlingen in een eentalig programma (voor een uitgebreidere beschrijving van dit onderzoek, zie Deel 2).

Deze bevindingen doen ons besluiten dat, in een context waar alle leerlingen met een taalachterstand in de onderwijstaal (Engels) dezelfde thuistaal (Spaans) spreken, een organisatievorm waarin zowel de thuistaal- als de onderwijstaal instructietalen zijn, meer geschikt is om de leesvaardigheid van deze leerlingen te verbeteren dan een organisatievorm waarin de onderwijstaal de enige instructietaal is.

De overige eigenschappen van taalbad, zoals de duurtijd, setting en intensiteit, worden noch in de review van Cheung en Slavin (2012a) noch in andere reviews expliciet onderzocht. Op basis van de overview vinden we met andere woorden geen evidentie voor of tegen het effect van taalbad zoals gedefinieerd door de overheid.

3. Verder zoeken naar onderzoeksevidentie voor taalbad

In de systematische zoektocht naar relevante publicaties voor de overview hielden we de scope bewust breed, vanuit de brede definiëring van het begrip ‘taaltraject’ en de verschillende profielen van leerlingen met nood aan taalsteun (zie Deel 1). We combineerden verschillende sets van zoektermen met betrekking tot de beoogde doelgroepen, taaldoelen, types studies en interventies (zie Appendix A) en vertrokken dus niet a priori van een bepaalde organisatievorm van een taaltraject.

Hoewel de gehanteerde zoekprocedure voldoende evidentie opleverde voor essentiële bouwstenen van kwaliteitsvolle taaltrajecten, stelde deze ons niet in staat om een uitspraak te doen over specifieke organisatievormen zoals een taalbad. Daarom voerden we na het afronden van Deel 2 een bijkomende systematische zoektocht uit die zich toespitste op het taalbad als organisatievorm. Gezien de vraag van de opdrachtgever, wilden we er zeker van zijn dat er, door de gekozen zoekprocedure, geen effectonderzoek naar taalbaden gemist zou kunnen zijn. Om de kwaliteit van onze adviezen over dit onderwerp te garanderen, kozen we er opnieuw voor om enkel naar reviews te zoeken, en dus niet naar primaire studies.

In maart 2021 doorzochten we de elektronische databank Web of Science Core Collection, op basis van volgende zoekopdracht.

TITLE	((submersion OR immersion OR “push-in” OR “pull-out” OR “push in” OR “pull out” OR “English-only” OR “instructional model” OR “dual language” OR “bilingual education” OR program OR “sheltered NEAR instruction” OR “transitional bilingual education” OR TBE OR EO OR SI OR SEI))
AND	
TOPIC	((“dual language” or DLL or bilingual or multilingual or “minority student” or EAL or ESL or ELL or “English as a second language” or “English as an additional language” or ESOL or “English language learner” or “English learner” or “limited English proficiency” or LEP or SLA or “second language” or “minority language” or “language minority” or “heritage language” or “linguistic diverse” or “culturally diverse” or CLD OR EL OR L2 OR immigrant OR refugee OR newcomer OR “newly NEAR arrived”))
AND	
TITLE	((meta-analysis OR “systematic review” or “research synthesis” or “systematic literature review” OR overview))
Timespan = 2000 – 2021	
Language = English	

Deze zoektocht leverde tien reviews op. Twee daarvan vonden we eerder ook tijdens de oorspronkelijke zoektocht. Beide reviews werden echter uitgesloten in de eerste fase van de selectieprocedure, aangezien ze op het vlak van methodologische vereisten niet aan de vooropgestelde inclusiecriteria voldeden (zie Deel 2 voor een uitgebreide rationale en beschrijving van deze criteria). De systematische review van Hur et al. (2020) focust op interventies rond beginnende geletterdheid bij tweetalige kleuters (dual language learners). De auteurs voeren een aparte analyse uit van (quasi-)experimentele studies en single case studies, maar brengen de baseline-equivalentie van de interventie- en controlegroep niet helder in kaart bij de studies met een quasi-experimenteel design. Reljić en collega’s (2015) onderzoeken het effect van tweetalige programma’s in Europa. Hoewel de auteurs initieel voorstellen dat de interventie- en controlegroepen in de geïncludeerde (quasi-)experimentele studies gerandomiseerd of gematcht moeten zijn (wat dus betekent dat de baseline-equivalentie nagegaan kan worden), schrappen ze dit criterium na een verkennende zoektocht omdat ze anders te weinig studies overhouden.

De acht overige reviews toetsten we aan dezelfde inclusiecriteria af. Opnieuw vonden we geen review die aan alle criteria voldeed. Vier studies vielen buiten de scope van dit onderzoek doordat ze focusten op het hoger onderwijs, buitenlandse studieprogramma’s in het kader van vreemdetalenonderwijs (study abroad programs) of op de opleiding van zorgverleners. De vier overige reviews focusten, net als de eerder weerhouden review van Cheung en Slavin (2012a), op het effect van de instructietaal, maar voldeden niet aan het inclusiecriteria ‘soorten studies’, hetzij doordat ze niet uitsluitend experimentele en quasi-experimentele studies met een treatment-control-design includeerden, hetzij doordat ze de baseline-equivalentie (zijn beide groepen bij de start van de interventie voldoende gelijk?) van de interventie- en controlegroep bij quasi-experimentele studies niet in kaart brachten.

Om geen Nederlandstalig onderzoek te missen, voerden we naast deze elektronische zoektocht ook een bijkomende manuele zoektocht uit. We gaven in Google Scholar zoektermen als 'review', 'taalbad' en 'schakelklas' in. We vonden enkele primaire effectstudies, maar geen systematische reviews die het effect van taalbaden of een gelijkaardige organisatievorm zoals een schakelklas op basis van (quasi-)experimentele studies samenvat.

Ook op basis van deze bijkomende elektronische en manuele zoektocht kunnen we besluiten dat tot op heden geen reviews met een rigoureuze methodologie (al dan niet met een component meta-analyse) bestaan die het effect van taalbaden meten.

4. Taalbaden onderbelicht in wetenschappelijk onderzoek, enkele oorzaken

Via de hierboven beschreven zoekprocedures vonden we noch in de internationale noch in de Nederlandstalige literatuur kwaliteitsvolle reviews die het effect van een taalbad meten volgens de definitie van het Decreet basisonderwijs. Hieronder concretiseren we enkele oorzaken van deze lacune aan de hand van voorbeelden.

Ten eerste is het **niet eenvoudig** om het effect van taalbaden (en alternatieve organisatievormen) **te meten**. Het meest geschikte design voor effectonderzoek is een experimentele studie met een interventie- en controlegroep waarin beide groepen bij de start van de interventie voldoende gelijk zijn (baseline-equivalentie) zodat een eventuele leerwinst aan de interventie toe te schrijven is en niet aan verschillen die vooraf al aanwezig waren (What Works Clearinghouse, 2017). Hoewel de initiële gelijkheid van de interventie- en controlegroep een aandachtspunt is voor elke effectstudie naar onderwijsprogramma's, vormt deze bij onderzoek naar taalbaden (en alternatieve organisatievormen) een extra uitdaging (Slavin & Cheung, 2005). Gerandomiseerde studies, waarin de leerlingen via een loterijstelsel aan de interventie- en controlegroepen toegekend worden, zijn bij dit type onderzoek doorgaans niet haalbaar. Bij het werken met bestaande groepen kan de baseline-equivalentie nagestreefd worden door een interventiegroep te koppelen aan een controlegroep die op een aantal fundamentele kenmerken gelijk scoort (*matching*), zoals percentage jongens/meisjes, leeftijd, vaardigheidsniveau in de onderwijstaal (op basis van pretest), vaardigheidsniveau in de thuistaal (op basis van pretest), aantal jaren blootstelling aan de onderwijstaal (bijvoorbeeld nieuwkomers of niet), SES-kenmerken enz. Zelfs wanneer onderzoekers met deze zichtbare verschillen rekening houden, kunnen echter ook onzichtbare verschillen ongelijkheid of *selection bias* veroorzaken (Dietrichson et al., 2020; Krowka et al., 2017; Slavin & Cheung, 2005). Een voorbeeld hiervan zijn programma's die (op vrijwillige basis) buiten de reguliere lestijden georganiseerd worden of programma's waarvoor ouders (of leerlingen) zelf kunnen kiezen (Slavin & Cheung, 2005). De kans bestaat dat leerlingen die aan dergelijke programma's deelnemen een ander profiel hebben dan leerlingen die niet aan deze programma's deelnemen, bijvoorbeeld omdat ze ouders hebben die meer betrokken zijn of omdat ze zelf meer intrinsiek gemotiveerd zijn. Dit kan op zijn beurt een invloed hebben op de schoolse prestaties. Hetzelfde geldt voor leerkrachten en/of scholen die bijvoorbeeld zelf beslissen om een bepaalde organisatievorm, zoals een taalbad, aan te bieden: het is niet ondenkbaar dat net die scholen die meer ervaring met de doelgroep hebben (en dus misschien meer expertise hebben), sterker geneigd zijn om dit te doen. Verder is ook attritie een uitdaging in het onderzoek naar taalbaden (en alternatieve organisatievormen), omdat een grote uitval in de interventie- of controlegroep ervoor kan zorgen dat groepen die bij de start van een interventie gelijk waren dit tegen het einde niet meer zijn (What Works Clearinghouse, 2017). In het kader van onderzoek naar taalbaden denken we hierbij aan leerlingen met een lage SES die vaker verhuizen dan leerlingen met een hoge SES (Scanlon & Devine, 2001), of leerlingen met bijzondere noden die tijdens een (langere) interventie van school veranderen.

Ten tweede wordt het onderzoeksveld rond taalbaden (en alternatieve organisatievormen) gekenmerkt door een **gebrek aan standaardterminologie**. Verschillende 'labels' worden door elkaar gebruikt, wat het samenvatten van onderzoek bemoeilijkt en de kans op methodologische fouten in reviewstudies vergroot (Hill, 2018; Reljić et al., 2015; Edvantia, 2007). Vooral de verschillende invullingen van het begrip immersie (*immersion*) zijn problematisch. In de Angelsaksische literatuur doelt men met immersie immers zowel op tweetalige programma's (*dual language immersion*) als op eentalige programma's (English immersion of de zogenaamde *English-only* aanpak). Binnen de tweetalige programma's bestaat verder nog het onderscheid tussen immersieprogramma's voor leerlingen met een andere thuistaal die de onderwijstaal leren (*language minority students*) en programma's voor leerlingen die thuis de dominante onderwijstaal spreken en daarnaast een tweede (of derde) taal leren (*language majority students*, zoals bijvoorbeeld het Canadese model of, dichter bij huis, de Nederlandstalige immersieprogramma's in Wallonië). Door dit inconsistente gebruik van terminologie worden soms positieve resultaten van programma's gerapporteerd voor groepen voor wie deze programma's in principe niet bedoeld zijn (Hill, 2018).

Ten derde zijn de concrete invullingen van taalbaden (en alternatieve organisatievormen) sterk **gebonden aan specifieke contexten** (Moughamian et al., 2009; Sierens & Van Avermaet, 2014). De verankering in tijd en/of ruimte zorgt ervoor dat het mogelijke effect van een bepaalde organisatievorm in de ene context niet zomaar te vertalen valt naar een andere context. Zo bestaat er in de VS een uitgebreide onderzoekstraditie naar het effect van tweetalig onderwijs, waarin de eerder beschreven review van Cheung en Slavin (2012a) gesitueerd moet worden. Omdat in de VS bijna 75 procent van de *English language learners* Spaans als thuistaal heeft (U.S. Department of Education, National Center for Education Statistics, 2018), kunnen de bevindingen uit onderzoek naar deze tweetalige programma's niet zomaar vertaald worden naar een context waarin veel scholen een zeer divers leerlingenpubliek hebben, zoals in Vlaanderen (Bialystok, 2018; Sierens & Van Avermaet, 2014). De Nederlandse context leunt met haar schakelklassen dan weer dichter aan bij de Vlaamse, maar ook hier is voorzichtigheid geboden bij het vertalen van eventuele onderzoeksresultaten. In een schakelklas krijgen leerlingen met een sterke taalachterstand sinds 2006 gedurende één schooljaar intensief taalonderwijs in een aparte groep, zodat ze daarna (opnieuw) in het reguliere onderwijs kunnen aansluiten (Mulder et al., 2008). Deze invulling van een taalbad benadert de Vlaamse definitie, al zijn er ook enkele belangrijke verschillen wat de intensiteit en doelgroep betreft. De schakelklassen kunnen voltijds, deeltijds of buiten het curriculum ('verlengde schooldag') ingericht worden en ouders moeten altijd toestemming voor inschrijving geven (zie eerder, *selection bias*). Verder worden de schakelklassen voor verschillende leeftijdsgroepen georganiseerd, zoals voor kleuters voorafgaand aan groep 3 (eerste leerjaar; VOET-klas), maar ook bijvoorbeeld tussen groep 3 en groep 4 (eerste en tweede leerjaar) of na groep 8 (zesde leerjaar; KOP-klas). Deze KOP-klas is meteen een bijzondere variant, omdat dit type schakelklas ingericht wordt voor leerlingen met 'groeipotentie' waarvan men verwacht dat zij na het traject in deze schakelklas hoger kunnen instromen in het secundair onderwijs (Mulder et al., 2008). Dit betekent dat de doelgroep van de Nederlandse schakelklassen niet helemaal representatief is voor de doelgroep die het Decreet voor ogen heeft.

Ten vierde wordt het onderzoek naar het effect van taalbaden gedomineerd door een **sterke focus op het effect van de instructietaal**. Deze dominantie is te verklaren vanuit het grote aantal studies uit de Angelsaksische context en de bewogen geschiedenis die de VS op dit vlak kent (voor een overzicht, zie bijvoorbeeld Bialystok, 2018; Edvantia, 2007; Parrish et al., 2006). Door deze eenzijdige focus op de instructietaal delen auteurs van reviewstudies bestaande organisatievormen doorgaans in twee categorieën in, ofwel tweetalig ofwel eentalig, zonder rekening te houden met de grote diversiteit aan programma's en contexten binnen beide categorieën. Dit heeft tot gevolg dat in sommige reviews programma's vergeleken worden die in principe niet met elkaar vergeleken kunnen worden, omdat ze zich bijvoorbeeld tot andere doelgroepen richten of omdat de context te verschillend is (opmerking van Hill, 2018; Sierens & Van Avermaet, 2014). Een concreet voorbeeld vinden we bij Reljić en collega's (2015) die in hun meta-analyse naar het effect van Europese tweetalige programma's op de leesvaardigheid van *language minority students* de volgende twee programma's includeren: een programma voor leerlingen met Turks als thuistaal in Nederland versus een programma voor leerlingen met Spaans en/of Baskisch als thuistaal in Spaans Baskenland. Verder zorgt de eenzijdige focus op het effect van de instructietaal ervoor dat het mogelijke effect van andere relevante eigenschappen van taalbaden, zoals de duurtijd, intensiteit (bijvoorbeeld voltijds of deeltijds), setting (buiten of binnen de reguliere onderwijsactiviteiten), doelgroep (leeftijd, vaardigheidsniveau in de thuistaal en/of de onderwijstaal) maar ook bijvoorbeeld de inhoud van het programma (enkel taalonderwijs of ook vakinhouden) niet onderzocht worden (National Academies of Sciences, Engineering, and Medicine, 2017). Hoewel er in de VS alleen al verschillende soorten taalbadprogramma's bestaan (bijvoorbeeld *Sheltered English Instruction*, *Structured English Immersion*, *Pull-out ESL*), elk met een eigen organisatie, inhoud en focus, worden eentalige programma's in de meeste reviews over het effect van de instructietaal enkel vernoemd als de *English-only* controlegroep (Slavin & Cheung, 2005; voor een beschrijving van de programma's, zie bijvoorbeeld Moughamian et al., 2009; National Academies of Sciences, Engineering, and Medicine, 2017; Office of English Language Acquisition, 2019).

Ten vijfde zorgt de focus op het effect van de instructietaal en de daarmee gepaard gaande binaire opdeling van programma's ervoor dat in reviews over dit thema **zelden aandacht is voor de gelijkenissen tussen programma's** in beide categorieën (Edvantia, 2007; Mulder et al., 2008). De verschillende eentalige en tweetalige programma's bevinden zich in werkelijkheid eerder op een continuüm met aan het ene uiterste het puur tweetalige model en aan het andere uiterste het puur eentalige model (Moughamian et al., 2009). Daartussen bevinden zich programma's met verschillende gradaties van het gebruik van de thuistaal. Zo zetten ook verschillende *English-only* programma's in op het gebruik van de thuistaal tijdens de instructie (Edvantia, 2007; Moughamian et al., 2009) of wordt in het *Transitional Bilingual Education (TBE)*-model de thuistaal enkel gedurende de eerste jaren als instructietaal ingezet, aangezien het doel van dit model uitsluitend de beheersing is van de onderwijstaal (in tegenstelling tot andere tweetalige programma's). Verder is het belangrijk om, los van de instructietaal of organisatievorm, te kijken naar wat er in de klas gebeurt, zoals bijvoorbeeld naar de kwaliteit van de instructie en het aangeboden curriculum (Cheung & Slavin, 2012a; Groothoff, 2020; National Academies of Sciences, Engineering, and Medicine, 2017; Sierens & Van Avermaet, 2014).

5. Besluit

Door een gebrek aan kwaliteitsvolle reviews kunnen we er geen uitspraak over doen in welke mate een taalbad zoals gedefinieerd in het Decreet basisonderwijs een geschikte organisatievorm is van taaltrajecten voor leerlingen met nood aan taalsteun. De sterke focus op het effect van de instructietaal in de afgelopen decennia laat in de wetenschappelijke literatuur weinig ruimte voor onderzoek naar andere relevante eigenschappen van taalbaden (en alternatieve organisatievormen), zoals duur, intensiteit, setting, doelgroep enz. Meer onderzoek, waarin het effect van verschillende organisatievormen van taaltrajecten in een Vlaamse context onderzocht wordt op basis van een (quasi-)experimenteel design met voldoende aandacht voor baseline-equivalentie, is nodig.

Daarnaast pleiten steeds meer onderzoekers er de laatste jaren voor om verder te kijken dan het type programma en te focussen op wat er in de klas gebeurt (Cheung & Slavin, 2012a; Groothoff, 2020; National Academies of Sciences, Engineering, and Medicine, 2017; Sierens & Van Avermaet, 2014). Niet enkel de krijtlijnen van een taaltraject zijn dus belangrijk; het is minstens even belangrijk om te kijken naar wat er tussen die krijtlijnen precies gebeurt.