KOHNSTAMM
INSTITUUT

NELI onderzocht in het gespecialiseerd onderwijs

EVALUATIEONDERZOEK NAAR DE NUFFIELD EARLY LANGUAGE INTERVENTION IN HET SO EN SBO



INEKE VAN DER VEEN
ANNEMIEK VEEN
JUDITH RISPENS
DOROTHÉ ELSHOF

© 2026 Kohnstamm Instituut UvA BV • ISBN 978-94-6321-223-6

CIP-gegevens KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Veen, van der I., Veen, A. & Elshof, D. Amstelveen: Kohnstamm Instituut.
Rispens, J.E. Amsterdam: Universiteit van Amsterdam.
NELI onderzocht in het gespecialiseerd onderwijs.
Evaluatieonderzoek naar de Nuffield Early Language Intervention in het so en sbo.
Amstelveen: Kohnstamm Instituut.
(Rapport 1158, projectnummer 40934)

ISBN 978-94-6321-223-6

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher.

Uitgave en verspreiding:
Kohnstamm Instituut
Keizer Karelplein 1, 1185 HL Amstelveen
Tel.: 020-214 1400
www.kohnstamminstituut.nl
Dataverwerking: Elion.nl
© Copyright Kohnstamm Instituut, 2026

Inhoudsopgave

1 Inleiding 1
1.1 (Verwachte) effecten en werkzame factoren 1
1.2 Programmakenmerken 2
1.3 Onderzoeksvragen 5

2 Methode 6
2.1 Onderzoeksontwerp 6
2.2 Verloop dataverzameling 10
2.3 Kenmerken van voor NELI geselecteerde leerlingen 12
2.4 Resultaten matching 13
2.5 Analysemethode 15

3 Resultaten effectiviteit interventie 19
3.1 Taalvaardigheid interventie- versus vergelijkingsgroep 19
3.2 Gedrag interventie- versus vergelijkingsgroep 21

4 Resultaten mediatieanalyse 23

5 Beschrijvende resultaten over de uitvoering van NELI 24
5.1 Achtergrond van de tutoren 24
5.2 Training 24
5.3 LanguageScreen 25
5.4 Organisatie van de begeleiding 25
5.5 Voortgang 25
5.6 Groeps- en individuele sessies 26
5.7 Registraties en screening 28
5.8 Didactische vaardigheden 29
5.9 Ervaren resultaten 29
5.10 Voor welke leerlingen is NELI het meest geschikt? 32
5.11 Wat is er nodig om NELI nog meer effecten te laten sorteren? 33
5.12 Transfer onderdelen van NELI naar de reguliere les 34
5.13 Knelpunten en pluspunten bij de uitvoering van NELI 35
5.14 Redenen om NELI aan te bevelen aan andere scholen 39
5.15 Aanbevelingen voor de ontwikkelaar 41

6 Resultaten effect interventie bij hogere uitvoeringskwaliteit 43
6.1 Uitvoering - kwantiteit 43
6.2 Uitvoeringskwaliteit 43
6.3 Aanwezigheid en betrokkenheid 44
6.4 Groep met gunstigere uitvoeringskwaliteit – indicatie van potentieel te behalen effect van NELI 44

7 Resultaten aanvullende taaltaken 46
7.1 Ontwikkeling per taaltaak 46
7.2 Ontwikkelingen naar achtergrondkenmerken 47
7.3 Ontwikkeling in relatie tot betrokkenheid en aanwezigheid 48

8 Haalbaarheid van een Randomized Controlled Trial in het s(b)o 50
8.1 Randvoorwaarden op schoolniveau 50
8.2 Inhoudelijke aansluiting 51
8.3 Implementatiekwaliteit en monitoring 52
8.4 Ontwerpkeuzes voor een RCT in het s(b)o 52
8.5 Belangrijkste inzichten 53

9 Samenvatting en slotbeschouwing 54
9.1 Samenvatting 54
9.2 Slotbeschouwing 56

Literatuur 59

Bijlagen 61
Bijlage I Aanvullende resultaten matching ten behoeve van LanguageScreen data 61
Bijlage II Resultaten matching ten behoeve van leerlingprofieldata 62
Bijlage III esultaten matching ten behoeve van mediatie-analyse 63
Bijlage IV – Interventiekinderen: ontwikkeling taaltaken en LanguageScreen naar achtergrondkenmerken 64
Bijlage V – Interventiekinderen: ontwikkeling taaltaken en LanguageScreen naar aanwezigheid en betrokkenheid bij deelname 65

1 Inleiding

Om scholen te helpen gefundeerde keuzes te maken en inzicht te geven in de effectiviteit van mogelijke aanpakken, worden meerdere als kansrijk getypeerde interventies onderzocht in het programma ‘Effectmeting kansrijke interventies in het po en vo’. Voor het speciaal (basis)onderwijs is echter weinig informatie beschikbaar over de effectiviteit van interventies. Daarom is de nieuwe NRO-ronde ‘Pilottrajecten interventies in het vmbo, pro en (v)so’ toegevoegd, gericht op deze sectoren. Een interventie waarvan de effectiviteit binnen het reguliere basisonderwijs al is onderzocht (Van Driel, Meeter & Van der Wilt, 2025), is de *Nuffield Early Language Intervention* (NELI), een in Engeland ontwikkeld taalverrijkingsprogramma waarin kleuters met de zwakste taalvaardigheid worden begeleid door een tutor met als doel de achterstand in woordenschat en mondelinge taalontwikkeling in te lopen. Dit is belangrijk, omdat achterstanden terugdringen het meest loont door te investeren in de vroege ontwikkeling van kinderen die dat het meeste nodig hebben (Heckman, 2013). In dit onderzoek is NELI geïmplementeerd voor kinderen van 4 t/m 7 jaar in het sbo en cluster 3 en 4. Omdat het hier gaat om een (deels) andere doelgroep, dan bij NELI namelijk kinderen met specifieke onderwijsbehoeften, en om ander onderwijs dan regulier basisonderwijs, is een apart onderzoek gevraagd naar de implementatie van NELI in het gespecialiseerd onderwijs. Dit onderzoek moet laten zien in hoeverre er aanwijzingen zijn dat deze interventie in het speciaal (basis)onderwijs effectief is (en zo ja voor wie en waardoor), of en onder welke condities het programma binnen het s(b)o uitvoerbaar en haalbaar is en of/hoe het programma middels een Randomized Controlled Trial (RCT: een onderzoek met random toewijzing) kan worden onderzocht.

1.1 De Nuffield Early Language Intervention (NELI)

NELI werd ongeveer 15 jaar geleden ontwikkeld door de Universiteit van Oxford. Het is een intensief tutoringprogramma waarin de zwakst presterende kinderen uit een kleutergroep gedurende 20 weken wekelijks drie kleinegroepsactiviteiten met maximaal zes kinderen en twee individuele activiteiten aangeboden krijgen aan de hand van een vast omschreven programma. Het programma is gestructureerd rondom zogenoemde ‘topic areas’, volgens onderbouwde principes voor het ondersteunen van luister- en vertelvaardigheid en het bevorderen van woordenschat. NELI functioneert als een aanvulling op het bestaande taalonderwijs. Spil van het programma is een speciaal getrainde tutor, doorgaans een onderwijsassistent. Het NELI-programma bevat ondersteunende hulpmiddelen, zoals een online training en een gedetailleerde set van lesplannen en lesmateriaal. Onderdeel van het programma is het screeningsinstrument LanguageScreen om snel en betrouwbaar de

mondelinge taalvaardigheid in beeld te brengen en te monitoren en mogelijke TOS-problematiek (TaalOntwikkelingsStoornis) te signaleren.
De online tutortraining is verplicht voor tutoren die met NELI gaan werken en bestaat uit twee onderdelen. In deel 1 staan de basisprincipes van taalverwerving bij jonge kinderen centraal. Deel 2 bereidt voor op de uitvoering. Elk deel duurt ongeveer vier uur.
Leerkrachten, directie en andere betrokkenen worden uitgenodigd en aangemoedigd om deel 1 van de onlinetraining te volgen, om een beeld te krijgen van de interventie en de tutoren te kunnen ondersteunen. In Groot-Brittannië is NELI beoordeeld als (kosten)effectieve aanpak voor het bestrijden van onderwijsachterstanden en is daarom grootschalig ingevoerd. Gerandomiseerde experimenten lieten zien dat een interventie van twintig weken, tot twee maanden vooruitgang oplevert in de taalontwikkeling die bij een interventie van dertig weken oploopt naar vier maanden (Sibieta, Kotecha & Skipp, 2016; West e.a., 2021). Uitgedrukt in effectgrootte bedroeg het effect van NELI-deelname in het onderzoek van West e.a. (2021) op taalvaardigheid .26 (klein effect). In de Nederland werd de effectiviteit van NELI onderzocht in een RCT onderzoek op reguliere basisscholen die extra financiering ontvangen vanuit het onderwijsachterstandenbeleid (Van Driel e.a., 2025). In dit onderzoek werd eveneens een positief, maar lager effect van .15 gevonden.
In het Britse onderzoek werd ook het effect van NELI op de LanguageScreen gerapporteerd; dit bedroeg .32 (West e.a., 2021). In de Nederlandse studie werd het effect op de LanguageScreen niet onderzocht. Wel werd in de Nederlandse studie, in tegenstelling tot de Britse studie, het effect van NELI op narratieve vaardigheden onderzocht. Hiervan werd een positief effect van .19 gevonden. Naast taalvaardigheid vonden West e.a. (2022) ook een verbetering in gedrag als gevolg van NELI-deelname (effectgrootte .23). Dit aspect werd in het Nederlandse onderzoek niet onderzocht.

1.2 Programmakenmerken: (verwachte) werkzame factoren

De effectiviteit van NELI kan naar verwachting worden verklaard vanuit enerzijds specifieke programmakenmerken en anderzijds mediërende factoren die de kenmerken van het programma verder ondersteunen. Daarnaast kunnen modererende factoren zijn: kenmerken die de effectiviteit van het programma beïnvloeden. In deze paragraaf worden deze factoren achtereenvolgens besproken.

Extra tijd en persoonlijke begeleiding
De interventie bestaat uit (dagelijks) extra tijd voor doelgerichte mondelinge taal-activiteiten (High Dosage Tutoring), in combinatie met een kleinschalige setting (Small Group Tutoring). Kinderen hebben zo niet alleen vaker en meer instructie en oefening dan andere leerlingen, maar ook beter op hun niveau afgestemde instructie, oefening en feedback, wat weer bijdraagt aan hogere concentratie en betrokkenheid. De *time on task* is dus hoger dan bij leerlingen die deze interventie niet volgen. Verwacht mag worden dat die componenten eveneens het succes van de interventie verklaren.

Effectieve didactieken voor taalstimulering
NELI is gebaseerd op eerder aangetoonde effectieve didactieken voor taalstimulering: een combinatie van taalaanbod, taalproductie en taalfeedback. Gericht taalaanbod is van belang om kinderen met nieuwe woorden in aanraking te brengen, die zij moeilijk zelf oppikken uit de context. De interventie werkt vanuit ‘robust vocabulary instruction’ (Beck, McKeown &

Kucan, 2013), die inzet op diepe woordkennis en waarin de didactische principes van voorbewerken, semantiseren, consolideren en controleren gevolgd worden. Naast het leren verstaan en begrijpen van nieuwe woorden, is taalproductie en samen doordenken (sustained shared thinking) van groot belang voor de taalontwikkeling (Siraj-Blatchford & Sylva, 2004; Henrichs, 2010). De interventie voorziet daarin door kinderen actief tot taalproductie, samen doordenken en narratieve taken uit te dagen. Door de kleinschalige setting hebben kinderen meer ruimte om zelf te spreken dan in een grotere groep. Dit bevordert goed afgestemde taalfeedback door de tutor. Vanuit deze effectieve didactieken kan verwacht worden dat zowel de receptieve als de productieve taalvaardigheid van kinderen toeneemt, op niveau van woordenschat, zinsconstructie, en narratieve vaardigheden.

Gunstige omstandigheden voor positief en aangepast gedrag
Naast de positieve effecten van meer *time on task* en effectieve taaldidactiek, kan NELI ook gunstige omstandigheden creëren voor het gedrag van leerlingen. Zoals beschreven, bleek dat NELI tot positieve gedragsverandering kan leiden (West e.a. 2022). In dat onderzoek werd een brede maat voor aangepast gedrag gehanteerd, waarin gedragsregulatie (met o.a. aspecten van impulscontrole), aandacht, externaliserend probleemgedrag en sociaal gedrag werden meegenomen. Het is daarmee onduidelijk op welk van de genoemde aspecten het effect vooral betrekking had.

Binnen NELI worden aandacht en actief luisteren gestimuleerd. De op niveau afgestemde taalactiviteiten en de herhaling van woorden/begrippen en activiteiten vergroten de kans op succeservaringen. Dit kan ertoe bijdragen dat leerlingen met meer aandacht aan hun taken werken. Tutoren krijgen instructie over hoe zij goed luistergedrag bij kinderen kunnen bevorderen en belonen. In de kleine groep krijgen leerlingen veel ruimte om zelf te spreken en hun gedachten te formuleren. Deze aspecten kunnen de werkhouding van kinderen en hun impulscontrole versterken.

De kleinschaligheid en regelmatigheid van de tutorsetting, inclusief één-op-één sessies en het samen doordenken bevordert het contact tussen de tutor en het kind en kan hun onderlinge relatie versterken. Dit kan vervolgens ook andere relaties positief beïnvloeden. Deelname aan NELI kan er dus toe leiden dat het gedrag van leerlingen positief verandert: prosociaal gedrag wordt gestimuleerd en probleemgedrag neemt af.

Mediërende factoren
Mogelijk zijn verbeteringen in taal door de interventie (deels) het gevolg van positieve veranderingen in gedrag die met de interventie samenhangen. In dat geval zou gedragsverbetering een mediërende factor zijn. West e.a. (2022) onderzochten dit voor de NELI-interventie, maar vonden hiervoor geen aanwijzingen. Een andere mogelijkheid is dat juist taalontwikkeling zelf fungeert als mediator. Achterblijvende taalontwikkeling kan leiden tot sociaal-emotionele en gedragsproblematiek (West e.a., 2021). Wanneer kinderen taalvaardiger zijn, kunnen zij beter communiceren over hun behoeften en wensen. Dat vergemakkelijkt de regulering van emoties en draagt bij aan een vermindering van internaliserend en externaliserend probleemgedrag (o.a. Curtis e.a., 2019). West e.a. (2022) onderzochten daarom ook of taalontwikkeling binnen de NELI-interventie een mediërende rol speelde bij gedragsverbeteringen. Ook hiervoor vonden zij geen aanwijzingen. De

positieve effecten van NELI op taalvaardigheid en gedrag bleken onafhankelijk van elkaar te bestaan en naast elkaar op te treden.

Modererende factoren
In de Engelse studie (West e.a., 2021) zijn geen significante verschillen in effectiviteit van de interventie gevonden tussen jongens en meisjes, en ook niet tussen eerste- en tweedetaalverwervers van het Engels. De groep die bij aanvang de zwakste taalvaardigheid had, profiteerde daarentegen wel iets meer van de interventie dan de groep leerlingen met een iets sterkere taalvaardigheid bij aanvang. Deze bevindingen suggereren dat de interventie in Nederland het meest effect kan bereiken bij de leerlingen met de zwakste taalvaardigheid in het Nederlands, ongeacht of dat hun moedertaal is. In het Nederlandse effectonderzoek naar NELI in het reguliere basisonderwijs werden hiervoor echter geen aanwijzingen gevonden: er bleek geen verband tussen de effectiviteit van NELI en het beginniveau van taalvaardigheid, noch met geslacht of het al dan niet spreken van Nederlands als thuistaal (Van Driel e.a., 2025).

Een andere factor die de mate van effectiviteit kan beïnvloeden is externaliserend en mogelijk ook internaliserend probleemgedrag bij de start van de interventie. De effectiviteit van de interventie wordt mogelijk negatief beïnvloed wanneer in hoge mate sprake is van probleemgedrag. Op cluster 4 scholen komt deze problematiek meer voor, aangezien deze scholen onderwijs bieden aan leerlingen met (ernstige) gedragsstoornissen en/of psychische problematiek. Internationaal onderzoek toont het belang van een positieve leerkracht-leerlingrelatie voor sociale gedragsregulatie en cognitieve ontwikkeling van kinderen. Een positieve relatie is van grote invloed op het welbevinden en er is steeds meer evidentie dat welbevinden een fundamentele voorwaarde is om tot leren en ontwikkelen te komen (Ryan & Deci, 2017; Lee-Pang e.a., 2025). Wanneer leerkrachten moeite hebben een positieve band te creëren met kinderen, neemt het welbevinden van leerkracht en kind af (Hargreaves, 2000). Daarmee verminderen de kansen voor kinderen om te leren (Hamre & Pianta, 2001). Met name externaliserend probleemgedrag is voor leerkrachten een belangrijke factor van stress, vanwege de regelmatige verstoringen van het groepsproces of het optreden van conflicten. Dat heeft een negatieve invloed op de relatie tussen leerkracht en kind en daarmee ook op de kwaliteit van interactie tussen leerkracht en leerling (Pianta, Hamre, Stuhlman, 2003). Internaliserend probleemgedrag van kinderen heeft minder directe invloed op de leerkracht zelf, maar heeft toch ook effecten op de relatie tussen leerkracht en leerling. Doordat de leerling weinig contact zoekt en weinig spontaan vertelt, heeft de leerkracht minder zicht op de persoonlijkheid en de ontwikkeling van het kind. Bovendien neemt het kind vaak minder actief deel in (kleine)groepsactiviteiten, waardoor het minder kans heeft om taal productief te gebruiken.

Effecten van de interventie hangen ook af van hoe goed de interventie geïmplementeerd wordt (Carroll e.a., 2007). In Engeland bleek de implementatiekwaliteit hoog. Onderdelen van deze zogenaamde *Implementation fidelity* zijn de getrouwheid aan de interventie, de dosering, en de kwaliteit van hoe de inhoud overgebracht wordt.

Ten slotte kunnen zich ook negatieve neveneffecten voordoen. Leerlingen worden bijvoorbeeld voor de interventie uit de klas gehaald en missen daardoor andere lesonderdelen.

Toepassing van NELI in het so en sbo
De implementatie van NELI in het gespecialiseerd onderwijs verloopt mogelijk anders dan in het reguliere basisonderwijs. In het so en sbo zijn relatief weinig 4-7-jarigen. De meeste kinderen worden namelijk op wat latere leeftijd vanuit het reguliere basisonderwijs naar het s(b)o verwezen (Borggreve & de Lange, 2019). De groep kinderen die daar al gelijk onderwijs volgt, heeft daarom een relatief zwaardere en complexere problematiek (Koopman & Ledoux 2020). Dat betekent bijvoorbeeld dat de tutoren, meer dan in het reguliere basisonderwijs, rekening moeten houden met gedragsproblemen en daarmee om moeten kunnen gaan. Dit wordt bij de implementatie van NELI in het gespecialiseerd onderwijs deels ondervangen door de inzet van eigen personeel (onderwijsassistenten). Binnen het reguliere basisonderwijs wordt binnen NELI met groepjes gewerkt van maximaal zes kinderen. In het gespecialiseerde onderwijs is een kleinere groepsgrootte mogelijk te prefereren. Daarnaast bevat cluster 3 verschillende soorten scholen: voor kinderen met lichamelijke en voor kinderen met cognitieve beperkingen (waarbij de laatstgenoemde groep het grootst is). De verschillen met het reguliere onderwijs roepen de vraag op onder welke condities NELI binnen het s(b)o uitvoerbaar en haalbaar is en of/hoe het programma middels een Randomized Controlled Trial (RCT) kan worden onderzocht.

1.3 Onderzoeksvragen

Gezien de diversiteit in specifieke onderwijsbehoeften van de leerlingen in het sbo en clusters 3 en 4 van het so, de kleinere leerlingaantallen, wordt een evaluatieonderzoek gevraagd, waarbij de volgende vragen centraal staan:

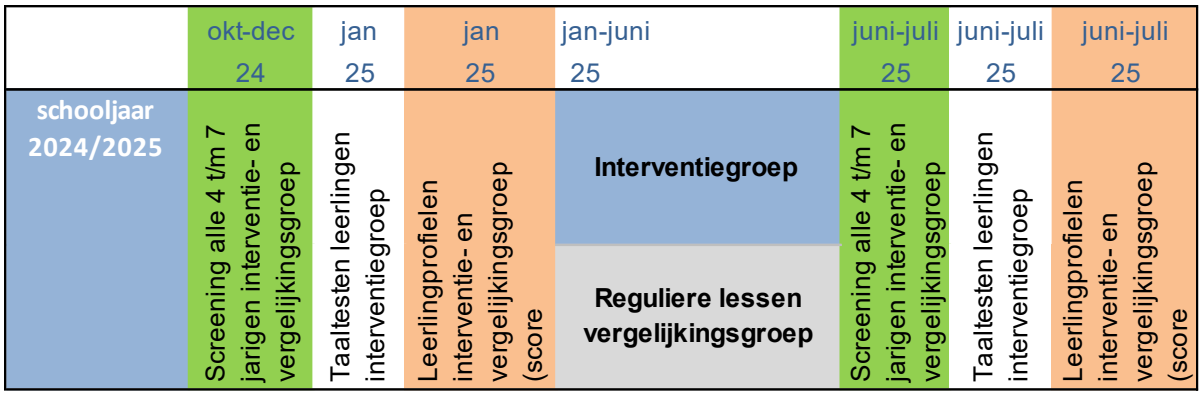
1. Wat zijn de (mogelijke) effecten van NELI op de taalvaardigheid en sociaal-emotionele ontwikkeling van leerlingen in het sbo en het so cluster 3 en 4? Welk type leerling is het meest gebaat bij de interventie?
2. Wat zijn de verklarende en werkzame factoren voor de effectiviteit? In hoeverre is er samenhang tussen de effecten op de sociaal-emotionele ontwikkeling en taalvaardigheid?
3. Wat is de relatie tussen implementatiekwaliteit en effectiviteit van NELI? Zijn er onbedoelde/negatieve effecten van de interventie?
4. In hoeverre is de interventie haalbaar en implementeerbaar in de genoemde schooltypen?
5. Kan de effectiviteit van de interventie onderzocht worden met een RCT-design en zo ja, hoe?

2 Methode

2.1 Onderzoeksontwerp

Onderdeel van de interventie is de afname van het screeningsinstrument LanguageScreen bij alle kleuters voorafgaand aan en na afloop van de uitvoering van de interventie. Om de belasting van scholen met dit onderzoek zo laag mogelijk gehouden, maken we voor de vergelijking van taalvaardigheid tussen de interventie- en vergelijkingsgroepleerlingen gebruik van dit instrument. Deelnemers aan NELI worden geselecteerd op basis van hun gestandaardiseerde score op de LanguageScreen. Daarbij wordt de volgende indeling gehanteerd. Alleen leerlingen met een score lager dan 90 komen in aanmerking voor deelname[1]. Het gemiddelde van de standaardscore is 100 de standaarddeviatie 15. Een score van 90 ligt dus tweederde standaarddeviatie onder het gemiddelde. De scores zijn in twee groepen ingedeeld: leerlingen die mogelijk extra ondersteuning kunnen gebruiken (score 82-89) en leerlingen die duidelijk extra ondersteuning kunnen gebruiken (score lager of gelijk aan 81). Aan de leerkrachten van alle leerlingen met een score lager dan 90 op de LanguageScreen, dus leerlingen die in aanmerking komen voor deelname aan NELI, is gevraagd om via een leerlingprofiel, voorafgaand aan en na afloop van de interventie per leerling een oordeel te geven over gedrag en informatie te geven over enkele achtergrondkenmerken. Ook deze oordelen kunnen we dus vergelijken tussen leerlingen in de interventie- en vergelijkingsgroep. Daarnaast zijn aanvullend taaltesten afgenomen bij de leerlingen in alleen de interventiegroep. Dit diende twee doelen: verdiepende informatie verkrijgen over de verandering in taalvaardigheid in de interventiegroep én het valideren van de LanguageScreen in het sbo en so. De aanvullende taaltesten werden uitgevoerd door medewerkers vanuit het onderzoek. Redenen om niet ook bij de vergelijkingsgroep extra taalmetingen uit te voeren, zijn ten eerste dat extra metingen te belastend zijn voor de vergelijkingsgroep zonder duidelijke opbrengst (geen interventieaanbod). Op basis van (alleen) de LanguageScreen bij de vergelijkingsgroep worden voldoende gegevens verkregen die nodig zijn om inzicht te krijgen in de eerste effecten bij de interventiegroep. Ten tweede kan deze opzet de haalbaarheid van een RCT voldoende worden onderzocht (onderzoeksvraag 4).

1 Tijdens het opstellen van het onderzoeksvoorstel waren we nog niet op de hoogte van het absolute inclusiecriterium voor deelname aan NELI. Daarom stelden we aanvankelijk voor om in een vroeg stadium een vergelijkingsgroep samen te stellen op basis van matching. Dit bleek echter niet nodig, omdat het absolute criterium al beschikbaar was en daarmee de onderzoeksgroep kon worden afgebakend.



Matchingsprocedure

Dat een RCT niet wordt ingezet, zorgt ervoor dat verschillen tussen wel- en niet-deelnemers niet zonder meer aan interventiedeelname zijn toe te schrijven. Het grote voordeel van RCT is namelijk dat -mits goed uitgevoerd- verschillen tussen wel- en niet-deelnemers daar wel aan kunnen worden toegeschreven (bijv. Hariton & Locasio, 2018). Om dit toch zoveel mogelijk te kunnen doen, hebben we matching uitgevoerd. Met deze techniek worden kenmerken van leerlingen in de interventie- en vergelijkingsgroep zoveel mogelijk gelijk gemaakt. We hebben hierbij gekozen voor *coarsened exact matching* (CEM, Iacus, King & Porro, 2012). Binnen deze matchingsmethode wordt -anders dan bij andere matchingsmethoden als *Propensity Score Matching* (PSM) naar exacte overeenkomsten gezocht op een set van vooraf bepaalde combinaties van covariaten, waaronder de score op de eerste screeningstoets. PSM, waarin in tegenstelling tot bij CEM wel aannames worden gedaan over verdelingen van de covariaten waarop wordt gematcht, is bekritiseerd, omdat het zeker (zoals hier) bij kleinere aantallen kan leiden tot onbetrouwbare resultaten (King & Nielsen, 2018). Hoewel CEM de vergelijkbaarheid tussen groepen aanzienlijk kan vergroten, blijft het mogelijk dat ongeobserveerde verschillen tussen leerlingen een rol spelen en daarmee (een deel van) de gevonden effecten kunnen verklaren.

Instrumenten

De *LanguageScreen* is een gevalideerd en (binnen Groot-Brittannië) genormeerd taalscreeningsinstrument dat binnen 10 minuten via een mobiele app is af te nemen door een onderwijsassistent of leerkracht (Hulme e.a., 2024). In Nederland wordt een door de aanbieder van NELI vertaalde versie gebruikt, die ook is ingezet binnen een RCT studie naar NELI in het reguliere basisonderwijs (Van Driel, Meeter & Van der Wilt, 2025). Er zijn vier subtests: expressieve en receptieve woordenschat, luisterbegrip en zinsherhaling. Bij de expressieve woordenschattaak worden kinderen gevraagd om afbeeldingen te benoemen. De receptieve woordenschattaak toetst het begrip van woorden door kinderen te laten kiezen uit meerdere plaatjes bij een gesproken woord. Bij de subtest luisterbegrip luisteren kinderen naar korte verhalen en beantwoorden daar vragen over (meting semantisch en syntactisch begrip). Bij de zinsherhalingstaak reproduceren kinderen zinnen van oplopende complexiteit, wat inzicht geeft in werkgeheugen en grammaticale verwerking. Hoewel de LanguageScreen ontwikkeld is om een totaaloordeel te vormen over de mondelinge taalvaardigheid, zijn ook de afzonderlijke subtests voldoende betrouwbaar bevonden (tussen .74 en .80; Hulme e.a., 2024).

Bij de interventiegroep zijn *extra taaltaken* afgenomen om de taalvaardigheid te meten, met twee doelen:

1. om de validiteit van de LanguageScreen voor Nederlandse leerlingen in het sbo en so te onderzoeken
2. om meer diepgaande informatie te verkrijgen over de taalontwikkeling van de interventiekinderen.

Er zijn vier taaltaken afgenomen. Deels komen deze overeen met de instrumenten die gebruikt zijn in de evaluatiestudie van Van Driel et al. (2025) naar de effecten van NELI bij kinderen in het regulier onderwijs en studies naar de effecten van NELI in Groot-Brittannië (Hulme et al., 2025 en Fricke et al., 2013). Deze taken zijn gekozen omdat de LanguageScreen ook (kort) deze taalvaardigheden meet (woordenschat, begrijpend luisteren en vertellen) en omdat dit de taalvaardigheden zijn waar de NELI interventie op gericht is. We sluiten met de gekozen taaltaken dus zo nauw mogelijk aan bij het Britse evaluatieonderzoek waarmee effectiviteit van NELI werd aangetoond (West e. a., 2021), en bij de Nederlandse studie (Van Driel et. al (2025). De testen zijn in het Nederlands verkrijgbaar, gestandaardiseerd en geschikt voor kinderen met een taalontwikkelingsstoornis. Hoewel de populatie met leerlingen in het s(b)o hier niet helemaal gelijk is aan die in het Britse evaluatieonderzoek, zijn de voorgestelde uitkomstmaten wel gestandaardiseerd, valide en betrouwbaar en geschikt voor kinderen in de leeftijd van onze doelgroep. Daarnaast hebben we een meting van narratieve vaardigheden en taalbegrip opgenomen, ook al zijn deze niet gemeten in het Engelstalige evaluatieonderzoek. De reden is dat beide vaardigheden een belangrijke rol spelen in de algemene taalvaardigheid van kinderen en de interventie hier deels op gericht is. De totale testtijd was niet langer dan 20 minuten, wat gemiddeld genomen haalbaar is voor deze leeftijdsgroep (Veen e.a., 2015).

De volgende vier taaltaken zijn afgenomen:

1. Actieve Woordenschat taak (CELF[2]-Preschool-2-NL; Wiig e.a., 2012)
   Deze taaltaak is onderdeel van de CELF-Preschool-2-NL en meet de expressieve woordenschat van jonge kinderen. Het kind krijgt afbeeldingen aangeboden en wordt gevraagd het bijbehorende woord te benoemen. Deze taak meet de vaardigheid om lexicale kennis actief te gebruiken en geeft inzicht in de productieve taalontwikkeling.
2. Zinnen begrijpen (CELF-Preschool-2-NL)
   De subtest Zinnen begrijpen is ook onderdeel van de CELF-Preschool-2-NL en richt zich op het receptieve taalbegrip van kinderen. Het kind hoort een zin en moet vervolgens de juiste afbeelding selecteren die overeenkomt met de betekenis. Deze taak toetst het vermogen om met name grammaticale structuren en semantische relaties te verwerken, en vormt daarmee een belangrijke indicator voor de ontwikkeling van syntactisch en semantisch begrip.
3. Het Busverhaal (RTNA; Jansonius e.a., 2014)
   Het Busverhaal is een narratieve taak binnen de RTNA (Renfrew Test voor Narratieve Analyse) die de productieve vertelvaardigheden van kinderen onderzoekt. Het kind krijgt een reeks afbeeldingen te zien en aan de hand van die plaatjes vertelt de testleider het kind een verhaal. Na afloop wordt het kind gevraagd dit verhaal na te vertellen. De taal

[2] Clinical Evaluation of Language Fundamentals, een test voor de diagnose en evaluatie van taalproblemen bij kinderen.

die het kind produceert wordt geanalyseerd op lengte en grammaticale rijkheid van de uiting. Het instrument biedt zo inzicht in de grammaticale productiviteit van een kind. Deze taak sluit aan bij de MAIN, in Van Driel et. Al (2025) en de 'Squirrel Story' in Fricke et al., (2013).

4. De Actieplaten (RTNA)
   De Actieplaten, ook van de RTNA, zijn visuele stimuli die ontworpen zijn om spontane taalproductie uit te lokken. Kinderen beschrijven wat er op de platen gebeurt, waarbij de focus ligt op het gebruik van werkwoorden, zinsstructuren en de vaardigheid om dynamische situaties verbaal te representeren. Deze taak geeft informatie over de expressieve syntaxis, het uitdrukken van betekenisvolle relaties, en het vermogen om taal in functionele contexten toe te passen.

In onderstaande tabel worden de samenhangen tussen de scores op de taaltaken en LanguageScreen voor de eerste meting weergegeven. Te zien is dat de correlaties hoog zijn, wat de validiteit van de LanguageScreen voor so en sbo leerlingen ondersteunt.

Tabel 2.1 - Meting 1 interventieleerlingen: samenhang tussen de ruwe scores op de LanguageScreen en taaltaken

		Actieve Woorden-schat taak	ActiePlaten semantiek	Zinnen begrijpen	ActiePlaten grammatica	Busverhaal lengte
LanguageScreen	corr.	0.723	0.519	0.718	0.589	0.415
	p	0.000	0.000	0.000	0.000	0.000
	n	106	92	105	101	91
Expressieve woordenschat	corr.	0.705	0.336	0.561	0.414	0.326
	p	0.000	0.001	0.000	0.000	0.002
	n	106	92	105	101	91
Receptieve woordenschat	corr.	0.499	0.389	0.621	0.372	0.178
	p	0.000	0.000	0.000	0.000	0.091
	n	106	92	105	101	91
Luisterbegrip	corr.	0.612	0.551	0.564	0.561	0.426
	p	0.000	0.000	0.000	0.000	0.000
	n	106	92	105	101	91
Zinnen herhalen	corr.	0.450	0.327	0.546	0.519	0.369
	p	0.000	0.001	0.000	0.000	0.000
	n	106	92	105	101	91

Leerlingprofiel: gedrag en achtergrondkenmerken. Voor het meten van gedrag maakten we gebruik van het gevalideerde pre-COOL instrumentarium (o.a. Veen e.a., 2015). Dit is ook geschikt voor kinderen in het s(b)o: een deel van de deelnemers aan pre-COOL ging naar het s(b)o. Het in pre-COOL gebruikte kleuter-/leerlingprofiel bevat items over onder meer gedrag, sociale competentie en taalvaardigheid. Specifiek zijn hieruit de volgende schalen voorgelegd aan de groepsleerkrachten: werkhouding (pre-COOL Consortium, 2012), CBQ (impulscontrole Rothbart e.a., 2001), CBCL (internaliserend, externaliserend probleemgedrag; Verhulst e.a., 1997), SCOL (opkomen voor jezelf, rekening houden met een ander; Joosten e.a., 2007). Aan het profiel van de eerste meting hebben we ten behoeve van het onderzoeken van modererende factoren ook een vraag naar thuistaal en specifieke onderwijsbehoefte toegevoegd. In het profiel van de tweede meting is een vraag opgenomen over algemene aanwezigheid in het afgelopen kalenderjaar: met name leerlingen in het so

kunnen hierin sterk verschillen vanwege bijvoorbeeld het volgen van therapie/behandelingen, wat hun onderwijsprestaties kan beïnvloeden.

Deelname en betrokkenheid. Leerlingen kunnen het meest profiteren van een programma wanneer zij intensief deelnemen en betrokkenheid tonen bij de activiteiten. Via een korte vragenlijst is de aanwezigheid en betrokkenheid (op individueel niveau) van de leerlingen bij de individuele en de groepssessies tweemaal gedurende de interventieperiode geïnventariseerd. Met deze gegevens kan onderscheid gemaakt worden tussen leerlingen die weinig en veel tutoruren hebben ontvangen en leerlingen die volgens de tutor in hoge mate of minder betrokken waren.

Tutorvragenlijst. Aan de tutoren werd een vragenlijst voorgelegd waarmee systematisch via grotendeels gesloten en aanvullend open vragen in kaart werd gebracht in hoeverre de interventie werd uitgevoerd zoals bedoeld, welke belemmeringen werden ervaren en welke omstandigheden en factoren de uitvoering bevorderden.

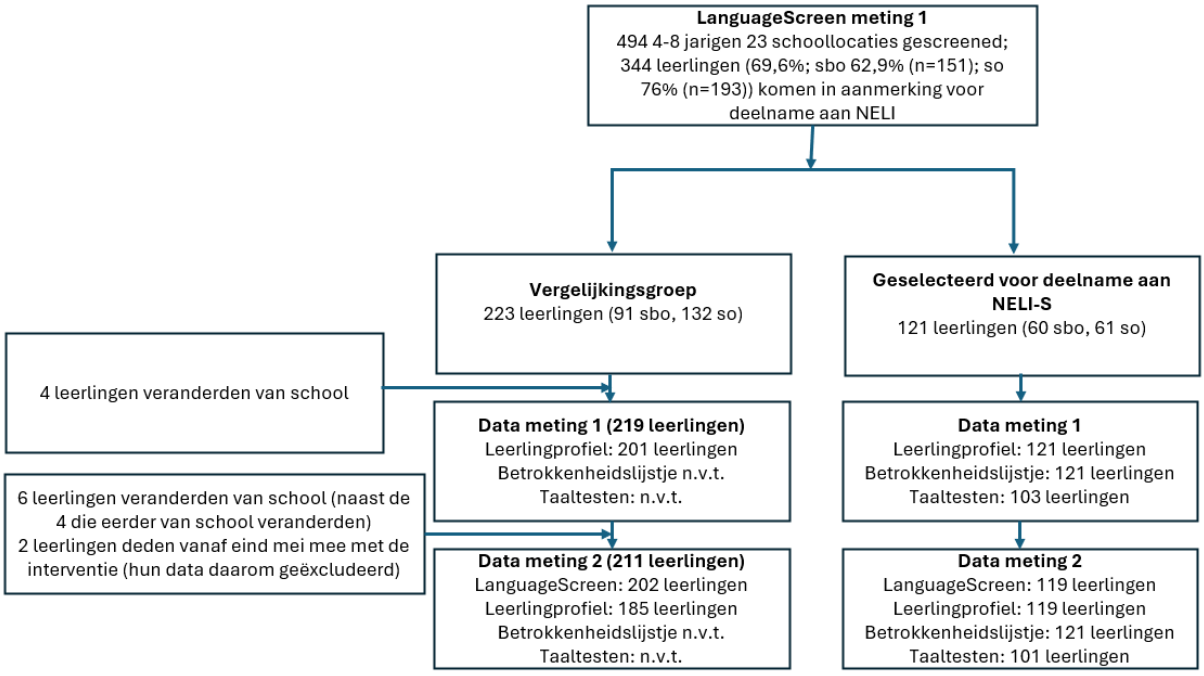
2.2 Verloop dataverzameling

In de maanden oktober-december 2024[3] werd op de deelnemende scholen de LanguageScreen afgenomen bij 494 leerlingen op 23 so en sbo schoollocaties. Op basis van hun score, kwam ruim twee derde (69,6%, n= 344) van de leerlingen in aanmerking voor deelname aan het NELI programma. 121 van de leerlingen die in aanmerking kwamen, werden geselecteerd voor deelname. Nagenoeg de helft van de leerlingen betrof so leerlingen, voor het grootste deel cluster 3[4], en de andere helft leerlingen uit sbo (61 so, 60 sbo). De uiteindelijke beslissing over deelname lag bij de school. Er werden vier leerlingen geselecteerd voor deelname, die hier op basis van hun score op de LanguageScreen niet voor in aanmerking kwamen. De gegevens van deze leerlingen nemen we niet mee in de effectanalyses.
Figuur 2.1 bevat een overzicht van de aantallen leerlingen in de interventie- en vergelijkingsgroep en de per meting verzamelde gegevens.

Meting 1. Voor alle 121 leerlingen werden bij de eerste meting (periode november 2024-januari 2025) leerlingprofielgegevens en gegevens over hun betrokkenheid bij deelname en aanwezigheid ingevuld. Bij 100 van de 121 deelnemende leerlingen werden bij meting 1 taaltesten afgenomen. Op één school werden bij een van de NELI-groepjes geen taaltesten afgenomen in verband met belasting; op een andere school met maar enkele deelnemende leerlingen werden geen taaltesten afgenomen.
Van 201 van de 223 leerlingen die niet voor deelname aan NELI werden geselecteerd, maar hiervoor op basis van hun score op de LanguageScreen wel in aanmerking kwamen, werd bij de eerste meting een leerlingprofiel ingevuld. Vier van de 22 leerlingen voor wie geen profiel werd ingevuld, bleken doorverwezen naar een andere school. Bij acht van de 22 leerlingen ging het om een klas waarvoor geen leerlingprofielen werden ingevuld; bij deze klas werd alleen de LanguageScreen afgenomen.

3 Bij enkele (6) leerlingen werd de LanguageScreen in de eerste helft van januari afgenomen.
4 Tien (ook) cluster 4 leerlingen werden geselecteerd voor deelname.

Meting 2. De tweede meting vond plaats in de periode juni-juli 2025. Bij 119 van de 121 aan NELI deelnemende leerlingen werd opnieuw de LanguageScreen afgenomen en werd het leerlingprofiel ingevuld. Bij 106 van de 121 deelnemende leerlingen werden de taaltjes afgenomen. Voor alle 121 aan NELI deelnemende leerlingen werden gegevens over betrokkenheid en aanwezigheid ingevuld. Eind mei startten twee leerlingen (uit de vergelijkingsgroep) met deelname aan NELI, ter vervanging van twee deelnemende leerlingen die deelnamen hoewel zij op grond van hun score op de LanguageScreen niet in aanmerking kwamen. Deze leerlingen zijn uit de vergelijkingsgroep verwijderd, maar niet toegevoegd aan de interventiegroep, omdat de deelname hiervoor te kort duurde. Uiteindelijk bleven in de vergelijkingsgroep 211 leerlingen over: zes leerlingen vertrokken van school, naast de vier leerlingen die al eerder van school waren veranderd. Bij 202 van deze leerlingen werd de LanguageScreen afgenomen en werd voor 185 leerlingen een leerlingprofiel ingevuld[5]. Van 118 leerlingen (59 sbo, 59 so) in de interventiegroep (87,3%) en van 172 leerlingen (80 sbo, 92 so) in de vergelijkingsgroep (81,5%) zijn gegevens van beide metingen beschikbaar van zowel de LanguageScreen als het leerlingprofiel.



Figuur 2.1 – Overzicht van aantallen leerlingen in de interventie- en vergelijkingsgroep en per meting verzamelde gegevens

Na de leerlingselectie door de scholen zijn de deelnemende scholen bezocht voorafgaand aan de start van de NELI interventie door een onderzoeker/testleider om de vier taaltaken af te nemen bij de geselecteerde leerlingen. In totaal zijn 21 van de 23 scholen bezocht om de taaltaken af te nemen. Op één school met drie deelnemende leerlingen werden geen taaltaken afgenomen.

5 Van twee klassen met in totaal 20 leerlingen ontbreken profielgegevens: de klas met 8 leerlingen waarvoor ook bij meting 1 geen profielen werden ingevuld en een klas waarvoor bij meting 1 wel, maar bij meting 2 geen leerlingprofielgegevens werden ingevuld.

Eind november 2024 zijn twee scholen bezocht (deze scholen zijn eerder gestart met de interventie); in januari 2025 zijn 18 scholen bezocht en in maart 2025 is één school bezocht (deze school startte wegens omstandigheden de interventie in maart). Deze 21 scholen waren of nog niet begonnen met de interventie, of de leerlingen hadden de eerste bijeenkomst gehad met de tutor. In totaal zijn bij 109 kinderen de 4 taaltaken afgenomen. Op één school werd bij één van de twee deelnemende groepjes kinderen de taaltaken afgenomen (betreft niet-deelname door zes kinderen). In totaal 10 deelnemende kinderen waren bij de testafname afwezig (meestal door ziekte).
Het duurde ongeveer 20 minuten per kind om het taalonderzoek af te nemen. Over het algemeen vonden kinderen het leuk om deze taakjes uit te voeren. Bij ongeveer 10% van de kinderen was een begeleider vanuit school aanwezig bij het taalonderzoek zodat de leerling zich op zijn/haar gemak voelde of om eventuele ernstige gedragsproblemen van leerlingen zoveel mogelijk te voorkomen.
De scholen werden bezocht door één van de vier testleiders. Deze testleiders hebben een gezamenlijke training gevolgd voorafgaand aan de taalonderzoeken in december 2024. De vier taken zijn in die training besproken en geoefend, net als het scoren van de deze taken. De instructies voor het afnemen van de taken zijn schriftelijk vastgelegd en alle vier de testleiders hebben een schriftelijk testprotocol gekregen. De pakketten met de materialen en scoreformulieren waren hetzelfde voor de vier testleiders.
Taken (3) en (4), waarbij de kinderen gevraagd werden om zelf iets te vertellen, zijn opgenomen (audio) en de audiofiles zijn bewaard in een veilige digitale omgeving (OneDrive van de UvA, beveiligd met een 2 staps authenticatie), zonder dat de namen van de leerlingen of de schoolnamen hieraan gekoppeld zijn. De audiofiles zijn meteen gewist van de opnameapparatuur zodra de opnames in de OneDrive opgeslagen waren. De data zijn anoniem opgeslagen. Op de scoreformulieren zijn alleen leerling ID nummers genoteerd. Namen van de leerlingen zijn niet schriftelijk vastgelegd tijdens de afname van de taalonderzoeken of daarna. Via de leerkracht op school werd voor de afname van het taalonderzoek de koppeling gelegd tussen naam leerling en leerlingnummer (mondeling). Er is een protocol opgesteld voor het transcriberen van de taaldata van taaltaken (3) en (4). In voorjaar 2025 zijn de transcripten onderzocht op de inter-rater betrouwbaarheid. De resultaten van de taken (1) en (2) zijn gescoord op basis van de handleiding van de taaltest. In mei en juni 2025 zijn de vier taaltaken weer afgenomen bij de dezelfde leerlingen (direct na afloop van de interventie), waarbij in de meeste gevallen de scholen bezocht werden door dezelfde testleiders die ook in januari de scholen bezocht hadden.

2.3 Kenmerken van voor NELI geselecteerde leerlingen

De toewijzing van leerlingen binnen de groep met lagere prestaties die in aanmerking kwamen voor deelname aan NELI vond uiteindelijk plaats op basis van het oordeel van de school. Om na te gaan in hoeverre leerlingen met specifieke kenmerken zijn geselecteerd voor de interventie, is gekeken of de achtergrondkenmerken van geselecteerde leerlingen verschillen van de achtergrondkenmerken van leerlingen die niet zijn geselecteerd. Dit levert niet alleen informatie op over mogelijke specifieke kenmerken van geselecteerde leerlingen, maar ook over welke achtergrondgegevens in elk geval meegenomen zouden moeten worden bij het samenstellen van de vergelijkingsgroep via matching. De matchingsprocedure is erop gericht om een vergelijkingsgroep samen te stellen die qua achtergrondkenmerken zo vergelijkbaar mogelijk is aan de interventiegroep. Tabel 3 bevat

de resultaten van de vergelijking. Wat taalvaardigheid betreft blijkt er geen verschil te zijn in de score op de LanguageScreen, maar -zonder controle voor andere kenmerken- wel in de inschatting van taalvaardigheid door de leerkracht. De leerlingen die voor de interventie zijn geselecteerd, zijn volgens het oordeel van de leerkracht wat minder taalvaardig dan de leerlingen die hier niet voor zijn geselecteerd. Na controle voor overige kenmerken verschilt dit niet significant. Voor leerlingen met een niet-Nederlandse thuistaal was de kans op interventiedeelname ruim tweemaal (2.043) zo hoog als voor leerlingen met een Nederlandse thuistaal en voor meisjes ruim anderhalf (1.686) maal zo hoog als voor jongens. De schoolsoort (sbo versus so), leeftijd, gedrag en ondersteuningsbehoeften verschilt niet tussen wel- en niet deelnemers aan de interventie.

Tabel 2.2 - Vergelijking tussen wel en niet voor NELI geselecteerde leerlingen (n=309; resultaten logistische regressie analyse met leerlingen die in aanmerking komen voor deelname)

	model 1		model 2		model 3		model 4	
	Exp(B)	p	Exp(B)	p	Exp(B)	p	Exp(B)	p
Standaardscore op de LanguageScreen	1.020	0.326					1.027	0.288
Inschatting taalvaardigheid door leerkracht	0.700	0.030					0.725	0.090
Leeftijd			1.022	0.805			0.999	0.988
Sbo (ref=so)			1.084	0.748			0.951	0.852
Niet-Nederlandse thuistaal (ref = geen)			2.043	0.005			2.226	0.006
Meisje (ref = jongen)			1.686	0.045			1.661	0.064
Ondersteuningsbehoefte:								
lichamelijk			1.179	0.562			1.071	0.816
cognitief			0.614	0.089			0.550	0.063
internaliserend probleemgedrag			0.959	0.876			0.844	0.568
externaliserend probleemgedrag			0.597	0.102			0.795	0.581
Werkhouding					0.846	0.362	0.742	0.153
Impulscontrole					1.222	0.348	1.425	0.130
Externaliserend probleemgedrag					1.104	0.653	1.206	0.437
Internaliserend probleemgedrag					1.065	0.731	1.200	0.395
Emotioneel reactief gedrag					0.866	0.420	0.865	0.479
Opkomen voor zichzelf					0.902	0.608	0.953	0.830
Rekening houden met anderen					1.174	0.420	1.174	0.466
Constante	0.288	0.389	0.514	0.294	0.393	0.490	0.044	0.132

2.4 Resultaten matching

Om een vergelijkingsgroep samen te stellen die qua achtergrondkenmerken zo vergelijkbaar mogelijk is aan de interventiegroep, is matching uitgevoerd.
De belangrijkste variabele bij de matching is de score op de LanguageScreen. Deze scores zijn gestandaardiseerd, zodat de resultaten van kinderen van verschillende leeftijden met elkaar vergeleken kunnen worden. Voor de matching is gewerkt met een indeling in drie groepen. Verder namen we onderwijssoort mee (sbo of so). Aangezien we binnen het so vanwege lagere aantallen in cluster 4 geen clusterindeling konden toevoegen, hebben we opgenomen of een leerling volgens de leerkracht een ondersteuningsbehoefte heeft met betrekking tot externaliserend probleemgedrag (ja/nee). Omdat geslacht en thuistaal samen bleken te hangen met selectie voor de interventie, hebben we die kenmerken ook bij de

matching betrokken (dichotome variabelen). Aangezien er bij enkele achtergrondkenmerken een aantal missende waarden waren (thuistaal 5, sekse 8, ondersteuningsbehoefte 10), zijn deze gegevens meervoudig geïmputeerd, met het pakket mice binnen R versie 4.4.2. Bij de imputatie werd rekening gehouden met de nesting van leerlingen binnen klassen. Er werden tien datasets aangemaakt. Vervolgens is de matching eveneens binnen R uitgevoerd met behulp van het pakket Matchthem op elk van deze tien datasets. Het type matching dat werd uitgevoerd, is coarsened exact matching, waarbij alleen observaties gematcht worden die exact in dezelfde categorieën vallen. Door middel van weging wordt gezorgd voor een evenwichtige verdeling van observaties in de analyse. Analyses worden vervolgens uitgevoerd op de tien datasets. Gepresenteerde resultaten betreffen gepoolde analyseresultaten over de tien datasets.
De imputatie- en matchingsprocedure is apart uitgevoerd voor de LanguageScreen en voor de gegevens verzameld met leerlingprofielen. Zoals genoemd is er in twee klassen voor gekozen door de leerkracht om geen leerlingprofielen in te vullen voor leerlingen in de vergelijkingsgroep, waardoor het aantal ingevulde leerlingprofielen wat lager is dan het aantal LanguageScreen-afnames. In de dataset voor de imputatie en matching werden alleen data van leerlingen opgenomen waarvoor gegevens beschikbaar waren van de tweede meting van respectievelijk de LanguageScreen en leerlingprofielen. Het aantal leerlingen in de dataset waarop de matching is uitgevoerd betreft voor de LanguageScreendata 119 leerlingen in de interventiegroep en 202 leerlingen in de vergelijkingsgroep. Voor de matching ten behoeve van de leerlingprofieldata is dit 119 leerlingen in de interventiegroep en 185 leerlingen in de vergelijkingsgroep. We presenteren hier de resultaten van de matching, uitgevoerd voor de LanguageScreen data. Resultaten van de matching ten behoeve van de leerlingprofieldata zijn vergelijkbaar en opgenomen in bijlage II. Bij de matching konden zeven leerlingen uit de interventiegroep niet worden gematcht met een vergelijkingsgroep-leerling, en 17 leerlingen uit de vergelijkingsgroep niet met een interventiegroep-leerling. Na weging is de effectieve steekproefgrootte (ESS) 112 leerlingen in de interventiegroep en 95 leerlingen in de vergelijkingsgroep. Door de matching zijn na weging de aandelen meisjes, leerlingen met een niet-Nederlandse thuistaal, leerlingen afkomstig uit sbo/so en met een ondersteuningsbehoefte op het gebied van externaliserend gedrag exact gelijk. De matching resulteerde dus in volledige balans op geslacht, thuistaal, schooltype (sbo/so) en ondersteuningsbehoefte met betrekking tot externaliserend gedrag.
Wat betreft de vergelijking tussen de interventie- en vergelijkingsgroep met betrekking tot de LanguageScreen score zijn de resultaten als volgt. Voorafgaand aan matching was het gemiddelde verschil in score minimaal en niet significant (-0.030), na matching nog steeds heel klein en niet significant (0.096).
De gemiddelde afstand tussen de gematchte paren op de LanguageScreen bedraagt 0.388, wat acceptabel is (vuistregel: waarden < 0.25 zeer goed, 0.25-0.50 acceptabel).
De spreiding blijft na matching vergelijkbaar tussen beide groepen: de variantie ratio vóór matching was 0.804 en na matching 0.820 (de waarde 1 is optimaal). De verdeling van LanguageScreen scores was voorafgaand aan de matching al vergelijkbaar (*empirical cumulative distribution function* (e*CDF*) gemiddeld 0.035 en maximaal 0.10), na matching nog iets meer vergelijkbaar (gem. 0.025 en maximaal 0.067). Waarden onder 0.1 worden doorgaans gezien als teken van goede balans. Een grafische weergave van de verdelingen is opgenomen in bijlage I, in de vorm van een dichtheidsplot voor de tien datasets.

2.5 Analysemethode

Vergelijking interventie- en vergelijkingsgroep
Met behulp van multilevel regressieanalyse met het pakket Lme4 binnen R zijn verschillen in uitkomstmaten tussen de interventie- en vergelijkingsgroep nagegaan. In multilevelanalyse wordt de clustering van leerlingen binnen klassen meegenomen in het model. De intraklassecorrelaties (ICC's) in de onderstaande tabel laten zien welk deel van de variantie per variabele en meting toe te schrijven is aan verschillen op klasniveau en op leerlingniveau. Omdat schooltype niet als extra niveau kon worden opgenomen, omvatten de ICC's ook variantie die toe te schrijven is aan verschillen tussen schoolsoorten. We tonen ICC's van de ongewogen en niet-geïmputeerde data en van gewogen en geïmputeerde data (gemiddelden over de 10 datasets). De ongewogen ICC's geven inzicht in de mate van clustering, terwijl de gewogen ICC's relevant zijn voor de effectschattingen. Uit de waarden blijkt klasgebonden variantie, wat bevestigt dat multilevel regressie-analyse de aangewezen methode is.

Tabel 2.3 - Overzicht van ongewogen en gewogen intraklassecorrelaties

	ongewogen		gewogen	
Gegevens LanguageScreen	***m1***	***m2***	***m1***	***m2***
LanguageScreen	0.098	0.191	0.083	0.164
Expressieve woordenschat	0.032	0.047	0.025	0.057
Receptieve woordenschat	0.047	0.089	0.101	0.116
Luisterbegrip	0.161	0.290	0.156	0.295
Zinnen herhalen	0.172	0.371	0.210	0.324
Gegevens leerlingprofiel				
Inschatting taalvaardigheid door leerkracht	0.187	0.101	0.149	0.216
Werkhouding	0.021	0.029	0.099	0.196
Impuls	0.097	0.067	0.160	0.091
Internaliserend	0.210	0.106	0.210	0.142
Externaliserend	0.030	0.053	0.204	0.156
Emotionele regulatie	0.065	0.087	0.120	0.246
Opkomen voor zichzelf	0.039	0.119	0.188	0.240
Rekening houden met anderen	0.132	0.135	0.188	0.193

Bij elke analyse was de uitkomstmaat bij de tweede meting de afhankelijke variabele en waren de uitkomstmaat bij de eerste meting en de indicator voor interventiedeelname de onafhankelijke variabelen. Om rekening te houden met clustering van leerlingen binnen klassen (n=37) werd een random intercept toegevoegd voor klasnummer. In de analyses zijn de bij de matching verkregen gewichten toegepast. De analyses zijn uitgevoerd op de tien datasets, waarvan de resultaten zijn gepoold om robuuste schattingen te verkrijgen. Vervolgens is met behulp van interactie-effecten nagegaan in hoeverre het verschil tussen de interventie- en vergelijkingsgroep anders is voor jonge en oudere leerlingen, sekse, thuistaal, de hoogte van de voormeting (score meting 1), de mate van externaliserend probleemgedrag en of sprake is van een ondersteuningsbehoefte op cognitief gebied (moderatie). Wat betreft taalvaardigheid is dit voor de score op de totale LanguageScreen

uitgevoerd. Bij de sociaal-emotionele uitkomstmaten is eveneens de interactie met de voormeting op de LanguageScreen nagegaan. De moderatieanalyses waren exploratief, er is daarom geen correctie toegepast voor multiple testing.

Mediatie
Met mediatie-analyse kan worden onderzocht hoe het effect van een interventie op een uitkomst tot stand komt. Dit biedt meer inzicht in werkzame mechanismen van de interventie (onderzoeksvraag 2). We onderzochten in hoeverre gedragsmaten het effect op taalvaardigheid mediëren en in hoeverre taalvaardigheid het effect op gedragsmaten medieert. In dit onderzoek is moderne mediatie-analyse (Pearl, 2009) toegepast. Moderne mediatie-analyse, nog weinig toegepast in onderwijskundig onderzoek, verdient de voorkeur boven traditionele mediatie-analyse, omdat dit robuustere causale analyse mogelijk maakt, doordat interacties en confounders systematisch kunnen worden gemodelleerd (Chi e.a., 2022; Whitlle e.a., 2017). Uitgegaan is van een dataset met 121 interventie- en 199 leerlingen in de vergelijkingsgroep. Dit betekent dat gegevens van de 12 vergelijkingsgroepleerlingen waarvoor zowel bij de voor- als nameting geen leerlingprofielgegevens werden ingevuld, niet zijn meegenomen. Na het imputeren van de (in beperkte mate aanwezige) missende gegevens waarbij rekening werd gehouden met de nesting van leerlingen binnen klassen, werd de matching met het pakket mice binnen R uitgevoerd op de tien geïmputeerde datasets (zie Bijlage III). Bij de matching konden zeven leerlingen uit de interventiegroep en 17 uit de vergelijkingsgroep niet één-op-één worden gekoppeld. De tien datasets zijn opgeslagen en ingelezen in Mplus 8.0, waarmee de mediatie-analyses zijn uitgevoerd met correctie voor de geneste structuur van leerlingen binnen klassen. Hierbij werd de volgende procedure toegepast. Eerst onderzochten we per gedragsmaat (werkhouding, impulscontrole, internaliserend en externaliserend probleemgedrag, emotioneel reactief gedrag, rekening houden met anderen en opkomen voor zichzelf) in hoeverre dit het effect van de interventie op de totale LanguageScreen en subtests medieert. Daarnaast onderzochten we in hoeverre effecten op de verschillende gedragsmaten werden gemedieerd door taalvaardigheid.
Steeds werden de scores op de nameting van een concept gecorrigeerd voor scores op de voormeting van hetzelfde concept.
Er zijn zowel totale als pure indirecte effecten berekend. Het *totale indirecte effect* geeft weer in welke mate de interventie via de mediator doorwerkt op de uitkomst, inclusief de mogelijkheid dat de interventie de relatie tussen mediator en uitkomst beïnvloedt. Het ***pure indirecte effect*** beschrijft hetzelfde pad, maar onder de aanname dat de relatie tussen mediator en uitkomst gelijk blijft in alle condities. Een voorbeeld: stel NELI (I) verhoogt de werkhouding (M) van kinderen, wat vervolgens zorgt voor een verhoging van hun taalvaardigheid (Y). Het pure indirecte effect geeft dan weer welk deel van het interventie-effect via werkhouding verloopt, bij een vaste relatie tussen werkhouding en taalvaardigheid. Het totale indirecte effect houdt daarnaast rekening met de mogelijkheid dat de interventie niet alleen de werkhouding verhoogt, maar ook de manier waarop werkhouding doorwerkt naar taalvaardigheid (dus ook de sterkte van de M–Y relatie beïnvloedt).
Daarbij is getoetst of er een significant verschil is tussen het pure en totale indirecte effect. Een verschil wijst er dus op dat de relatie tussen mediator en uitkomst in de interventiegroep anders is dan in de vergelijkingsgroep.

Hoe de uitvoering van NELI verliep
Effecten van een interventie zijn mede afhankelijk van de kwaliteit van de implementatie (Carroll e.a., 2007). Om inzicht te krijgen in zowel de uitvoering als de uitvoerbaarheid in het so en het sbo is een vragenlijst afgenomen onder de tutoren zijn groepsgesprekken gevoerd met medewerkers van tien scholen. Op basis van de uitkomsten van de vragenlijsten en de gesprekken beschrijven we hoe de interventie op de deelnemende scholen is uitgevoerd en ervaren. Daarbij gaan we in op de mate van interventietrouw, bevorderende en belemmerende factoren en de opbrengsten voor zowel de school als de leerlingen en maken we onderscheid tussen sbo en so.
Verschil NELI-deelnemers met vergelijkingsgroep bij optimalere uitvoeringskwaliteit
Om het effect van NELI bij optimale uitvoering te onderzoeken, is een Complier Average Causal Effect (CACE)-analyse uitgevoerd (o.a. Terza, Basu & Rathouz, 2008). Optimale uitvoering en deelname kunnen samenhangen met kenmerken die niet (volledig) gemeten zijn, zoals klasklimaat of didactische vaardigheden. Deze niet-gemeten kenmerken kunnen ook samenhangen met de nameting van taal. Er is dan sprake van endogeniteit. Als hier niet voor wordt gecontroleerd, kan dit zorgen voor een vertekening van het effect van optimale uitvoering.
Bij een CACE-analyse wordt dit voorkomen door het deel van optimale uitvoering dat samenhangt met niet gemeten kenmerken uit het effect te halen. Dit wordt gedaan door de oorspronkelijke toewijzing tot de interventie te gebruiken als hulpmiddel om verschillen in optimale uitvoering te verklaren. Toewijzing beïnvloedt namelijk wel de kans op optimale uitvoering, maar niet rechtstreeks de taaluitkomst.
De analyse is uitgevoerd met een twee-staps instrumentele variabele (IV) methode, een zogenaamde two-stage residiual inclusion (2SRI) analyse. In de **eerste stap** werd optimale uitvoering (afhankelijke variabele) met een logistische regressieanalyse voorspeld door toewijzing tot de interventie en de variabelen die ook werden gebruikt voor de matching: voormeting van taalvaardigheid, geslacht (meisje/jongen), thuistaal (Nederlands of (ook) een niet-Nederlandse taal), onderwijssoort (sbo of so) en of een leerling volgens de leerkracht een ondersteuningsbehoefte heeft met betrekking tot externaliserend probleemgedrag (ja/nee).
In de **tweede stap werd** met het pakket lme4 binnen R een lineair multilevelmodel met random intercepts per klas geschat. In deze analyse was de afhankelijke variabele de nameting van de LanguageScreen en werden als onafhankelijke variabelen de indicator voor optimale uitvoeringskwaliteit, het residu uit de eerste stap om te corrigeren voor endogeniteit en dezelfde covariaten als in stap 1 opgenomen.

Taalontwikkeling interventieleerlingen: de aanvullende taaltaken
De resultaten van het aanvullend afgenomen taalonderzoek bij de interventieleerlingen bieden rijkere informatie over het verloop van hun ontwikkeling. Wat betreft actieve woordenschat is de CELF-meting uitgebreider dan de meting van dit concept (expressieve woordenschat) binnen de LanguageScreen. Ook de CELF-subtest Zinnen Begrijpen is uitgebreider dan de subschaal Luisterbegrip van de LanguageScreen, terwijl beide receptief zinsbegrip betreffen. Daarnaast zijn met de Action Picture Test en het Busverhaal aanvullende concepten gemeten die niet in de LanguageScreen zijn opgenomen, namelijk narratieve spontane taalproductie en grammaticale complexiteit in vrije taal. Van een deel van de kinderen viel de leeftijd buiten het bereik van de normtabellen, waardoor de ruwe scores van deze kinderen niet worden omgezet naar gestandaardiseerde scores. Daarom zijn we in de hier beschreven analyses (ook bij de LanguageScreen) uitgegaan van ruwe scores.

In totaal zijn van 101 interventiekinderen toetsgegevens beschikbaar. Soms kon geen score worden vastgesteld, omdat er onvoldoende taaluitingen beschikbaar waren om betrouwbaar een score toe te kunnen kennen. Daarnaast ontbraken bij de nameting gegevens van twee van de 101 kinderen. In het volgende overzicht staan de aantallen beschikbare scores per taaltaak en meting.

	voormeting	nameting
Actieve Woordenschat	99	94
APT semantiek	87	87
APT grammatica	94	90
Zinnen Begrijpen	98	96
Busverhaal	86	81

Geen van de scores op de LanguageScreen miste voor de 101 kinderen. Missende gegevens op de taaltaken zijn tienvoudig geïmputeerd met het pakket mice binnen R 4.4.2 waarbij rekening werd gehouden met de nesting van leerlingen binnen klassen. De scores zijn gestandaardiseerd waarbij de voormeting als referentiepunt is gebruikt. Om de ontwikkelingen te beschrijven, zijn per taaltaak multilevelanalyses uitgevoerd met drie niveaus (metingen genest binnen leerlingen, genest binnen klassen), met de nameting als afhankelijke variabele en de voormeting als covariaat. Gemiddeld bedroeg de tijd tussen de voor- en nameting 0.454 jaar (std. 0.047). Hoewel de variantie in de tijd tussen beide metingen beperkt was, hebben we dit als covariaat opgenomen om te onderzoeken of het tot vertekening van resultaten zou kunnen leiden. Dit bleek niet het geval: de tijd tussen de metingen hing niet significant samen met de ontwikkeling in de taaltaken en is daarom in de verdere analyses niet meegenomen.
Om te onderzoeken of taalniveau en -groei verschilden naar sekse, thuistaal, onderwijssoort (sbo/so), leeftijd en externaliserend probleemgedrag (bij de voormeting), zijn deze variabelen als covariaten en in interactie met tijd aan het model toegevoegd. Daarnaast is de relatie met betrokkenheid bij en aanwezigheid tijdens de interventie gerelateerd aan de ontwikkeling in een model met controle voor de achtergrondkenmerken leeftijd, sekse en onderwijssoort (sbo/so).
Alle analyses zijn per geïmputeerde dataset uitgevoerd; de gepresenteerde resultaten betreffen de gepoolde uitkomsten van deze tien analyses.

3 Resultaten effectiviteit interventie

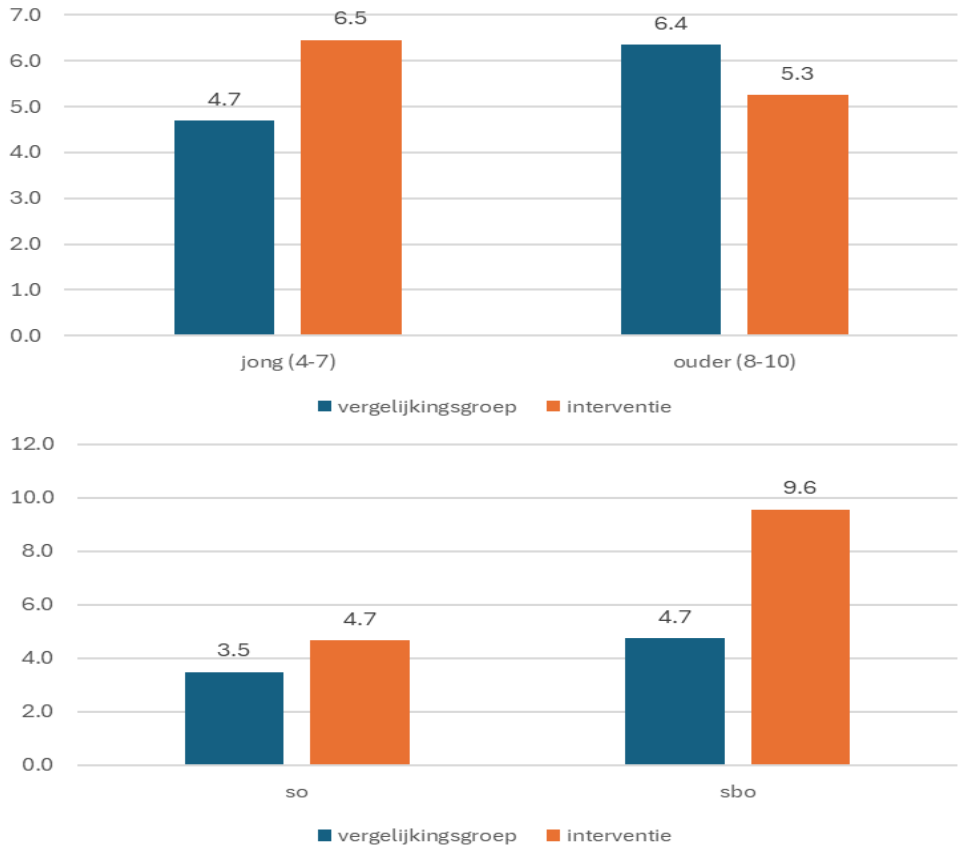
3.1 Taalvaardigheid interventie- versus vergelijkingsgroep

Leerlingen die aan NELI deelnamen, zijn meer vooruitgegaan in taalvaardigheid dan leerlingen die daar niet aan deelnamen. Het verschil is ruim drie punten op de gestandaardiseerde LanguageScreen met een effectgrootte van 0.338. Kijken we naar de subtestscores, dan is te zien dat het verschil het grootst is voor luisterbegrip (effectgrootte 0.500) en dat ook zinsherhaling (effectgrootte 0.282) en expressieve woordenschat (effectgrootte 0.186) meer zijn toegenomen in de interventie- dan in de vergelijkingsgroep. De toename in receptieve woordenschat verschilde niet tussen de interventie- en vergelijkingsgroep. De verandering in de inschatting van de taalvaardigheid door de groepsleerkracht verschilt niet tussen de interventie en controlegroep.

Tabel 3.1 - Resultaten taalvaardigheid: interventie- versus vergelijkingsgroep (ongestandaardiseerde resultaten; voor het interventie-effect ook gestandaardiseerde resultaten met 95% betrouwbaarheidsinterval)

					beta met 95% betrouwbaarheidsinterval			
(Sub)schaal		**estimate**	**std.error**	**p**	**beta**	**(laag**	**-**	**hoog)**
LanguageScreen (totaal)	intercept	2.749	(4.234)	0.517				
	voormeting	1.013	(0.053)	0.000				
	interventie	3.096	(0.740)	**0.000**	**0.338**	(0.180	-	0.496)
Expressieve woordenschat	intercept	11.917	(3.649)	0.001				
	voormeting	0.887	(0.044)	0.000				
	interventie	1.903	(0.868)	**0.029**	**0.186**	(0.020	-	0.353)
Receptieve woordenschat	intercept	35.111	(5.136)	0.000				
	voormeting	0.649	(0.057)	0.000				
	interventie	2.427	(1.353)	0.074	0.183	(-0.017	-	0.383)
Luisterbegrip	intercept	32.114	(5.266)	0.000				
	voormeting	0.636	(0.065)	0.000				
	interventie	5.763	(1.167)	**0.000**	**0.500**	(0.302	-	0.698)
Zinnen herhalen	intercept	15.915	(4.972)	0.002				
	voormeting	0.839	(0.060)	0.000				
	interventie	3.009	(0.997)	**0.003**	**0.282**	(0.103	-	0.462)
Inschatting taalvaardigheid	intercept	0.816	(0.132)	0.000				
	voormeting	0.756	(0.047)	0.000				
	interventie	0.004	(0.079)	0.957	0.005	(-0.167	-	0.176)

Moderatie: verschillen in effecten naar achtergrondkenmerken?
In eerdere studies naar NELI werden weinig verschillen in effectiviteit gevonden naar kenmerken van leerlingen. Wel bleek in de Engelse effectstudie (West e.a., 2001) dat NELI effectiever was naarmate leerlingen bij de voormeting een zwakkere taalvaardigheid hadden. In het Nederlandse effectonderzoek in het reguliere basisonderwijs werd dit niet gevonden.
Net als in de reguliere studie zijn we nagegaan in hoeverre er verschillen zijn in effectiviteit van NELI op de voormeting van taalvaardigheid, thuistaal en sekse. Daarnaast zijn verschillen met betrekking tot kenmerken onderzocht die specifiek relevant zijn voor de uitvoering van NELI in het gespecialiseerd onderwijs: een leeftijd van 8-10 jaar (ja, nee jonger), onderwijssoort (sbo/so), het hebben van een ondersteuningsbehoefte op cognitief gebied en op basis van de score op de schaal externaliserend probleemgedrag bij de voormeting. Net als binnen de Nederlandse effectstudie vonden we geen significante interacties met taalvaardigheid bij de voormeting, geslacht en thuistaal. Evenmin vonden we verschillen naar mate van effectiviteit en externaliserend probleemgedrag en of sprake is van ondersteuningsbehoeften op cognitief terrein. Wel vonden we dat de effectiviteit van NELI verschilt naar leeftijd en onderwijssoort. De interactie tussen interventie en oudere leeftijdsgroep was significant, $b = -2.85$, $SE = 1.36$, $t(288) = -2.09$, $p = .037$, evenals de interactie tussen interventie en onderwijssoort, $b = 3.63$, $SE = 1.47$, $t(274) = 2.46$, $p = .014$ (zie Figuur 3.1). Voor jongere leerlingen in de interventiegroep ontwikkelde de taalvaardigheid zich gunstiger dan voor leerlingen in de vergelijkingsgroep, terwijl oudere leerlingen in de interventiegroep juist lager scoorden dan oudere leerlingen in de controlegroep. Dit patroon suggereert dat NELI in de huidige vorm minder geschikt is voor de oudere leeftijdsgroep. Daarnaast laat Figuur 2 zien dat in het sbo het verschil tussen wel en niet interventiedeelname duidelijk groter is dan in het so: in het so nam de taalvaardigheid in de interventiegroep met 1.2 punten (4.7 - 3.5) meer toe dan in de vergelijkingsgroep, terwijl dit verschil in het sbo 4.9 punten (9.6 - 4.7) bedroeg. We merken op dat onderwijssoort en leeftijd deels overlappen: het aandeel oudere kinderen was in het so hoger (21.1%) dan in het sbo (9.9%).



Figuur 3.1 – Geschatte ontwikkeling in de totaalscore op de LanguageScreen bij gemiddelde score op de voormeting voor de interventie- en controlegroep naar leeftijdsgroep (4-7 en 8-10) en onderwijssoort (sbo en so)

3.2 Gedrag interventie- versus vergelijkingsgroep

Deelname aan NELI hangt samen met een gunstigere ontwikkeling in de mate waarin leerlingen voor zichzelf opkomen (effectgrootte 0.257) en rekening houden met anderen (effectgrootte 0.195), zoals beoordeeld door leerkrachten. In leerkrachtoordelen over werkhouding, impulscontrole, internaliserend en externaliserend probleemgedrag en emotioneel reactief gedrag vonden we geen verschil.

Tabel 3.2 - Resultaten gedrag: interventie- versus vergelijkingsgroep (ongestandaardiseerde resultaten; voor het interventie-effect ook gestandaardiseerde resultaten met 95% betrouwbaarheidsinterval)

					beta met 95% betrouwbaarheidsinterval			
Schaal		**estimate**	**std.error**	**p**	**beta**	**(laag**	**-**	**hoog)**
Werkhouding	intercept	1.125	(0.146)	0.000				
	voormeting	0.622	(0.045)	0.000				
	interventie	0.084	(0.080)	0.299	0.099	(-0.087	-	0.285)
Impulscontrole	intercept	1.453	(0.155)	0.000				
	voormeting	0.551	(0.046)	0.000				
	interventie	0.031	(0.081)	0.700	0.040	(-0.163	-	0.242)
Internaliserend probleemgedrag	intercept	1.193	(0.128)	0.000				
	voormeting	0.514	(0.044)	0.000				
	interventie	-0.062	(0.063)	0.330	-0.099	(-0.299	-	0.100)
Externaliserend probleemgedrag	intercept	0.619	(0.127)	0.000				
	voormeting	0.736	(0.038)	0.000				
	interventie	0.002	(0.075)	0.982	0.002	(-0.159	-	0.163)
Emotioneel reactief gedrag	intercept	0.769	(0.118)	0.000				
	voormeting	0.667	(0.037)	0.000				
	interventie	-0.109	(0.077)	0.156	-0.118	(-0.280	-	0.045)
Opkomen voor zichzelf	intercept	1.442	(0.162)	0.000				
	voormeting	0.542	(0.046)	0.000				
	interventie	0.190	(0.072)	**0.009**	0.257	(0.066	-	0.447)
Rekening houden met anderen	intercept	1.224	(0.149)	0.000				
	voormeting	0.609	(0.044)	0.000				
	interventie	0.144	(0.071)	**0.043**	0.195	(0.007	-	0.382)

Moderatie: verschillen in gedragseffecten naar achtergrondkenmerken?
Interventiedeelname bleek niet systematisch sterker of zwakker samen te hangen met gedragsontwikkeling voor specifieke groepen leerlingen of leerlingkenmerken. We gingen dit na voor de volgende kenmerken: leeftijd 4-7 of 8-10 jaar, sekse, thuistaal, de hoogte van de voormeting (score meting 1), de score op de LanguageScreen (meting 1), de mate van externaliserend probleemgedrag en of sprake is van een ondersteuningsbehoefte op cognitief gebied. Voor werkhouding, impulscontrole, internaliserend probleemgedrag, opkomen voor zichzelf en rekening houden met anderen werden geen significante interacties gevonden. Twee interacties bleken significant. Voor de volledigheid beschrijven we deze hieronder. Voor het leerkrachtoordeel van emotioneel reactief gedrag was er een significante interactie met leeftijdsgroep (b = −0.36, SE = 0.14, $t(272)$ = −2.58, p = .010): bij oudere leerlingen hing interventiedeelname samen met een afname van emotioneel reactief gedrag; bij jongere leerlingen was dit nauwelijks het geval.
Daarnaast was er significante interactie tussen interventie en de (gecentreerde) LanguageScreen-score (b = −0.03, SE = 0.02, $t(270)$ = −1.98, p = .048). Dit betekent dat het interventie-effect op externaliserend probleemgedrag afneemt naarmate de LanguageScreen-score bij de voormeting hoger is: voor leerlingen met relatief hoge LanguageScreen-scores op de voormeting hangt deelname aan de interventie samen met iets minder externaliserend probleemgedrag dan in de vergelijkingsgroep, terwijl dit verschil bij leerlingen met lagere LanguageScreen-scores minimaal is of ontbreekt.

4 Resultaten mediatieanalyse

Om meer inzicht te krijgen in de mogelijke werkzame mechanismen van NELI (onderzoeksvraag 2) voerden we mediatie-analyses uit. Daarbij onderzochten we in hoeverre gedragsmaten het effect van de interventie op taalvaardigheid (totaalscore LanguageScreen) mediëren, en in hoeverre taalvaardigheid het effect van de interventie op gedragsmaten medieert. De totale indirecte effecten zijn weergegeven in Tabel 4.1. Geen van deze effecten bleek significant. Dit betekent dat er geen aanwijzingen zijn dat gedragsverandering een verklarende rol speelt in de positieve effecten van NELI op taalvaardigheid, noch dat verbeterde taalvaardigheid gedragsverbeteringen binnen de interventie verklaart. De effecten op taal en gedrag lijken daarmee onafhankelijk van elkaar te bestaan en parallel op te treden. Hoewel de indirecte effecten niet significant zijn, wijzen de effectgroottes er eerder op dat verbeterde taalvaardigheid samenhangt met gedragsverbetering dan dat gedragsverandering een verklaring vormt voor de effectiviteit van NELI op taalvaardigheid. Deze patronen zijn echter niet sterk genoeg om als bewijs voor mediatie te worden geïnterpreteerd.

Tabel 4.1 - Totale indirecte effecten van de mediatie-analyses

	Gedrag als mediator			LanguageScreen als mediator		
	estimate	s.e.	p	estimate	s.e.	p
Werkhouding	0.024	0.028	0.383	0.044	0.029	0.135
Impuls	0.005	0.020	0.800	0.052	0.028	0.062
Internaliserend	0.023	0.020	0.257	-0.052	0.030	0.083
Externaliserend	0.008	0.013	0.527	-0.048	0.027	0.076
Emotionele regulatie	0.028	0.028	0.316	-0.041	0.028	0.148
Opkomen voor zichzelf	0.025	0.016	0.122	0.010	0.019	0.615
Rekening houden met anderen	0.028	0.021	0.193	0.026	0.023	0.264

5 Beschrijvende resultaten over de uitvoering van NELI

In dit hoofdstuk beschrijven we de bevindingen over de uitvoering en uitvoerbaarheid van NELI gebaseerd op vragenlijsten onder tutoren en groepsgesprekken met medewerkers van tien scholen.

5.1 Achtergrond van de tutoren

Op alle 23 deelnemende scholen vulden één of meer tutoren een vragenlijst in (gemiddeld twee, minimum één, maximum vijf). De meeste tutoren (69%) zijn onderwijsassistent en de meesten (58%) werken minimaal vier jaar op de school. Er is geen verschil tussen so en sbo in de functie van de tutor en het aantal jaren dat de tutor werkzaam is op de school.
De groepsgesprekken werden uitgevoerd op tien scholen, waarvan vier sbo-scholen, twee so scholen cluster 4, drie so-scholen cluster 3 en een combinatieschool cluster 3 en 4. De groepssamenstelling en het aantal deelnemers varieerden per school. Veelal namen meerdere tutoren (doorgaans twee) en meerdere groepsleerkrachten deel, naast een NELI-coördinator. In totaal waren 43 professionals betrokken bij de tien groepsgesprekken.

5.2 Training

Voorafgaand aan de uitvoering wordt een online training aangeboden. Deze bestaat uit twee onderdelen. In deel 1 staan de basisprincipes van taalverwerving bij jonge kinderen centraal. Deel 2 bereidt voor op de uitvoering. Elk deel duurt ongeveer vier uur. Beide onderdelen kunnen zelfstandig, op een zelfgekozen moment worden doorlopen. Een tutor moet beide onderdelen van de training hebben afgemaakt voordat met NELI kan worden gestart. Op basis hiervan krijgt de tutor een certificaat als bevoegd NELI- tutor. De ontwikkelaar raadt aan dat zoveel mogelijk betrokkenen op de deelnemende school deel 1 van de training volgen. Deel 2 is specifiek voor de tutoren. Uit de antwoorden op de vragenlijst komt naar voren dat de tutoren over het algemeen een halve tot een hele dag aan de training hebben besteed. Uit de groepsgesprekken blijkt dat vaak ook andere betrokkenen dan de tutoren (bijvoorbeeld leerkrachten) beide onderdelen van de training volgden. In sommige gevallen volgde bijna de hele onderbouw de training, onder meer om een back-up te hebben bij ziekte of uitval van een tutor.
Voor verreweg de meeste tutoren (91%) gaf de training naar eigen zeggen voldoende (60%) of ruim voldoende (31%) voorbereiding op de uitvoering van het programma. Uit de gesprekken blijkt dat ook het voorgesprek met Vinci bijdraagt aan een goede voorbereiding. Hoewel de start soms onwennig was en er in het begin vragen waren, werd de structuur

gaandeweg duidelijker en makkelijker te begrijpen. Een goede voorbereiding, zoals het ordenen en knippen van materialen per sessie, bleek essentieel voor een soepele uitvoering, al kostte dit wel tijd.
Sommige tutoren waren nog bezig met de training terwijl de uitvoering al startte, waardoor het begin wat rommelig verliep. Sommigen ondervonden problemen met online toegang, maar konden zich inwerken via documenten. Het online materiaal werd door meerdere tutoren als prettig ervaren, omdat dat het werken in eigen tempo mogelijk maakte. Over het algemeen kwamen teams er onderling goed uit, soms met ondersteuning van Vinci.

5.3 LanguageScreen

De LanguageScreen is een digitale tool om op een snelle en betrouwbare manier de mondelinge taalvaardigheid in beeld te brengen en te monitoren. Zoals in hoofdstuk 2 genoemd, wordt dit instrument als onderdeel van NELI ingezet om het taalniveau van de leerlingen vast te stellen en leerlingen te selecteren voor deelname. Afname gebeurt door de leerkracht/tutor op school. De afname van de LanguageScreen werd door bijna tweederde (62%) van de tutoren als (heel) gemakkelijk ervaren, de meeste overige tutoren vonden het noch lastig noch gemakkelijk. Het voorbeeldfilmpje over de afname heeft hierbij geholpen. Ook wat betreft tijdsduur wordt de afname als goed uitvoerbaar ervaren, al vraagt afname bij een groot aantal leerlingen aanzienlijke inzet, zo blijkt uit de groepsgesprekken. De uitkomsten werden door verreweg de meeste tutoren (89%) informatief (83%) of zeer informatief (6%) gevonden. Er zijn geen verschillen tussen sbo en so in ervaringen met de LanguageScreen.

5.4 Organisatie van de begeleiding

In zowel so als sbo begeleidden verreweg de meeste tutoren (87%) één groepje leerlingen, de overige tutoren begeleidden twee groepjes. In het sbo verzorgden de meeste tutoren (53%) de begeleiding alleen, in het so meestal samen (79%). De meest voorkomende vorm was dat de tutoren de begeleiding over verschillende dagen verdeelden, soms werden groepjes samen begeleid en soms nam de ene tutor de groepsbegeleiding op zich terwijl de ander zich richtte op individuele begeleiding. Sommige scholen hebben specifieke oplossingen gevonden, zoals tutoren die volledig voor NELI werden ingeroosterd of een tutor die alle voorbereiding en coördinatie op zich nam, wat door de andere teamleden als zeer prettig werd ervaren. De coördinatie van het programma werd vaak door één persoon gedaan, die fungeerde als 'spin in het web'.

5.5 Voortgang

NELI omvat 20 weken, waarin de leerlingen wekelijks drie keer per week een groepsactiviteit van 30 minuten per keer aangeboden krijgen, in groepjes van ongeveer vier leerlingen. Daarnaast krijgt elke leerling twee individuele sessies van 15 minuten. De sessies hebben een vaste structuur.
De uitvoering van het programma was zo gepland dat de tutoren, bij een start van het programma in januari, het programma aan het eind van het schooljaar zouden hebben doorlopen. In de praktijk is dit echter op de meeste scholen niet haalbaar gebleken. Op het moment dat de vragenlijst werd afgenomen, verwachtte twee-derde (67%) van de tutoren

het programma nog niet volledig afgerond te hebben; gemiddeld dachten zij tot vier vijfde (84%) van het programma te zullen komen (minimum 53%, maximum 96%). Er zijn geen verschillen tussen sbo en so in de mate van voortgang. Als toelichting wordt in zowel so als sbo een latere start genoemd, vanwege omstandigheden, vrije dagen en vakanties. Binnen het so wordt ook een tekort aan personeel als reden genoemd: de tutor was bijvoorbeeld nodig om andere lessen te geven waardoor sessies niet door konden gaan.

5.6 Groeps- en individuele sessies

Groepssessies: aantal en duur
Op verreweg de meeste scholen werden, zoals voorgeschreven, drie groepssessies per week gegeven, op twee scholen (één so en één sbo) lukte dit niet. Als reden werd 'te weinig beschikbare tijd' gegeven. Ruim een derde (38%) van de tutoren gaf aan enkele groepssessies herhaald te hebben, dus vaker dan één keer uit te voeren. Dit verschilde niet tussen so en sbo. Het vaakst werd hierbij genoemd dat de woorden nog onvoldoende beklijfden bij de kinderen. Ook werd genoemd dat herhaling goed is voor de kinderen. Een reden die alleen binnen het so werd gegeven is dat door een te korte spanningsboog en problematisch gedrag herhaling nodig bleek: *'Wat je in een NELI les moet behandelen is heel veel en 30 minuten groepssessie duurt voor onze leerlingen te lang. De spanningsboog op deze school ligt veel lager dan in het reguliere onderwijs',* aldus een tutor in het so, cluster 4.
De meerderheid van de tutoren besteedde gemiddeld een half uur aan een groepssessie (so 60%; sbo 88%), zoals voorgeschreven. In het sbo besteedde de overige 12% gemiddeld minder dan een half uur aan een sessie. Binnen het so besteedde eveneens 12% van de tutoren hier minder dan een half uur aan, terwijl ruim een kwart (28%) hier gemiddeld meer dan een half uur aan besteedde, namelijk rond de 40 minuten. In het so werd dus gemiddeld meer tijd aan de groepssessies besteed dan in het sbo.

Individuele sessies: aantal en duur
De meeste tutoren (83%) voerden, zoals voorgeschreven, voor elke leerling twee individuele sessies per week uit. Waar werd afgeweken, werden vrijwel altijd minder dan twee individuele sessies uitgevoerd. Genoemde redenen waren onderbezetting en ziekte van leerlingen.

Groeps- en individuele sessies: kwaliteit uitvoering per onderdeel
Alle sessies bestaan uit een vast aantal terugkerende onderdelen. Voor zowel de groeps- als individuele sessies vroegen we de tutoren om aan te geven in hoeverre het lukt om deze onderdelen aan bod te laten komen. Antwoordcategorieën waren 1 lukt helemaal niet, 2 lukt niet, 3 lukt soms wel/soms niet, 4 lukt wel, 5 lukt heel goed. Tabel 5.1 toont per onderdeel apart voor so en sbo de gemiddelde score en het percentage van de tutoren dat aangeeft dat het (helemaal) niet lukt (% 1+2) en het percentage dat aangeeft dat het (heel) goed lukt (% 4+5). Over het geheel genomen lukte het de tutoren goed om de meeste onderdelen aan bod te laten komen. Een uitzondering vormt het onderdeel luisterspel. De NELI handleiding bevat een 'Luisterspelletjesbibliotheek', met suggesties voor spelletjes waarmee de aandacht van de kinderen wordt gericht op het luisteren. Van alle onderdelen lukt de uitvoering van het luisterspel het minst goed in zowel sbo als so en individuele als groepssessies. In de groepssessies geeft 21% van de tutoren aan dat het helemaal niet goed of niet goed lukt en 54% dat het goed of heel goed lukt, in het sbo geeft 6% van de tutoren

aan dat het (helemaal) niet goed lukt en 29% dat het (heel) goed lukt. In de individuele sessies is het patroon vergelijkbaar. Een tutor in het so formuleert: *'In het luisterspel gaat veel tijd zitten, wat ten koste gaat van de concentratie, die nodig is voor andere activiteiten'.* Een andere geeft aan: *'Het luisterspel heb ik vaak overgeslagen, omdat ik maar beperkt tijd had om de sessie te geven. Als ik het luisterspel wel toevoegde, duurde de sessie te lang'*. Het is ook een kwestie van prioritering: *'Ik kies vaak voor 'woordenschat' en 'narratief' en dan laat ik het luisterspel achterwege'*, aldus een tutor in het sbo.

In elke sessie worden nieuwe woorden geïntroduceerd, of nieuwe woorden uit vorige sessies herhaald. Het onderdeel 'Introductie en woordenschatinstructie en -herhaling' lukt het beste. Op één onderdeel zien we een verschil tussen het sbo en het so: in het so lukt het narratieve onderdeel minder goed dan in het sbo. Dit deel van de sessie is gericht op het ontwikkelen van narratieve vaardigheden, die nodig zijn om een begrijpelijk verhaal te vertellen. Aan de hand van zogenoemde vraagwoorden (wie, wat, waar, waarom) worden kinderen vertrouwd gemaakt met de belangrijkste bouwstenen van een verhaal. Deze vragen vormen een houvast bij het begrijpen en structureren van narratieven. In de praktijk blijkt dat dit onderdeel echter niet altijd wordt uitgevoerd. Als reden wordt onder meer aangevoerd dat het cognitieve niveau van de kinderen hiervoor ontoereikend zou zijn. In plaats daarvan verschuift de nadruk naar herhaling en consolidatie van speciale woorden. Enkele citaten ter illustratie: *'Het narratief is het lastigste onderdeel van de les. Bij ons hebben de leerlingen te weinig kennis en mogelijkheden om dit goed, of enigszins te kunnen uitvoeren. Vereenvoudigen werkt hier ook niet'* (tutor so). *'Een heel lastig onderdeel is het narratief. Dat is echt heel moeilijk bij onze doelgroep. Het verhaal dat je moet voorlezen sluit vaak minder of niet aan bij de belevingswereld van de kinderen. En het taalgebruik is heel moeilijk'* (tutor cluster 3). In het sbo wordt overigens ook genoemd dat het niet goed lukt om het narratieve onderdeel voldoende goed te laten slagen in de groepssessie.

Het onderdeel 'plenair' is bedoeld om samen aan het einde van een sessie terug te blikken op de sessie en te herhalen wat de leerlingen aangeboden hebben gekregen, zodat de opgedane kennis wordt verdiept en bestendigd. Zowel binnen het so als het sbo blijkt dat dit onderdeel in de praktijk niet altijd voldoende kan worden uitgevoerd. Vaak ontbreekt het aan tijd, of is de spanningsboog van de kinderen aan het einde van de sessie onvoldoende om de gezamenlijke afronding met voldoende aandacht te doorlopen.

Tabel 5.1 - Oordeel tutoren over uitvoerbaarheid interventieonderdelen binnen de individuele en groepssessies apart voor so en sbo: de gemiddelde score, het aandeel van de tutoren dat aangeeft dat het onderdeel (helemaal) niet lukt (% 1+2) en het aandeel dat aangeeft dat het (heel) goed lukt (% 4+5)

	So (n=24)			Sbo (n=17)		
	gem.	% (helemaal) niet goed	% (heel) goed	gem.	% (helemaal) niet goed	% (heel) goed
Groepssessies						
Introductie	4.2	4	88	4.5	0	94
Luisterspel (luistervaardigheden; luisterregels)	3.5	21	54	3.3	6	29
Instructie woordenschat	4.2	0	92	4.5	0	100
'Speciale woorden', 'overlearning'[6]	4.0	0	71	4.2	0	88
Narratief*	3.5	8	46	4.4	6	88
Plenair	3.8	4	63	3.6	0	53
Individuele sessies						
Luisterspel (luistervaardigheden, luisterregels)	3.3	29	54	3.2	6	24
Introductie	4.3	0	92	4.2	6	82
Woordenschatherhaling (zelfstandige naamwoorden uit de groepssessie herhalen, aan de hand van Flashcards)	4.4	0	92	4.5	0	94
Narratief*	3.8	8	67	4.6	0	100

**significant verschil tussen so en sbo*

5.7 Registraties en screening

Het is de bedoeling dat tutoren registraties bijhouden van de groeps- en individuele sessies. Aantekeningen uit de groepssessies worden gebruikt in de individuele sessies om deze goed aan te laten sluiten. Het maken van aantekeningen op het registratieblad na de groepssessies en het bijhouden van administratie van de individuele sessies lukte de tutoren over het algemeen goed (zowel bij de groeps- als individuele sessies 81% (heel) goed). Wanneer het niet of minder goed lukte kwam dit vooral door tijdgebrek, bijvoorbeeld omdat men snel weer door moest naar een volgende activiteit: *'Ik noteer alleen iets als een leerling afwezig is of als mij iets speciaals opvalt. Ik ben na de sessies ook weer hard nodig in de lessen'*, aldus een tutor in het sbo.
Daarnaast zijn er screeningsformulieren om de vooruitgang van de kinderen op het gebied van narratieve spraak en grammatica, en gedrag en houding bij te houden. Het lukte de tutoren over het algemeen goed om ook deze formulieren in te vullen (85% (heel) goed). Er zijn geen verschillen tussen so en sbo in de mate waarin de genoemde registraties en screening slaagden.

6 Overlearning: het blijven oefenen en herhaald gebruik~~en~~ van woorden of begrippen na~~dat~~ initiële beheersing ~~is bereikt~~. Door deze herhaling worden de woorden in het langetermijngeheugen verankerd, wat leidt tot betere retentie, snellere recall en automatische uitvoering.

5.8 Didactische vaardigheden

Aan de tutoren werd gevraagd in hoeverre het hen lukt om tijdens de tutorlessen te differentiëren, woordenschatdidactiek toe te passen en om narratieve vaardigheden van kinderen te ondersteunen. In Tabel 5.2 staan de resultaten apart voor het so en sbo. Over het algemeen lukt het de tutoren goed om de didactische vaardigheden toe te passen. Er is één (significant) verschil tussen so en sbo: differentiëren, het aanpassen van activiteiten aan het niveau, interesses en behoeften van het kind, lukte de tutoren in het so wat minder goed dan in het sbo. Een so-tutor formuleert: *'Voor onze groep is herhalen waardevoller dan differentiëren.'* En een andere tutor: *'Ik verlaag alle activiteiten al, omdat het sowieso te lastig is.'* Ook wordt differentiatie tijdrovend genoemd: *'Vaak bereid je kort op de uitvoering voor, vanwege de intensiteit van het programma. Dan is er te weinig tijd voor differentiatie'*, aldus een so-tutor. Uit de groepsgesprekken blijkt dat met name logopedisten de sessies afstemmen op de specifieke onderwijs- en ondersteuningsbehoeften van de groep waarmee zij werken. Sommigen vertellen actief op zoek te gaan naar de onderliggende doelen van een sessie en daar op eigen wijze invulling aan te geven, vaak minder strikt dan het programma voorschrijft. Deze werkwijze sluit aan bij hun professionele praktijk, waarin het gebruikelijk is om creatief en flexibel met lesmateriaal om te gaan en interventies aan te passen aan de mogelijkheden van de leerlingen, in het bijzonder binnen cluster 3. Zo kiezen zij er naar eigen zeggen bijvoorbeeld eerder voor om één activiteit verdiepend en intensief uit te voeren, dan om meerdere activiteiten meer oppervlakkig te behandelen.

Tabel 5.2 – De mate waarin het tutoren lukt om drie didactische vaardigheden toe te passen, apart voor so en sbo: de gemiddelde score en het aandeel tutoren dat aangeeft dat dit (helemaal) niet lukt (% 1+2) en het aandeel dat aangeeft dat het (heel) goed lukt (% 4+5)

	So (n=24)			Sbo (n=17)		
	gem.	% (helemaal) niet goed	% (heel) goed	gem.	% (helemaal) niet goed	% (heel) goed
differentiëren (activiteiten aanpassen aan het niveau, interesses en behoeften van het kind)*	3.9	4.2	70.8	4.6	0.0	100.0
toepassing woordenschatdidactiek	4.0	0.0	83.3	4.4	0.0	94.1
ondersteunen van narratieve vaardigheden	3.8	0.0	62.5	4.2	0.0	88.2

**significant verschil tussen so en sbo*

5.9 Ervaren resultaten

De meerderheid van de tutoren gaf desgevraagd aan enigszins tot duidelijk resultaten van NELI te zien (Tabel 5.3). Wat betreft productieve woordenschat, informatieoverdracht, werkhouding en zelfvertrouwen worden in het so minder duidelijk positieve resultaten waargenomen dan in het sbo.

Tabel 5.3 - Voor zowel so als sbo door tutoren waargenomen resultaten van het NELI programma (rij%)

	So (n=24)			Sbo (n=17)		
	niet	enigszins	duidelijk	niet	enigszins	duidelijk
Productieve woordenschat*	4	50	46	0	18	82
Zinsbouw en informatieoverdracht*	29	50	21	12	35	53
Narratieve vaardigheden	22	52	26	6	41	53
Taalbegrip	13	63	25	0	65	35
Sociaal gedrag	23	41	36	0	69	31
Werkhouding*	9	61	30	6	25	69
Luisterhouding	13	57	30	6	38	56
Zelfvertrouwen*	5	50	45	0	24	76

**significant verschil tussen so en sbo*

De tutoren in het so lichten hun antwoorden in de vragenlijst als volgt toe: *'De kinderen zijn moeilijk lerend. Je kunt niet verwachten dat ze in zo'n korte tijd groei door zouden maken'* (tutor so). En: *'Ik vind, voor de tijd die we aan NELI besteden, de opbrengst tegenvallen'* (tutor so). Een so-tutor noteert hierbij ook een reden: *'Het programma gaat nu te snel voor de kinderen. Zij hebben meer verwerkings- en leertijd nodig'.* Maar er zijn ook positieve geluiden: *'Je zag de leerlingen, zeker nu je terugkijkt, groeien'* (tutor so). En: *'De leerlingen zijn meer gaan praten'* (tutor so). Een sbo-tutor licht toe: *'Ik merk dat de leerlingen die de NELI lessen volgen meer zelfvertrouwen hebben gekregen om te vertellen. De meeste kinderen kunnen een logisch verhaaltje vertellen'.* En: *'Drie* [van de vier] *leerlingen zijn er gegroeid, van amper praten naar gesprekjes voeren'* (tutor sbo). En: *'De woordenschat van leerlingen is vergroot en dit hoor je terug in het kringgesprek of het narratieve gedeelte'* (tutor sbo). In het sbo wordt onder meer opgemerkt dat de leerlingen door NELI meer plezier hebben in taal en dat ze de aangeboden woorden goed onthouden.

Ook in de groepsgesprekken hoorden we veel positieve geluiden over de effecten van NELI. Zo werd naar voren gebracht dat sommige leerlingen een duidelijke sprong hebben gemaakt van, eerst het uiten van enkele woorden, tot later in het traject het formuleren van volledige zinnen en het begrijpen van verhaalstructuren. Een opvallend voorbeeld is een leerling die voorheen nauwelijks sprak, maar nu een verhaal van begin tot eind kan navertellen, zelfs met geluiden, wat volgens de geïnterviewde duidt op een goed begrip van de verhaalstructuur. Wat betreft de effecten op de sociaal-emotionele ontwikkeling wordt waargenomen dat sommige leerlingen duidelijk mondiger zijn geworden en meer zelfvertrouwen hebben gekregen. Dit uit zich onder meer in meer spontane gesprekken en het durven delen van verhalen. De betrokkenen brengen deze ontwikkeling in verband met een grotere woordenschat en het ontstane besef bij de leerlingen dat zij complexe woorden beheersen, wat een positief effect heeft op hun zelfbeeld. Leerlingen die voorheen stil of teruggetrokken waren, laten zich nu vaker zien als sociaal vaardiger en nemen zelf het initiatief tot contact met andere kinderen. Volgens de geïnterviewden draagt NELI dus, naast een grotere winst in taalvaardigheid, ook zichtbaar bij aan het zelfbeeld van leerlingen, doordat zij ontdekken dat ze daadwerkelijk iets te vertellen hebben en in staat zijn complexere woorden te gebruiken.
Hoewel er hier en daar duidelijk vooruitgang wordt waargenomen in het gedrag van leerlingen, zoals een toename in praten en interactie, is het niet altijd eenvoudig deze

ontwikkeling rechtstreeks toe te schrijven aan NELI, omdat leerlingen vaak ook andere vormen van begeleiding krijgen, zoals logopedie. Men heeft de indruk dat de effecten sterker zijn bij extraverte leerlingen, omdat introverte leerlingen minder snel geneigd zouden zijn taal actief in te zetten.

We vroegen of NELI inzichten heeft opgeleverd die zonder deze interventie niet, of minder snel aan het licht zouden zijn gekomen. Het zowel in de vragenlijst als in de interviews meest genoemde inzicht betreft een verdiept en breder zicht op de capaciteiten van leerlingen. Zo werd in de interviews naar voren gebracht dat sommige leerlingen meer woorden blijken te kennen, of meer praten dan ze normaal gesproken laten zien. Een leerling die begon met een laag startniveau, verraste door zijn grote inzet en door een onverwachte voortuitgang in taalgebruik. Volgens sommigen komen de betere prestaties eerder dan verwacht aan het licht door de individuele en systematisch aandacht die kenmerkend is voor NELI. De intensieve begeleiding in kleine groepjes maakt specifieke hulpvragen zichtbaar die in een grotere setting minder snel aan het licht komen. Soms ontstaat er meer inzicht in de specifieke behoeften van leerlingen. Het gebruik van screeningsinstrumenten en het invullen van tussenevaluaties maken de individuele ontwikkeling inzichtelijk en bieden concrete aanknopingspunten voor gerichte ondersteuning: *'Voor sommige leerlingen hebben wij ingezien dat bijvoorbeeld logopedie helpend zou zijn'* (tutor so).
Men ervaart dat de aanpak bijdraagt aan een verdiept inzicht in wat voor de leerlingen belangrijk is. Dat geldt bijvoorbeeld voor de consequente, terugkerende lesstructuur. Hoewel dit principe al wel bekend was, bevestigt en versterkt het programma dit inzicht door het systematisch toe te passen binnen elke les. Het besef dat herhaling en een duidelijke, vaste structuur essentieel zijn voor het leerproces van de leerlingen, is hierdoor gegroeid.
Een tutor uit cluster 3 merkt op dat leerlingen, doordat zij tijdens de sessies worden gestimuleerd om meer te spreken en minder terug te vallen op gebaren, daadwerkelijk tot meer verbale productie komen.
Hoewel de narratieve activiteiten uitdagend worden genoemd, laten de leerlingen, bijvoorbeeld bij het navertellen van een verhaal of het in de juiste volgorde leggen van plaatjes, zien dat zij hier met passende ondersteuning wel degelijk toe in staat zijn: *'Ik vind het heel leerzaam om te zien hoe onze leerlingen toch leerbaar zijn, ook in het narratieve deel. Als we niet met NELI hadden gewerkt, had ik dit niet ontdekt'* (tutor cluster 3). Door intensiever met leerlingen in gesprek te gaan, ontstaat daarnaast meer zicht op hun woordenschat en zinsbouw.
In plaats van te vertrekken vanuit wat een kind nog niet kan, kiest NELI volgens sommigen voor een benadering die uitgaat van het eigen verhaal van de leerling. Deze aanpak levert een rijker en genuanceerder beeld op van zowel de sterke als de ontwikkelpunten van een kind en wordt ervaren als een leuke en effectieve manier van werken.

Sommige betrokkenen ervaren een duidelijke meerwaarde van NELI voor de eigen professionele ontwikkeling. Uit Tabel 5.4 blijkt dat men door het werken met NELI vooral meer kennis heeft opgedaan over specifieke leerlingen, waarvan hierboven voorbeelden zijn gegeven. Dat geldt het sterkst voor de professionals in het sbo. Uit de groepsgesprekken blijkt dat NELI tutoren en leerkrachten geholpen heeft hun begrip van taalontwikkeling te verdiepen en het aanbieden van nieuwe woorden te verbeteren. Zo heeft NELI bijgedragen aan hun professionele groei, vooral ook bij degenen die nog maar kort in het onderwijs

werkzaam zijn (voor een enkeling was dit het geval). Ook wordt genoemd dat men bewuster is gaan letten op aspecten zoals actief luisteren en het belang van routine. Een tutor geeft aan meer taal te gebruiken tijdens spel met kinderen en zich meer bewust is geworden van het belang van herhaling van woorden en begrippen. De gestructureerde aanpak van NELI wordt daarbij gewaardeerd: het programma biedt helderheid en houvast, terwijl het tegelijkertijd ruimte laat voor eigen invulling en creativiteit. In meerdere interviews benadrukken de deelnemers dat ze zich meer bewust zijn geworden van het belang van de consistente structuur die NELI biedt, vooral in vergelijking met de doorgaans variabele invulling van lessen die zij zelf ontwikkelen.

Tabel 5.4 - Door tutoren waargenomen resultaten van het NELI programma, voor zowel so als sbo (rij%)

	So (n=24)			Sbo (n=17)		
	niet	enigszins	duidelijk	niet	enigszins	duidelijk
Meer kennis over specifieke leerlingen*	4	50	46	0	12	88
Specifieke kennis over spraak-taalproblemen	13	58	29	0	59	41
Meer zelfvertrouwen in de omgang met deze leerlingen	26	30	43	24	29	47
Inzicht in hoe leerlingen met taal/communicatie-problemen vooruitgang boeken	4	54	42	12	29	59
Bredere kennis over het soort leeromgeving dat deze leerlingen nodig hebben	8	50	42	12	41	47

**significant verschil tussen so en sbo*

5.10 Voor welke leerlingen is NELI het meest geschikt?

NELI is vooral bedoeld voor de kleutergroepen en de geïnterviewden vinden dit over het algemeen passend. Naar hun idee is het programma het best uitvoerbaar met kleuters in groep 2, met eventueel een uitloop naar het begin van groep 3. Voor jongere net ingestroomde kleuters, die vaak minder concentratie kunnen opbrengen en nog moeten wennen aan schoolse regels, zoals stilzitten en op je beurt wachten, is het programma naar hun idee minder geschikt. De jongste kleuters hebben over het algemeen meer moeite om de inhoud te volgen dan oudere kleuters en leerlingen van groep 3. Dit geldt bijvoorbeeld voor vrij complexe opdrachten waarbij ze een woord moeten omschrijven zonder het woord zelf te benoemen. Sommigen raden NELI af in de kleutergroepen en bevelen het vooral aan voor oudere leerlingen in de middenbouw, die doorgaans meer rust en een groter concentratievermogen hebben. NELI wordt over het algemeen als passend gezien voor leerlingen met verschillende onderwijsbehoeften. Maar er zijn wel verschillen tussen de clusters. Met name in cluster 3 zijn er leerlingen bij wie de ontwikkeling stagneert, of bij wie door een fysieke afwijking of gehoorproblemen de verstaanbaarheid en klankontwikkeling worden belemmerd. Voor deze leerlingen zijn soms individuele aanpassingen nodig. Een voorbeeld is dat van een leerling met een beperkte mondmotoriek, die thuis geen Nederlands spreekt en beperkt Nederlandse taalinput krijgt. Deze leerling heeft zoveel moeite met het programma, dat men zich afvraagt of het niet te hoog gegrepen is. Specifiek voor cluster 4 leerlingen blijken groepssessies over het algemeen uitdagend, omdat deze leerlingen snel zijn afgeleid en moeite hebben hun aandacht bij de activiteit te houden. Individuele sessies worden daarom als effectiever ervaren, aangezien deze ruimte bieden voor een meer op maat gemaakte aanpak, die zorgt dat leerlingen meer betrokken

meedoen. Zo liet een meisje tijdens groepssessies dermate weinig betrokkenheid zien dat besloten is haar uitsluitend individuele begeleiding te geven. Hoewel dit voor haar ontwikkeling goed leek uit te pakken, was de keerzijde dat ze de sociale interacties in de groepssessies zichtbaar miste.
Veel leerlingen die aan NELI meedoen, spreken thuis een andere taal dan Nederlands. Hoewel deze leerlingen wel nieuwe woorden verwerven, blijkt het toepassen in zinnen uitdagend. Sommigen noemen NELI voor deze groep weinig effectief als het gaat om zinsopbouw. Anderen zijn juist van mening dat NELI heel geschikt is voor NT2 kleuters, vanwege het systematische aanbod aan nieuwe worden en de expliciete ondersteuning van narratieve vaardigheden. Nóg meer focus op zinsopbouw, eventueel ondersteund door aanvullende methoden zoals 'Zien is Snappen', zou voor deze groep een waardevol aanbod zijn, zo suggereert één van de schoolteams.

5.11 Wat is er nodig om NELI nog meer effecten te laten sorteren?

Om NELI (nog) meer effect te laten sorteren is volgens de respondenten onder meer het volgende nodig. In de eerste plaats een betere afstemming tussen NELI en wat er in het reguliere lesprogramma gebeurt. Deze afstemming kan thematisch van aard zijn, maar wanneer dat niet haalbaar is, kan ook worden aangesloten bij de speciale woorden. NELI bevat bijvoorbeeld een afbeelding van een bloemenwinkel met het doelwoord *'betalen'*. Dit woord kan echter net zo goed worden geïntegreerd binnen een ander thema, zoals *'vervoer'*, als dat op dat moment centraal staat in de reguliere klas. Op die manier ontstaat samenhang tussen beide contexten, zonder dat de inhoud volledig thematisch hoeft samen te vallen. Een suggestie is om de Flashcards[7] met speciale woorden ook te gebruiken in de klas. De samenhang kan verder worden versterkt wanneer leerkrachten een NELI sessie bijwonen. Door mee te kijken raken ze bekend en beter vertrouwd met de werkwijze van het programma, wat kan zorgen voor meer betrokkenheid en een betere aansluiting met het reguliere onderwijs.
Ook aanpassingen in het NELI-programma worden genoemd om meer effect te sorteren. Sommigen opperen minder lesstof aan te bieden en meer te herhalen. Enkele tutoren in het so geven aan behoefte te hebben aan handvatten om het narratieve deel goed uit te kunnen voeren met hun groep. Meermaals komt naar voren dat de inhoud van NELI hier en daar verouderd is, bijvoorbeeld als het gaat om plaatjes, woorden, maar ook thema's. Dat is voor sommigen ook reden geweest om verouderde of te ingewikkelde woorden te vervangen door meer alledaagse termen. Zo is 'luchthaven' vervangen door 'vliegveld' en 'verkeerslicht' door 'stoplicht', zodat de taal meer past bij wat de kinderen dagelijks horen. Een tutor heeft de oefening *'treinkaartje kopen'* vervangen door *'inchecken met een OV-chipkaart'*, in lijn met de huidige praktijk. Ook is er een tendens om de inhoud aan te passen aan het hogere niveau van oudere kinderen, aangezien de thema's soms als te kleuterachtig worden ervaren. Het programma zou aan kracht winnen als er beter wordt aangesloten bij 'thema's van deze tijd': *'Ik zou graag wereldoriëntatie-onderwerpen willen behandelen'*.
Ten slotte zou NELI volgens sommigen meer effect hebben als een oudercomponent wordt toegevoegd. Aanbevolen wordt om naast het informeren van ouders te zorgen voor

7 Onderdeel van NELI: kleine kaartjes met een woord en het bijbehorende plaatje, als leermiddel om de woorden effectief te onthouden.

materiaal voor de ouders, bijvoorbeeld bladen met nieuwe woorden, waarmee ze thuis aan de slag kunnen.

5.12 Transfer onderdelen van NELI naar de reguliere les

Gevraagd is of onderdelen, aspecten of suggesties uit het NELI-programma zijn overgenomen in de reguliere les. Deze vraag is om twee redenen van belang. Ten eerste om inzicht te krijgen in welke didactische principes of elementen van NELI als zo waardevol worden ervaren, dat men ze zou willen integreren in het reguliere onderwijs. Ten tweede geeft het een indicatie van de mate waarin leerlingen die niet deelnemen aan NELI, maar wel in de reguliere klas zitten, indirect met (elementen van) het programma in aanraking zijn gekomen en ook hiervan hebben kunnen profiteren, het zogenoemde *spill-over*-effect.
Uit Tabel 5.5 blijkt dat het vooral gaat om twee aspecten. Ten eerste: ruim een derde (35%) van de so-tutoren en ruim de helft (54%) van de sbo-tutoren noemt strategieën of handvatten die voor (bepaalde) leerlingen werken (geldt significant meer voor het sbo). Ten tweede: 40% van de so- en 53% van de sbo-tutoren noemt overname van doelwoorden en narratieve aspecten in de reguliere les.
Uit de groepsgesprekken blijkt over het algemeen dat het karakter van de reguliere lessen niet is veranderd en leerkrachten in de reguliere lessen niet anders zijn gaan werken als gevolg van NELI. Een belangrijke reden waarom leerkrachten in de reguliere les niet voortborduren op NELI, is het gebrek aan inzicht in de precieze inhoud en uitvoering van NELI-sessies. Dat heeft deels te maken met tijdgebrek; de meesten hebben geen tijd om zich echt in NELI te verdiepen.
Sommigen merken op dat het niet altijd past om nog een extra werkwijze toe te voegen aan de bestaande praktijk: scholen werken al met diverse aanpakken en methoden, waardoor het lastig kan zijn om NELI hierin betekenisvol te integreren. Hoewel er in het speciaal onderwijs veel ambulante begeleiding is voor taalonderwijs in de klas, is een directe koppeling met NELI niet altijd mogelijk.
Sommige tutoren zijn zelf ook leerkracht en staan op de groep, naast hun NELI-tutoring. Dan is de aanpak vaak wat meer geïntegreerd. Enkele leerkrachten geven aan dat NELI een goede aanvulling vormt. Zij hebben tips uit de training meegenomen die goed toepasbaar zijn in de klas en bij de leerlingen aanslaan, zoals het aanwijzen van 'de beste luisteraar'. Hiervoor is plaats ingeruimd in het afsluitende (plenaire) deel van een sessie, waarin aan de bij NELI behorende handpop Ted wordt gevraagd welk kind hij vandaag als beste luisteraar heeft gekozen en waarom.
Verder hebben sommigen gezocht naar thematische overlap in onderwijsaanbod bij bepaalde onderwerpen, bijvoorbeeld wanneer zowel in de reguliere les als in NELI over onderwerpen als beroepen of groenten wordt gesproken. Een leerkracht probeerde aan te haken bij de NELI-onderwerpen en woorden terug te laten komen in de klas door een relatie te leggen met het klassenthema. Een andere leerkracht heeft een overzicht van de NELI-doelwoorden gedeeld met alle betrokken leerkrachten en logopedisten, die de woorden wilden proberen te integreren in hun lessen. Maar in het algemeen vinden de betrokkenen thematische afstemming tussen NELI en het reguliere curriculum lastig, omdat thema's in de verschillende situaties niet altijd samenvallen. Er is over het algemeen dus geen sprake van een bewuste, structurele afstemming van thema's tussen NELI en de reguliere lessen. Sommige leerkrachten zouden dit wel graag zien, omdat een afgestemd aanbod elkaar juist kan versterken en herhaling meer effect kan hebben. Sommige groepsleerkrachten zouden

ook wel graag willen zien hoe NELI in de praktijk werkt, om er vervolgens in de klas mee aan de slag te gaan. Anderen vinden juist dat een verschillend aanbod waardevol is.

Tabel 5.5 - Het aandeel van de tutoren dat aangeeft dat de groepsleerkracht onderdelen, aspecten of suggesties uit het programma heeft overgenomen in de reguliere les, voor zowel so als sbo (%)

	So (n=24)	Sbo (n=17)
Strategieën of handvatten die voor (bepaalde) leerlingen werken*	35	54
Thema's	9	20
Verhalen	18	27
Doelwoorden en narratieve aspecten	40	53
Principes, zoals 'overlearning', multi-zintuigelijke technieken	31	27

**significant verschil tussen so en sbo*

5.13 Knelpunten en pluspunten bij de uitvoering van NELI

In de tutorvragenlijst is de tutoren een lijst met knel- en pluspunten voorgelegd met de vraag in hoeverre zij deze ervaren bij de uitvoering van NELI. De knelpunten staan in Tabel 5.6, de pluspunten in Tabel 5.7.

Knelpunten: belemmerende factoren en omstandigheden
Bij de knelpunten waren de antwoordalternatieven 1 'geen knelpunt', 2 'enigszins een knelpunt' en 3 'een duidelijk/sterk knelpunt'. Uit Tabel 5. 6 blijkt dat het ontbreken van een geschikte ruimte in het so vaker een knelpunt is dan in het sbo. Het feit dat de leerlingen uit de les worden gehaald en dat de interventie iedere dag moet worden uitgevoerd, wordt op beide schooltypen het vaakst genoemd.
Ook in de groepsgesprekken komt het knelpunt 'uit de les halen' vaak naar voren. Door hun deelname aan NELI zijn deze leerlingen minder aanwezig in de reguliere klas. Hierdoor missen zij onderdelen van het reguliere programma tijdens de NELI-sessies. Bovendien ziet de leerkracht deze leerlingen minder vaak, waardoor het risico bestaat dat bepaalde aspecten van hun ontwikkeling minder goed in beeld blijven. Sommige leerlingen nemen naast NELI deel aan andere interventies, zoals logopedie. Dit leidt tot een vol rooster. Het vraagt van de verschillende professionals dat zij zorgen voor een goede afstemming tussen de verschillende activiteiten in het programma van deze leerlingen.

Groepsleerkrachten uit de reguliere les
Net als voor leerlingen die uit de klas worden gehaald, geldt voor reguliere leerkrachten die NELI-sessies verzorgen dat zij door hun afwezigheid bepaalde zaken niet meemaken. Zij zien de leerlingen minder in de klas en lopen daardoor het risico een deel van het groepsgebeuren en bepaalde ontwikkelingsaspecten van individuele leerlingen te missen. Wanneer leerkrachten vanwege het verzorgen van de NELI tutoring uit de klas zijn, heeft dit invloed op de klassenbezetting. Er zijn dan minder beschikbare handen in de groep. Het wegvallen van een leerkracht in de klas legt daarmee extra druk op de achterblijvende collega.

Organisatie en inpassing van het programma
Vrijwel overal wordt de planning en logistiek van NELI als een grote uitdaging ervaren. Het inroosteren van leerlingen en tutoren kost veel tijd en is een complexe puzzel. Leerlingen

nemen vaak deel aan meerdere vormen van ondersteuning en tutoren hebben soms ook andere taken
waarmee in de planning rekening moet worden gehouden. De intensiteit van het programma en de vereiste continuïteit maken het moeilijk om meer dan één of twee groepen kwalitatief goed te begeleiden, terwijl hier wel behoefte aan bestaat, aangezien meer leerlingen baat zouden hebben bij NELI. Op de meeste scholen wordt NELI ingezet naast andere methoden ter bevordering van taal en geletterdheid bij kinderen die niet of nauwelijks spreken. Deze combinatie maakt de inpassing van NELI in de dagelijkse onderwijspraktijk complex. NELI vergt een aanzienlijke tijdsinvestering, onder meer voor herhaling en het bijhouden van de administratie (registratie). Dit maakt het moeilijk om het programma binnen de gestelde 20 weken af te ronden. De dagelijkse routine in het speciaal onderwijs is vaak minder voorspelbaar dan in het regulier onderwijs, waardoor NELI minder goed inpasbaar is in het so. De planning van sessies en de selectie van leerlingen voor NELI vormen een logistieke uitdaging, met name wanneer leerlingen gedurende het schooljaar in- of uitstromen. De intensiteit van het NELI-programma, vooral de vele activiteiten in groepssessies en het voortdurende schakelen, vraagt veel van vooral de jongste leerlingen in het speciaal onderwijs.

Personele bezetting
Bij personeelstekorten, bijvoorbeeld door ziekte, komt de uitvoering van NELI in het gedrang. Als back-up zijn soms groepsleerkrachten geschoold in NELI, maar in de praktijk blijkt hun inzet in NELI lastig, omdat zij immers (ook) een klas moeten begeleiden.

Ruimte
Een knelpunt betreft de beschikbaarheid van geschikte ruimtes. De huidige praktijk van het gebruiken van gangen of een tijdelijk ingericht plekje in de klas is niet ideaal.

Het projectmateriaal
Het programmaonderdeel 'wie is de beste luisteraar' wordt enerzijds geroemd, anderzijds zorgt het regelmatig voor onrust en onderlinge conflicten onder de kinderen. Om die reden wordt het soms overgeslagen of afgedaan met een algemeen compliment, in plaats van een gerichte waardering voor een specifieke leerling. Er is zelfs voorgesteld dit onderdeel te schrappen. Maar anderzijds opperen sommigen om het ook op te nemen in de reguliere lessen.
Het uitdelen van beloningsstickers leidt tot scheve ogen. Om jaloezie en gevoel van uitsluiting te voorkomen, hebben sommigen besloten om de verdiende stickers direct in de tas van de leerlingen te doen en niet mee de klas in te nemen. Het registreren van voortgang en gedrag na elke sessie kost veel tijd, waarvoor in de praktijk niet altijd ruimte is.

Tabel 5.6 - Het aandeel van de tutoren dat aangeeft knelpunten te ervaren bij de uitvoering van het NELI programma, voor zowel so als sbo (rij%)

	So (n=24)			Sbo (n=17)		
	Geen	enigszins	duidelijk /sterk	geen	enigszins	duidelijk /sterk
Onvoldoende voorbereidingstijd	26	52	22	53	41	6
Onvoldoende tijd voor uitvoering	42	29	29	35	41	24
Het feit dat de interventie elke dag moet worden uitgevoerd	15	50	35	13	63	25
Dat leerlingen uit de reguliere les worden gehaald	22	43	35	29	29	41
Onvoldoende ondersteuning binnen de school	59	36	5	63	31	6
Onvoldoende ondersteuning door interventieaanbieder (of andere externe ondersteuner)	78	22	0	88	13	0
Projectmateriaal onvoldoende	57	30	13	71	29	0
Geen geschikte ruimte/klaslokaal*	48	35	17	88	6	6
Te weinig draagvlak in het reguliere team	63	25	13	63	38	0
Te weinig middelen	76	19	5	81	19	0
weinig of geen medewerking of belangstelling van de ouders van de leerlingen	58	32	11	54	31	15

**significant verschil tussen so en sbo; De kolommen 'geen', 'enigszins' en 'duidelijk/sterk' geven de mate weer waarin tutoren aangaven het genoemde aspect als knelpunt te ervaren.*

Pluspunten: positieve ervaringen en punten van succes

Bij de pluspunten (Tabel 5.7) zijn de antwoordalternatieven 1 'niet positief', 2 'enigszins positief' en 3 'duidelijk/zeer positief'. De motivatie bij alle betrokkenen om NELI te laten slagen is bij het sbo duidelijk positiever dan bij so, maar is bij het so evengoed positief. Opvallend is dat de aansluiting tussen NELI en het reguliere lesprogramma vrij vaak als niet zo positief wordt beoordeeld. Het enthousiasme van de leerlingen over NELI valt op als heel duidelijk positief. Dit wordt door de tutoren in beide schoolsoorten als één van de belangrijkste pluspunten benoemd. Ook de begeleiding door Vinci wordt duidelijk als positief ervaren. In het overzicht van uitspraken was ook een uitspraak opgenomen over de ouders ('het enthousiasme van de ouders van de leerlingen'). Dit was geen duidelijk positief punt en in het so zelfs wat negatiever dan in het sbo. Het contact met ouders over NELI is vooral in het so beperkt, doordat veel ouders nauwelijks op school komen, en leerlingen vaak met leerlingvervoer naar school worden gebracht. Zoals genoemd hebben de scholen de ouders geïnformeerd, en over het algemeen reageren deze positief op het feit dat hun kind extra ondersteuning krijgt, maar verder blijft het contact met ouders over NELI doorgaans minimaal.

Volgens de geïnterviewden zijn de volgende factoren of omstandigheden bevorderend voor de uitvoering van NELI. Deels zijn deze punten een tegenhanger van al genoemde negatieve punten.

Personele bezetting

Essentieel voor een goede uitvoering van NELI is voldoende personele bezetting. Het werken met een vaste tutor die zich specifiek bezighoudt met de NELI-groep wordt als essentieel beschouwd. De tutor moet de tijd krijgen om na de sessies de voortgang

zorgvuldig vast te leggen. Met het oog op structuur en efficiëntie verzorgt idealiter één tutor zowel de groeps- als individuele sessies.

Ruimte en tijd
Voor een goede uitvoering van NELI is het van belang te kunnen beschikken over een rustige, vaste ruimte. Een aparte NELI-ruimte waarin materialen kunnen blijven staan en posters en andere visuele hulpmiddelen permanent aanwezig zijn, wordt essentieel genoemd. Het ontbreken van een dergelijke ruimte, of het steeds moeten wisselen of uitwijken naar minder geschikte plekken zoals een gang, leidt tot verstoringen. Ook tijd voor gedegen voorbereiding is van belang voor een goede uitvoering.

Enthousiasme van de leerlingen
Het plezier dat NELI bij de leerlingen teweegbrengt, werkt motiverend.

Groepsgrootte
Op de meeste scholen is gewerkt met groepjes van vier leerlingen, wat over het algemeen als ideaal en werkbaar wordt ervaren. Deze groepsgrootte biedt ruimte voor voldoende individuele aandacht en voor stimulerende onderlinge interacties. Bij een grotere groepsgrootte zijn de omstandigheden minder gunstig: er is meer onrust en minder ruimte voor betekenisvolle interacties en intensieve oefening van de doelwoorden. Bij een grotere groep van bijvoorbeeld zes leerlingen is de inzet van twee tutoren noodzakelijk om de groep goed te kunnen begeleiden en eventueel storend gedrag aan te pakken. Het werken in kleine groepjes is voor de meesten overigens geen nieuwe praktijk, aangezien zij hier in het s(b)o al aan gewend zijn.

Combinatie van groeps- en individuele sessies
De combinatie van groeps- en individuele sessies wordt over het algemeen als waardevol ervaren. De sessies vullen elkaar goed aan. De groepssessies dragen bij aan de ontwikkeling van sociale vaardigheden en een positieve groepsdynamiek, terwijl individuele sessies ruimte bieden voor verdieping, herhaling en gerichte ondersteuning die aansluit bij het ontwikkelingsniveau van het kind. Daarbij is de (voorgeschreven) spreiding van de sessies over de week wel belangrijk. Wanneer beide sessies op één dag worden gepland (wat soms organisatorisch noodzakelijk is), blijkt dit voor leerlingen over het algemeen te vermoeiend. Er is een duidelijk verschil in participatie en concentratie tussen groeps- en individuele sessies: in één-op-één situaties zijn kinderen minder snel afgeleid en is diepgaander doorvragen mogelijk, terwijl in groepen de onrust van één kind de hele sessie kan beïnvloeden. Verschillende respondenten benoemen de kracht van de herhaling en verdieping, als gevolg van het feit dat in de individuele sessies wordt voortgebouwd op de groepslessen.

Het projectmateriaal
De beschikbare materialen, zoals kant-en-klare woordkaartjes met plaatjes en de handpop Ted, worden positief ervaren en dragen bij aan de motivatie. Een vaste start en duidelijke lesstructuur, met afwisseling in werkvormen en kortere onderdelen, blijken prettig te werken.
De flexibiliteit om het programma aan te passen aan de behoeften van de kinderen, bijvoorbeeld door sessies te verdiepen of de volgorde van onderdelen te wijzigen, wordt als positief ervaren.

Draagvlak team

NELI sluit volgens de geïnterviewden goed aan bij de taalbeleidsplannen van scholen, omdat deze erop gericht zijn de taalontwikkeling van kinderen via een gevarieerd aanbod systematisch te stimuleren. NELI past goed bij bestaande praktijken, zoals het werken met taalgroepjes en een pop, zoals Ted, maar in vergelijking met de gangbare praktijk kent NELI meer herhaling en een langer durende thematische aanpak, wat als positief wordt beoordeeld.

Op de meeste deelnemende scholen is de schoolleiding nauw betrokken bij de implementatie van NELI, is deze hier enthousiast over en stimuleert de uitvoering ervan actief. De geïnterviewden ervaren duidelijke bereidheid van de schoolleiding om mee te denken over de organisatorische inbedding. Zo wordt er tijd en ruimte vrijgemaakt voor de voorbereiding van groepslessen, een essentiële randvoorwaarde voor de uitvoering. Ook zijn afspraken gemaakt over de ondersteuning door het hele schoolteam, waaronder de verwachting dat de NELI-leerlingen zoveel mogelijk aanwezig zijn en dat er tijdens de sessies geen andere activiteiten worden gepland.

Tabel 5.7 - Het aandeel van de tutoren dat aangeeft positieve punten te ervaren bij de uitvoering van het NELI programma, voor zowel so als sbo (rij%)

	So (n=24)			Sbo (n=17)		
	niet positief	enigszins positief	duidelijk/ zeer positief	niet positief	enigszins positief	duidelijk/ zeer positief
De motivatie bij alle betrokkenen om NELI te laten slagen*	4	46	50	0	19	81
Goed verlopend overleg met/ondersteuning door de reguliere leerkracht(en)	0	45	55	0	38	62
Goede ondersteuning vanuit Vinci	0	32	68	0	27	73
De aansluiting tussen NELI en het reguliere lesprogramma	48	43	10	43	50	7
Goede communicatie tussen alle betrokkenen	0	42	58	0	38	63
Enthousiasme van de leerlingen over NELI	0	29	71	0	12	88

**significant verschil tussen so en sbo*

5.14 Redenen om NELI aan te bevelen aan andere scholen

Zowel in de vragenlijst als in de groepsgesprekken is gevraagd of men NELI zou aanbevelen aan andere scholen. Bijna alle tutoren, 95%, beveelt andere scholen aan om met NELI te gaan werken. Dit verschilt niet tussen so en sbo.

Door tutoren in het so genoemde redenen om NELI aan te bevelen zijn onder meer:

'NELI heeft een duidelijke structuur en voldoende handvatten om leerlingen beter te leren communiceren, met meer inhoud.'

'Het is een duidelijk en leuk programma. Veel materialen en suggesties.'

'Het is erg leerzaam. Je biedt hiermee begeleiding in kleinere groepjes en doet effectieve en actieve opdrachten met leerlingen.'

'Leerzaam, volledig verzorgd programma. Ook leerzaam om te zien hoe (zml/cluster 3) leerlingen toch leerbaar zijn in het narratieve gedeelte. Als we niet met NELI waren gaan werken, had ik dit niet ontdekt.'

'Kinderen kunnen je echt verrassen hoe ze woorden en verhalen oppakken als je het structureel aanbiedt.'

Door tutoren in het sbo genoemde redenen om NELI aan te bevelen zijn onder meer:

'Als er ruimte voor is zou ik het zeker aanbevelen voor de taalzwakke leerlingen, omdat er duidelijke vooruitgang te zien is.'
'Het is een duidelijke en gestructureerde methode. Heel prettig om mee te werken. Ook vind ik het heel fijn dat er niet van alles omheen bedacht wordt, denk dan aan zaken die er eigenlijk niet toe doen maar zo 'gezellig of leuk' zijn.'
'Het is een doordachte methode die je na de cursus kunt volgen zonder externen te hoeven aantrekken.'
'Het is een leuke en effectieve manier om taal bij jonge kinderen te bevorderen.'
'Het is een programma waar je (mits je goed rapporteert) goed en snel een duidelijk resultaat ziet en hoort.'
'Ik merk aan de kinderen dat ze veel meer zijn gaan praten en op een logische manier een verhaal kunnen vertellen. De kinderen hebben meer zelfvertrouwen gekregen door het NELI-programma.'
'Omdat wij duidelijk vooruitgang zien, zeker op het gebied narratief.'
'Omdat de structuur van het NELI programma zeer helpend is als ook de aandacht voor de narratieve vaardigheden/mondeling taalgebruik.'

Tevredenheid over NELI

In de vragenlijst is gevraagd hoe tevreden men al met al is over het programma NELI. De antwoordalternatieven konden worden ingevuld op een schaal van 1 (helemaal mee oneens) tot en met 5 (helemaal mee eens). Tabel 5.8 laat zien dat de tutoren in hoge mate tevreden zijn over het programma. Het hoogst scoort de uitspraak 'het is inspirerend om met NELI te werken'. Er zijn geen verschillen tussen het so en het sbo.

Tabel 5.8 - Ervaringen met NELI: apart voor so en sbo de gemiddelde score en het percentage van de tutoren dat het met de uitspraak eens is: (helemaal) oneens (% 1+2) en (helemaal) eens (% 4+5)

	so			sbo		
	gem.	(helemaal) oneens	(helemaal) eens	gem.	(helemaal) oneens	(helemaal) eens
Het is inspirerend om met het NELI-programma te werken	4.0	0	83	4.1	0	94
Ik voel me bij dit project ondersteund door de coach	3.6	4	58	3.6	12	76
NELI helpt mij om me in mijn werk verder te ontwikkelen	3.7	4	58	3.9	6	76
Ik ben tevreden over hoe NELI loopt	3.6	8	63	3.9	0	82
Dit project maakt verschil voor mijn werk	3.3	8	29	3.8	6	65
Praktische zaken zijn goed geregeld	3.6	8	63	3.8	6	71

Voortzetting NELI

Over het geheel genomen ervaren de betrokkenen duidelijk in sterkere mate pluspunten dan knelpunten bij de uitvoering van NELI. Gezien de positieve ervaringen en de zichtbare vooruitgang bij de kinderen zouden verschillende scholen het programma dan ook graag willen voorzetten. Er zijn echter sterke twijfels over de haalbaarheid. Deze hangt af van organisatorische randvoorwaarden, zoals de beschikbaarheid van voldoende personeel (met name een tutor), geschikte ruimtes en de benodigde tijdsinvestering.

De implementatie van NELI vraagt om organisatorische aanpassingen, aangezien de werkwijze afwijkt van de normale gang van zaken. Op sommige scholen wordt nagedacht over specifieke toepassingen, zoals in de kleine kring en voor NT2-kinderen. Ook wordt opgemerkt dat NELI goed past binnen een structuurgroep (kleine, zeer gestructureerde klas voor leerlingen met intensieve ondersteuningsbehoeften), zoals die worden ingezet in het so. Dit zou volgens betrokkenen effectiever kunnen zijn dan NELI aanbieden in een reguliere groep met grotere niveauverschillen. Ook wordt gedacht over aanpassingen in de opzet van de sessies. Geopperd wordt om de individuele sessies te combineren met logopedie. Sommigen benadrukken de noodzaak van meer financiële ondersteuning voor een duurzame inbedding. Ook denken sommige scholen na over de inzet van vrijwilligers voor de uitvoering, zodat onderwijsassistenten de reguliere groep kunnen blijven ondersteunen.

5.15 Aanbevelingen voor de ontwikkelaar

De begeleiding en ondersteuning vanuit Vinci worden positief beoordeeld, met name vanwege de snelle respons op vragen. Aanbevelingen om de uitvoering te verbeteren betreffen vooral het materiaal. Sommigen ervaren registratiedruk en hebben behoefte aan meer duidelijkheid over het doel van de registratie, bijvoorbeeld of deze puur voor eigen gebruik is of ook een ander doel dient. Vinci zou duidelijker moeten maken wat precies het doel is van de registratie en wat er met alle verzamelde informatie gebeurt. Daarbij worden de papieren formulieren onhandig genoemd en is er een sterke voorkeur voor een digitale registratietool.
Er is behoefte aan aanvullend beeldmateriaal, zoals filmpjes en geluiden, ter illustratie van de thema's. Dit zou volgens sommigen de betrokkenheid van de kinderen kunnen vergroten. Ook wordt voorgesteld om te zorgen voor bladen met nieuwe woorden, om mee naar huis te geven met als doel de ouderbetrokkenheid bij het programma te vergroten.
Sommigen vinden het narratieve deel van de verhaalkaart te hoog gegrepen voor de kinderen; een alternatief is om kinderen te laten tekenen en over hun tekening te laten vertellen.
Er is behoefte aan een betere en praktischere aanlevering van de materialen, bijvoorbeeld in een map in plaats van een doos. Zo wordt voorgesteld om een ordner mee te leveren waarin de materialen per week overzichtelijk kunnen worden opgeborgen. Bij voorkeur gaat het om materiaal dat direct inzetbaar is, aangezien het lamineren van materialen veel tijd kost.
De pop Ted wordt als minder aantrekkelijk ervaren dan in de filmpjes en er is een wens voor een mooiere versie.

Verschillen tussen so en sbo
Op basis van de uitkomsten uit de vragenlijsten zien we de volgende verschillen tussen het so en sbo als het gaat om de uitvoering. In het sbo wordt de begeleiding meestal door één tutor uitgevoerd, terwijl in het so groepjes vaker samen worden begeleid. Het so is later gestart met de uitvoering van het programma, vooral door personeelstekorten, en kampt bovendien vaker met een gebrek aan geschikte ruimtes dan het sbo.
Tijdens de uitvoering van de sessies blijkt in het so meer herhaling nodig te zijn, onder andere door de kortere spanningsboog van leerlingen. Hierdoor gaat er relatief meer tijd naar groepssessies. Onderdelen zoals het narratieve deel verlopen in het so minder goed dan in het sbo. Ook differentiëren (het afstemmen op niveau, interesses en behoeften van leerlingen) blijkt in het so lastiger, al wordt hier wel actief naar gezocht. Met name

logopedisten proberen doelen op hun eigen, vaak creatievere en minder strikte manier vorm te geven, bijvoorbeeld door één activiteit intensief te behandelen in plaats van meerdere oppervlakkig.
Wat betreft resultaten worden in het so minder duidelijke effecten waargenomen dan in het sbo, bijvoorbeeld op het gebied van productieve woordenschat, informatieoverdracht, werkhouding en zelfvertrouwen. Daarnaast geven professionals in het sbo aan dat zij meer kennis over specifieke leerlingen hebben opgedaan en meer concrete strategieën hebben ontwikkeld die werken. Ook is hun motivatie om het programma te laten slagen iets positiever dan in het so, al is de houding in het so eveneens overwegend positief.

6 Resultaten effect interventie bij hogere uitvoeringskwaliteit

Zoals beschreven in de vorige paragraaf, verschillen scholen in de mate waarin het NELI programma is uitgevoerd zoals bedoeld. Ook de betrokkenheid en mate van deelname door kinderen varieerden. Voor de scores op de LanguageScreen hebben we de effectiviteit van het programma onderzocht in situaties waarin de uitvoering en deelname van hogere kwaliteit waren.
De resultaten van de tutorvragenlijst geven informatie over de mate waarin het programma is uitgevoerd zoals bedoeld: in hoeverre alle verschillende onderdelen zijn aangeboden, met voldoende duur ('kwantiteit uitvoering') en hoe de kwaliteit van uitvoering is ervaren. Voor beide aspecten is een indicator samengesteld.

6.1 Uitvoering - kwantiteit

Omdat de begeleiding vaak verzorgd werd door meerdere tutoren samen, is aan de hand van de oordelen van alle tutoren per school samen beoordeeld in hoeverre het aantal aangeboden onderdelen en de duur het volledige programma betroffen. Hiervoor werd één score gevormd op basis van de onderstaande indicatoren. Steeds werd een score van 1 toegekend wanneer het volgens plan verliep en een lagere score wanneer dit niet het geval is. Bij elke categorie vermelde we de score. De totaalscore is gevormd door per school alle scores op te tellen.

-aantal groepssessies: drie zoals bedoeld (1) of minder (0.5)
-duur groepssessies: minimaal half uur zoals bedoeld (1) of wat minder (0.5[8])
-aantal individuele sessies: twee zoals bedoeld (1) of minder (0.5)
-duur individuele sessies: 15 minuten zoals bedoeld (1) of wat minder (0.5[9])
-mate waarin het programma volledig is uitgevoerd: volledig (1), bijna volledig (0.75), grotendeels (0.5), gekomen tot halverwege (0).

Gemiddeld is de score 4.4 (maximum 5): vier van de 23 scholen hebben de maximale score; de laagste score is 3.5, wat op drie scholen voorkomt.

6.2 Uitvoeringskwaliteit

Wat de kwaliteit van uitvoering betreft, zijn indicatoren gevormd voor de mate waarin twee centrale programmaonderdelen naar het oordeel van de tutoren lukten: aandacht voor

8 Op twee scholen was het meer, dit is niet apart onderscheiden.
9 Op zes scholen was het meer, dit is niet apart onderscheiden.

luistervaardigheid en aandacht voor narratieve vaardigheden, beide in groeps- en individuele sessies (antwoordcategorieën 1 lukt helemaal niet – 5 lukt heel goed). Daarnaast is een score gevormd op basis van zes tutoroordelen over ervaren pluspunten bij de uitvoering, wat betreft de communicatie met en enthousiasme van betrokkenen, ervaren ondersteuning en communicatie, en aansluiting op het reguliere programma (antwoordcategorieën 1 niet positief – 3 duidelijk/zeer positief). Ook is een score gevormd met betrekking tot 11 ervaren knelpunten in de uitvoering (qua organisatie en draagvlak) zo gecodeerd dat een hogere score gunstiger is (antwoordcategorieën lopen van 1 een duidelijk/sterk knelpunt – 3 geen knelpunt). Voor de genoemde indicatoren voor uitvoeringskwaliteit zijn eerst gemiddelde scores per tutor berekend over de afzonderlijke items en vervolgens per school gemiddelde scores berekend over de tutoroordelen. Onderstaande tabel toont beschrijvende gegevens van de scores voor de 23 deelnemende scholen.

Tabel 6.1 - Gemiddelde uitvoeringskwaliteit voor de 23 scholen (tussen haakjes het theoretisch bereik van de scores)

	gem.	std.	min.	max.
Ervaren kwaliteit luistervaardigheid (1-5)	3.4	0.9	1.5	5.0
Ervaren kwaliteit narratieve vaardigheden (1-5)	4.0	0.7	2.0	5.0
Ervaren pluspunten (1-3)	2.6	0.3	2.0	3.0
Ervaren minpunten (1-3 gespiegeld: hoger is gunstiger)	2.4	0.3	1.9	3.0

6.3 Aanwezigheid en betrokkenheid

Gegevens over aanwezigheid en betrokkenheid geven inzicht in de mate waarin kinderen aan de interventie hebben deelgenomen. Per kind is voor zowel **aanwezigheid (1=nooit – 5=altijd)** als **betrokkenheid** (1 = helemaal niet- 5 = sterk) een gemiddelde score berekend op basis van het tutoroordeel met betrekking tot de groeps- en individuele sessies tijdens twee meetmomenten, namelijk halverwege en aan het eind van de interventie (vier items per construct). De kinderen die aan de interventie deelnamen waren gemiddeld genomen bijna altijd aanwezigheid en vrij sterk betrokken bij de individuele en groepssessies (zie onderstaande tabel).

Tabel 6.2 - Gemiddelde aanwezigheid en betrokkenheid bij de individuele en groepssessies van de aan de interventie deelnemende kinderen (n=119)

	gem.	std.	min.	max.
aanwezigheid bij de individuele en groepssessies (1-5)	4.2	0.5	2.5	5.0
betrokkenheid bij de individuele en groepssessies (1-5)	4.1	0.6	2.0	5.0

6.4 Groep met gunstigere uitvoeringskwaliteit – indicatie van potentieel te behalen effect van NELI

Om een indicatie te krijgen het effect op taalvaardigheid dat potentieel met NELI behaald kan worden, is het verschil vooruitgang in taalvaardigheid op de LanguageScreen berekend voor een subgroep van interventieleerlingen die naar een school gingen met gunstige uitvoeringskwaliteit en die in voldoende mate deelnamen. Voor de selectie van deze leerlingen zijn de volgende criteria gehanteerd: (1) een uitvoeringskwantiteit van 4,5 of

hoger; (2) een uitvoeringskwaliteit die gemiddeld of hoger is; (3) leerlingen die minimaal 'bijna altijd' aanwezig en (4) minimaal 'bijna altijd' betrokken waren bij deelname. De subgroep betreft in totaal 32 leerlingen: 19 sbo en 13 so leerlingen. De resultaten voor deze subgroep staan links in de onderstaande tabel[10]. Ter vergelijking is rechts in de tabel het interventie-effect voor de gehele interventiegroep weergegeven. Dit zijn enigszins andere verschillen dan die eerder werden gerapporteerd. De resultaten rechts in de tabel betreffen aanvullende analyses dit beter vergelijkbaar zijn. Hierbij is namelijk net als bij de analyse met de subgroep– geen gebruik gemaakt van matching, maar van een analyse met (dezelfde set) controlevariabelen. Hierdoor zijn de uitkomsten beter vergelijkbaar. In de tabel worden de gestandaardiseerde (voor leeftijd gecorrigeerde) LanguageScreen scores weergegeven die een standaarddeviatie van 10 hebben. De resultaten zien dat wanneer de interventie optimaal wordt uitgevoerd, de effecten aanzienlijk groter dan het gemiddelde effect in de totale interventiegroep. De toename in de score op de LanguageScreen is voor NELI-deelnemers bij optimalere uitvoeringskwaliteit ongeveer één standaarddeviatie hoger dan voor niet-deelnemers, wat aanzienlijk hoger dan het gemiddelde effect voor de totale interventiegroep (0.27 standaarddeviatie). Voor de subtest luisterbegrip geldt zelfs een effect van 1.6 standaarddeviatie bij optimale uitvoering, vergeleken met 0.45 standaarddeviatie voor de totale interventiegroep.

Tabel 6.3 - Scores op de totale LanguageScreen en subschalen: verschil tussen deelnemers en niet-deelnemers aan de interventie voor leerlingen bij gunstigere uitvoeringskwaliteit en voor de totale interventiegroep (n=119)

	Groep met optimalere implementatie			Totale interventiegroep		
	est.	s.e.	p	est.	s.e.	p
LanguageScreen (totaal)	10.045	2.582	0.000	2.701	0.760	0.000
Expressieve woordenschat	6.672	3.001	0.027	1.633	0.897	0.070
Receptieve woordenschat	7.406	4.493	0.100	2.539	1.313	0.054
Luisterbegrip	15.796	4.009	0.000	4.499	1.174	0.000
Zinnen herhalen	9.648	3.282	0.003	2.341	0.976	0.017

10 Alle residuen die werden gevormd bij stap 1 van de analyse, op die bij receptieve woordenschat na hingen in stap 2 van de analyse significant negatief samen met de uitkomst. Dit ondersteunt de keuze voor het gekozen type analyse waarin werd gecorrigeerd voor endogeniteit: niet-gemeten factoren die zowel deelname als de uitkomst beïnvloeden.

7 Resultaten aanvullende taaltaken

In dit hoofdstuk beschrijven we eerst de ontwikkeling van de interventieleerlingen op de aanvullend afgenomen taaltaken en vergelijken we deze met de ontwikkeling op de LanguageScreen (7.1). Vervolgens gaan we in op verschillen naar achtergrondkenmerken (7.2) en op de relatie tussen ontwikkeling en deelname aan de interventie, namelijk aanwezigheid en betrokkenheid hierbij (7.3).

7.1 Ontwikkeling per taaltaak

Wat de aanvullende taaltaken betreft, zijn in de periode van de voor- naar de nameting (gemiddeld 5.4 maanden) kinderen het meest vooruitgegaan in Zinnen Begrijpen (0.49 standaarddeviatie), gevolgd door Actieve Woordenschat (0.46 standaarddeviatie), zie Tabel 7.1. Wat het aspect semantiek van de Actie Platen Test (APT) betreft (hoe goed kinderen in staat zijn om betekenisvolle informatie uit te drukken passend bij een afbeelding) was de toename ongeveer even groot als bij Zinnen Begrijpen, maar bleek dit door de hogere standaardfout niet significant. De ontwikkeling in de LanguageScreen is groter, maar ook daar is de toename in Luisterbegrip het sterkst.

Tabel 7.1 - Ontwikkeling van de interventieleerlingen in de aanvullende taaltaken en scores op de LanguageScreen

	Est.	std.error	p
Taaltaak			
Actieve Woordenschat taak	**0.462**	0.097	0.000
APT semantiek	0.459	0.287	0.149
APT grammatica	0.260	0.196	0.215
Zinnen Begrijpen	**0.485**	0.096	0.000
Busverhaal	0.357	0.279	0.233
LanguageScreen			
Totaalscore	**0.876**	0.063	0.000
Expressieve woordenschat	**0.612**	0.066	0.000
Receptieve woordenschat	**0.632**	0.079	0.000
Luisterbegrip	**0.857**	0.088	0.000
Zinnen herhalen	**0.802**	0.081	0.000

7.2 Ontwikkelingen naar achtergrondkenmerken

Voor de aanvullende taaltaken laten de analyses zien dat leerlingen die thuis niet uitsluitend Nederlands spreken, bij de voormeting 0.80 standaarddeviatie lager scoren op de Actieve Woordenschat Taak (AWT) en 0.39 standaarddeviatie lager op Zinnen Begrijpen dan leerlingen met uitsluitend Nederlands als thuistaal (zie Tabel 7.2; uitgebreidere resultaten in bijlage IV). Voor de overige aanvullende taaltaken is er geen significante relatie met thuistaal. Leeftijd bij de voormeting hangt voor alle aanvullende taken positief samen met het niveau: per extra jaar leeftijd liggen de beginscores tussen 0.20 en 0.30 standaarddeviatie hoger (AWT 0.24, APT Semantiek 0.26, APT Grammatica 0.20, Zinnen Begrijpen 0.30 en Busverhaal 0.20). Er zijn geen verschillen naar sekse, onderwijssoort (sbo/so) en externaliserend probleemgedrag in beginniveau op de aanvullende taken.

De groei in score tussen de voor- en nameting op de aanvullende taaltaken verschilde niet significant naar sekse, thuistaal, onderwijssoort, leeftijd en externaliserend probleemgedrag; de interacties tussen tijd en deze achtergrondkenmerken zijn niet significant.
De resultaten voor de LanguageScreen laten een vergelijkbaar beeld zien, met wat meer verschillen in ontwikkeling naar achtergrondkenmerken. Leerlingen die thuis niet uitsluitend Nederlands spreken, scoren 0.71 standaarddeviatie lager op de totaalscore van de LanguageScreen. Dit hangt vooral samen met een lagere expressieve en receptieve woordenschat. Ook hier is een consistent positieve relatie tussen beginniveau en leeftijd: per extra jaar leeftijd liggen de scores 0.44 standaarddeviatie hoger (voor de subtests tussen de 0.25 en 0.42). Wat sekse betreft is er alleen bij de subtest expressieve woordenschat een significant verschil: meisjes scoren gemiddeld 0.42 standaarddeviatie lager dan jongens. Net als bij de aanvullende taaltaken is er geen verschil naar onderwijssoort en externaliserend probleemgedrag in het beginniveau op de aanvullende taken.
De groei in de LanguageScreen verschilt tussen so en sbo: leerlingen in het sbo gingen 0.59 standaarddeviatie meer vooruit op de LanguageScreen dan leerlingen het so. Daarnaast neemt de groei in de subtest zinnen herhalen van de LanguageScreen iets af met de leeftijd: ongeveer –0.19 standaarddeviatie per jaar. Er zijn geen verschillen naar sekse, thuistaal en externaliserend probleemgedrag in de mate van groei in LanguageScreenscores.

Tabel 7.2 - Ontwikkeling taaltaken en LanguageScreen naar achtergrondkenmerken (gestandaardiseerde scores)

	AWT	APT semantiek	APT grammatica	Zinnen Begrijpen	Busverhaal
(Intercept)	**0.483**	-0.088	-0.076	0.005	-0.279
nameting (tijd)	**0.566**	0.537	0.367	**0.376**	0.409
meisje (ref. jongen)	-0.161	0.049	-0.036	-0.139	-0.047
thuistaal niet alleen NL (ref alleen NL)	**-0.798**	-0.170	-0.062	**-0.392**	0.177
sbo (ref = so)	-0.070	0.337	0.247	0.471	0.381
leeftijd voormeting	**0.242**	**0.259**	**0.196**	**0.297**	**0.199**
externaliserend probleemgedrag (gecentreerd)	-0.037	0.016	0.044	-0.137	-0.027
tijd*meisje	-0.103	-0.063	-0.014	-0.068	-0.170
tijd*thuistaal	-0.021	-0.011	-0.053	0.268	-0.077
tijd*sbo	-0.004	-0.068	-0.129	-0.015	0.045
tijd*leeftijd voormeting	0.075	-0.047	0.017	-0.053	-0.133
tijd*externaliserend probleemgedrag	0.137	0.044	0.034	0.095	-0.068
	LanguageScreen (totaal)	**Expressieve Woordenschat**	**Luisterbegrip**	**Receptieve woordenschat**	**Zinnen herhalen**
(Intercept)	0.344	**0.648**	-0.035	0.320	0.076
nameting (tijd)	**0.688**	**0.324**	**0.789**	**0.468**	**0.658**
meisje (ref. jongen)	-0.216	**-0.419**	0.026	-0.363	0.106
thuistaal niet alleen NL (ref alleen NL)	**-0.713**	**-0.962**	-0.291	**-0.636**	-0.203
sbo (ref = so)	0.161	-0.074	0.291	0.218	0.061
leeftijd voormeting	**0.436**	**0.249**	**0.416**	**0.373**	**0.343**
externaliserend probleemgedrag (gecentreerd)	-0.007	0.046	0.001	-0.007	-0.105
tijd*meisje	0.027	0.112	-0.312	0.126	0.205
tijd*thuistaal	0.065	0.111	-0.036	0.098	0.010
tijd*sbo	**0.592**	**0.493**	**0.637**	0.167	**0.632**
tijd*leeftijd voormeting	**-0.115**	0.005	-0.136	-0.088	**-0.187**
tijd*externaliserend probleemgedrag	-0.039	-0.128	-0.007	-0.022	0.081

7.3 Ontwikkeling in relatie tot betrokkenheid en aanwezigheid

Vervolgens is onderzocht in hoeverre de aanwezigheid en betrokkenheid van leerlingen bij de interventieactiviteiten samenhingen met hun taalontwikkeling. De resultaten staan in Tabel 3 (zie bijlage V voor de volledige resultaten). Voor de aanvullende taaltaken werd geen significante relatie gevonden. Voor de LanguageScreen hing betrokkenheid wel significant samen met het niveau van taalvaardigheid: per punt hoger op de betrokkenheidsschaal (bereik 1-5), scoorden kinderen gemiddeld 0.30 standaarddeviatie hoger gescoord op de LanguageScreen (totaalscore). Dit was vooral gerelateerd aan het onderdeel luisterbegrip,

waar per punt hoger in betrokkenheid de score 0.32 standaarddeviatie hoger was. Betrokkenheid hing echter niet samen met de groei in scores op de LanguageScreen-schalen. Aanwezigheid hing niet samen met het niveau van de LanguageScreen-scores, maar wel met groei in expressieve woordenschat: kinderen die vaker aanwezig waren (per punt hoger op de vijfpuntsschaal) gingen 0.32 standaarddeviatie meer vooruit. Voor de andere LanguageScreen-(sub)schalen werd geen relatie tussen de aanwezigheid en groei gevonden.

Tabel 7.3 - Ontwikkeling taaltaken en LanguageScreen in relatie tot de mate van aanwezigheid en betrokkenheid bij de interventie na controle voor leeftijd, onderwijssoort en sekse (gestandaardiseerde scores)

		aanwezigheid			betrokkenheid		
		est.	**s.e.**	**p**	**est.**	**s.e.**	**p**
Actieve Woordenschat	niveau	0.094	0.204	0.647	0.280	0.169	0.098
	groei	-0.084	0.162	0.604	0.074	0.141	0.602
APT semantiek	niveau	-0.189	0.194	0.332	0.329	0.173	0.060
	groei	0.025	0.207	0.905	-0.218	0.213	0.311
APT grammatica	niveau	-0.144	0.218	0.510	0.415	0.212	0.059
	groei	0.083	0.206	0.689	-0.081	0.202	0.689
Zinnen Begrijpen	niveau	0.194	0.190	0.309	0.081	0.157	0.608
	groei	-0.068	0.168	0.686	0.043	0.158	0.787
Busverhaal	niveau	-0.121	0.201	0.548	0.361	0.189	0.059
	groei	0.408	0.251	0.109	-0.238	0.256	0.359
LanguageScreen (totaal)	niveau	0.026	0.169	0.880	**0.304**	0.136	0.027
	groei	0.187	0.120	0.119	0.199	0.106	0.062
Expressieve woordenschat	niveau	-0.078	0.195	0.691	0.255	0.164	0.121
	groei	**0.323**	0.123	0.009	0.135	0.109	0.217
Luisterbegrip	niveau	0.024	0.165	0.886	**0.323**	0.135	0.017
	groei	0.211	0.174	0.226	0.099	0.155	0.522
Receptieve woordenschat	niveau	0.030	0.184	0.871	0.235	0.155	0.130
	groei	0.142	0.154	0.358	0.215	0.137	0.118
Zinnen herhalen	niveau	0.183	0.180	0.311	0.141	0.143	0.327
	groei	-0.138	0.160	0.390	0.230	0.142	0.107

8 Haalbaarheid van een Randomized Controlled Trial in het s(b)o

Vraag in dit onderzoek is ook of/hoe het programma middels een Randomized Controlled Trial (RCT) kan worden onderzocht. In dit hoofdstuk formuleren we op basis van de bevindingen in het sbo en het so de belangrijkste randvoorwaarden voor de uitvoering van een dergelijk RCT-onderzoek. Daarbij gaan we in op randvoorwaarden op schoolniveau, eisen aan implementatiekwaliteit en ontwerpkeuzes.

8.1 Randvoorwaarden op schoolniveau

De beschrijvende resultaten laten zien dat de uitvoerbaarheid van NELI afhankelijk is van personele en organisatorische randvoorwaarden. Voor een RCT-onderzoek is het noodzakelijk dat deze randvoorwaarden vooraf worden geadresseerd.

Voldoende leerlingen die in aanmerking komen.
Voor het uitvoeren van een RCT is het belangrijk dat er voldoende leerlingen zijn die aan de inclusiecriteria voldoen. Voorafgaand aan het onderzoek was onzeker of er, mede door latere instroom door verwijzing vanuit het regulier onderwijs, voldoende leerlingen met de vereiste leeftijd en lage taalprestaties beschikbaar zouden zijn voor deelname aan NELI en of er voldoende kinderen zouden zijn om een vergelijkingsgroep mee te kunnen vormen. Vooraf werd als aanpassing van het programma aan so en sbo het leeftijdscriterium verhoogd. Hierdoor kwamen meer leerlingen in aanmerking voor deelname. Daarnaast bleek dat van alle gescreende leerlingen de meerderheid, namelijk ongeveer tweederde in aanmerking kwam voor deelname. Zeker in het sbo met ruim 260 sbo-vestigingen, is naar verwachting het leerlingenaantal voldoende hoog om een betrouwbare RCT naar de effectiviteit van NELI uit te kunnen voeren.

Personele bezetting
Een belangrijke voorwaarde is voldoende personele bezetting op de deelnemende scholen. Uit de groepsgesprekken blijkt dat het werken met een vaste tutor, die zich specifiek richt op NELI, als essentieel wordt gezien. Idealiter verzorgt één tutor zowel de groepssessies als de individuele sessies, en heeft deze persoon de ruimte om de voortgang zorgvuldig te registreren. Personeelstekorten – bijvoorbeeld doordat de tutor elders in de school wordt ingezet – vormen in de huidige praktijk een belangrijke reden dat sessies uitvallen en het programma niet volledig kan worden afgerond. Voor een RCT-onderzoek moet vooraf geborgd zijn dat tutoren voldoende uren hebben voor uitvoering, voorbereiding en registratie, en dat vervanging geregeld is bij ziekte of uitval. Vaak waren er op scholen

meerdere tutoren die samen de begeleiding verzorgden, bijvoorbeeld door de sessies van verschillende dagen te verdelen. Dit maakt de uitvoering minder kwetsbaar voor uitval, maar vergt tegelijkertijd een goede overdracht tussen tutoren.

Ruimte en tijd
Daarnaast is de beschikbaarheid van een rustige, vaste ruimte een belangrijke randvoorwaarde. Scholen -in het so vaker dan in het sbo- geven aan dat het ontbreken van een aparte NELI-ruimte leidt tot verstoringen en ad-hoc oplossingen (gebruik van gangen of hoekjes in de klas), die de kwaliteit van de sessies ondermijnen. Voor een RCT-opzet moet vaststaan dat deelnemende scholen gedurende de gehele interventieperiode kunnen beschikken over een geschikte ruimte.
Wat betreft tijd is het van belang dat de planning van de 20 weken NELI realistisch is. In het huidige onderzoek verwachtte twee derde van de tutoren het programma niet volledig te kunnen afronden en komen scholen gemiddeld tot circa vier vijfde van de geplande sessies. Enerzijds heeft dit te maken met organisatorische aspecten (vakanties, studiedagen en andere ondersteuningstrajecten), anderzijds met een niet optimale inhoudelijke aansluiting van het programma bij leerlingen in het gespecialiseerd onderwijs (zie 8.2).

Draagvlak en inbedding
Draagvlak en inbedding in de school zijn belangrijke voorwaarden voor een goede uitvoering. Over het geheel genomen ervoeren betrokkenen meer pluspunten dan knelpunten; de motivatie om NELI te laten slagen en het enthousiasme van leerlingen waren groot, vooral in het sbo. Tegelijkertijd zijn er de al genoemde organisatorische randvoorwaarden, zoals voldoende personeel, geschikte ruimtes en tijdsinvestering, als bepalend voor de haalbaarheid genoemd.
Een RCT-onderzoek vraagt daarom om:
- actieve betrokkenheid van de schoolleiding, die tijd en ruimte vrijmaakt voor voorbereiding, uitvoering en registratie en meedenkt over de organisatorische inbedding;
- draagvlak in het team, waarbij afspraken worden gemaakt over aanwezigheid van NELI-leerlingen, het vermijden van concurrerende activiteiten tijdens sessies en;
- structurele ondersteuning door de aanbieder, in de vorm van toegankelijke training, heldere materialen, digitale registratiemogelijkheden en een laagdrempelige helpdesk, zoals door tutoren in dit onderzoek positief is ervaren.

8.2 Inhoudelijke aansluiting

De resultaten wijzen erop dat NELI niet voor alle leerlingen even passend is. NELI is primair ontwikkeld voor kleuters; betrokkenen geven aan dat het programma het best uitvoerbaar is in de kleutergroepen, met een mogelijke uitloop naar het begin van groep 3. Tegelijkertijd worden er, met name in het so, groepen leerlingen genoemd voor wie het programma in de huidige vorm (te) moeilijk of is, bijvoorbeeld vanwege cognitieve of motorische beperkingen of een beperkte aandachtsspanne. Dus naast dat het programma qua leeftijd niet altijd aansluit bij een oudere leeftijdsgroep voor wie een dergelijk programma wel wenselijk is, zou in het programma meer ruimte voor herhaling van geleerde woorden moeten worden geboden. Wil een RCT haalbaar zijn, dan is het nodig dat er meer ruimte komt voor het herhalen van geleerde woorden. Dit kan door bijvoorbeeld het programma

meer te spreiden en meer dan 20 weken hiervoor te nemen, of door een selectie van onderdelen aan te bieden. Kortom, voor een RCT is het wenselijk om inhoudelijke aanpassingen te doen in thema's, woordkeuze en tempo, zodat het programma geschikter wordt voor oudere leerlingen en leerlingen met complexere problematiek, zodat de inhoud niet als te 'kleuterachtig' of te compact (met te weinig ruimte voor herhaling) wordt ervaren.
Deze aanpassingen dienen vooraf te worden vastgelegd, en niet ad-hoc tijdens de RCT, om vergelijkbaarheid tussen scholen en condities te waarborgen.

8.3 Implementatiekwaliteit en monitoring

De analyses in dit rapport laten zien dat de mate waarin NELI volgens plan wordt uitgevoerd sterk samenhangt met de effecten op de LanguageScreen. Scholen verschillen zowel in de hoeveelheid (aantal en duur van sessies) als in de ervaren kwaliteit van uitvoering. Voor een subgroep leerlingen op scholen met gunstigere uitvoeringskwaliteit – voldoende sessies, hogere kwaliteitsbeoordeling, hoge aanwezigheid en betrokkenheid – worden grotere positievere effecten gevonden dan voor de totale interventiegroep. Dit onderstreept dat implementatiekwaliteit een cruciale determinant is van effectiviteit.
Voor een RCT-onderzoek betekent dit dat implementatiefidelity systematisch moet worden gemonitord. De tutorvragenlijsten en aanwezigheidsregistraties uit het huidige onderzoek bieden hiervoor een praktisch instrumentarium. Door deze monitoringsgegevens te koppelen aan de uitkomstmaten kan in een RCT niet alleen het gemiddelde effect worden geschat, maar ook de relatie tussen implementatiekwaliteit en effectiviteit, zoals in de CACE-analyses in dit rapport is ingezet.

8.4 Ontwerpkeuzes voor een RCT in het s(b)o

Uit de praktijkbeschrijvingen en effectanalyses komt naar voren dat leerkrachten en tutoren soms onderdelen of didactische principes uit NELI overnemen in hun reguliere onderwijs. Zij noemen bijvoorbeeld dat zij strategieën of handvatten die voor bepaalde leerlingen werken, doelwoorden en narratieve elementen ook in de klas gebruiken, of dat zij dit graag meer zouden willen doen. Daarmee ontstaat de mogelijkheid dat leerlingen die niet aan NELI deelnemen, indirect toch met elementen van het programma in aanraking komen (spillover-effect). Wanneer deze spillover in sterke mate optreedt, kan dit de kans verkleinen om duidelijke verschillen tussen interventie- en controlegroep te vinden, omdat leerlingen in de controlegroep indirect profiteren van aanpassingen in het klassenaanbod. In dat geval wordt het gevonden effect van NELI waarschijnlijk onderschat. Het is daarom belangrijk om in de opzet van een RCT expliciet rekening te houden met mogelijke spillover: door de omvang ervan goed in kaart te brengen en/of hier actief op te sturen. Dit kan bijvoorbeeld door afspraken te maken over het (tijdelijk) beperken van het gebruik van specifieke NELI-elementen in de reguliere les op controlescholen, of door te kiezen voor randomisatie op schoolniveau in plaats van binnen klassen. Een nadeel van laatstgenoemde optie is dat er dan meer scholen nodig zijn om voldoende statistische power te behalen. Gegeven de hogere leerlingenaantallen en de minder heterogene populatie in het sbo, vergeleken met het so, is een RCT daar methodologisch én praktisch het meest haalbaar. Daarnaast laten de huidige effectanalyses zien dat de effecten van NELI-S in het sbo duidelijker zijn dan in het so, wat het aannemelijk maakt dat een verschil tussen

interventie- en controlegroep in een RCT bij voldoende implementatiekwaliteit detecteerbaar is.
De LanguageScreen is een bruikbaar screeningsinstrument gebleken, aangezien de uitkomsten in dezelfde richting lagen als die van de taaltaken. Een kanttekening is echter dat het instrument relatief oppervlakkig blijft. De CELF bijvoorbeeld geeft een breed profiel van taalvaardigheid, waarmee veel nauwkeuriger inzicht kan worden verkregen in specifieke taalvaardigheden, zoals expressieve woordenschat en receptief taalbegrip. Een storytellingtaak (zoals het Busverhaal) levert rijkere en meer ecologisch valide informatie op dan meer oppervlakkige taalmaten, omdat zichtbaar wordt hoe taalvaardigheden geïntegreerd worden gebruikt in een betekenisvolle context. Het nadeel van deze testes is dat ze omvangrijk zijn en daardoor minder goed inpasbaar in een testbatterij binnen een RCT.

8.5 Belangrijkste inzichten

Op basis van de bevindingen uit dit onderzoek kan worden geconcludeerd dat een RCT-onderzoek naar de effectiviteit van NELI-S in het sbo in principe mogelijk is, mits aan een aantal randvoorwaarden wordt voldaan. Deze randvoorwaarden betreffen voldoende personele capaciteit, stabiele organisatorische condities (tijd, ruimte, planning), een goede inhoudelijke aansluiting op de doelgroep, draagvlak in schoolteams en systematische monitoring van de implementatiekwaliteit. Veel van deze condities bleken in de deelnemende sbo-scholen relatief gunstig, zo was de motivatie om aan NELI deel te nemen groot. Belangrijk aandachtspunt is wel de aansluiting bij de wat oudere leerlingen die ook in het sbo nog in de kleutergroepen aanwezig zijn, en voor wie een taalstimuleringsprogramma even wenselijk is als voor jongere kleuters. Het materiaal bleek onvoldoende goed aan te sluiten bij deze leeftijdsgroep; het ligt daarom voor de hand om in een RCT vooraf na te denken over passende aanpassingen in thema's, tempo en moeilijkheidsgraad.
In het so, waar vaker niet werd voldaan aan organisatorische randvoorwaarden (personeel, ruimte, voorspelbare roosters) en waar sprake is van kleinere aantallen leerlingen per doelgroep en grotere diversiteit in problematiek, ligt een RCT minder voor de hand dan in het sbo.

9 Samenvatting en slotbeschouwing

9.1 Samenvatting

In dit onderzoek evalueerden we het NELI-programma voor kleuters in het speciaal (basis)onderwijs (sbo en so clusters 3 en 4). NELI is een gestructureerde, in Engeland ontwikkelde intensieve taalinterventie gericht op het versterken van de woordenschat en mondelinge taalvaardigheid van kleuters. NELI is in Nederland in het regulier onderwijs onderzocht (door middel van een RCT-studie), maar was nog niet in het gespecialiseerd onderwijs toegepast. Dit onderzoek richt zich daarom specifiek op NELI in het gespecialiseerd onderwijs, op (1) de effecten van NELI op taal en sociaal-emotioneel functioneren en voor welke leerlingen de interventie het meest effectief is, (2) verklarende en werkzame factoren, (3) de relatie tussen implementatiekwaliteit en effectiviteit, (4) haalbaarheid en implementeerbaarheid, en (5) de vraag of een RCT-design ook in het gespecialiseerd onderwijs uitvoerbaar zou zijn.
Het onderzoek is uitgevoerd op 23 scholen (sbo en so, clusters 3 en 4) met een kwantitatieve effectmeting met een interventie- en een vergelijkingsgroep, vragenlijsten onder tutoren en aanvullende groepsgesprekken op tien scholen. Taalvaardigheid is primair gemeten met de LanguageScreen (twee meetmomenten) en daarnaast met vier aanvullende taaltaken (actieve woordenschat, vertelvaardigheid, omschrijven van plaatjes en zinsbegrip) bij alleen de interventieleerlingen. Sociaal-emotionele uitkomsten (bijv. probleemgedrag, opkomen voor zichzelf, rekening houden met anderen) zijn eveneens op twee momenten gemeten.

(1) Effecten van NELI. Leerlingen in de interventiegroep gingen gemiddeld meer vooruit in taalvaardigheid (gemeten met de LanguageScreen) dan leerlingen in de vergelijkingsgroep, wat het sterkst het geval was in het sbo en bij jongere leerlingen (4–7 jaar). Ook uit de tutorvragenlijst en groepsgesprekken komt naar voren dat de waargenomen opbrengsten in het sbo doorgaans duidelijker zijn dan in het so.
Daarnaast hing NELI-deelname samen met een gunstigere ontwikkeling in de mate waarin leerlingen voor zichzelf opkomen en rekening houden met anderen, zoals beoordeeld door leerkrachten. In leerkrachtoordelen over werkhouding, impulscontrole, internaliserend en externaliserend probleemgedrag en emotioneel reactief gedrag vonden we geen verschil.

(2) Werkzame en verklarende factoren. Het NELI-programma is vooral gericht op het ondersteunen van luister- en vertelvaardigheden en woordenschat. Dit zien we terug in de resultaten op afzonderlijke taaldomeinen: de toename was het sterkst in luistervaardigheid en betrof verder zinnen herhalen en expressieve woordenschat. Meer aanwezigheid bij de groeps- en individuele sessies was bovendien gerelateerd aan een grotere toename in

expressieve woordenschat. Uit de groepsgesprekken en tutoroordelen komt ook naar voren dat vooral narratieve vaardigheden en productieve woordenschat verbeteren. Mediatieanalyses zijn uitgevoerd om na te gaan of veranderingen in gedrag (bijvoorbeeld verbetering in sociaal gedrag of werkhouding) de effecten van NELI op taalvaardigheid lijken te verklaren, dan wel of juist taalontwikkeling fungeert als mediator voor veranderingen in gedrag en sociaal-emotioneel functioneren. In lijn met eerdere internationale studies op NELI worden hiervoor in dit onderzoek geen duidelijke aanwijzingen gevonden: positieve effecten op taal en eventuele verbeteringen in gedrag en zelfvertrouwen blijken grotendeels onafhankelijk van elkaar op te treden, en niet als elkaars mediërende schakel. Hoewel de indirecte effecten niet significant zijn, wijzen de effectgroottes er eerder op dat verbeterde taalvaardigheid samenhangt met gedragsverbetering dan dat gedragsverandering een verklaring vormt voor de effectiviteit van NELI op taalvaardigheid. Dit wordt ondersteund door tutoroordelen: tutoren signaleerden dat leerlingen mondiger worden en meer zelfvertrouwen tonen, wat wordt gekoppeld aan hun gegroeide woordenschat.

(3) Relatie tussen implementatiekwaliteit en effectiviteit. Scholen verschillen in de mate van interventietrouw, de gerealiseerde intensiteit en de betrokkenheid en aanwezigheid van leerlingen. Op basis van tutorvragenlijsten is een kwantitatieve indicator geconstrueerd voor zowel de hoeveelheid (aantal en duur van sessies) als de ervaren kwaliteit van uitvoering; gemiddeld scoren scholen hoog, maar het programma wordt lang niet overal volledig afgerond en veel scholen komen in de praktijk tot circa vier vijfde van het programma. Om na te gaan wat NELI onder gunstige uitvoeringsomstandigheden kan opleveren, is een subgroep geselecteerd van leerlingen op scholen met hogere uitvoeringskwaliteit (voldoende sessies, hogere kwaliteitsbeoordeling, hoge aanwezigheid en betrokkenheid). Voor deze subgroep worden grotere positieve effecten op de LanguageScreen gevonden dan voor de totale interventiegroep. Dit wijst erop dat de uitvoeringskwaliteit een cruciale rol speelt in de mate van effectiviteit: het potentieel van NELI komt pas volledig tot zijn recht wanneer het programma volledig en volgens de richtlijnen wordt uitgevoerd. Met andere woorden, NELI kan duidelijke effecten sorteren, mits het programma conform de instructies en onder gunstige randvoorwaarden wordt toegepast.

(4) Haalbaarheid en implementeerbaarheid. NELI blijkt in het sbo en so implementeerbaar en haalbaar, met een aantal kanttekeningen. De duidelijke structuur en het houvast dat NELI daardoor biedt, wordt gewaardeerd. Tutoren voelen zich overwegend goed voorbereid na de training en begeleiding door de ontwikkelaar. Tegelijkertijd blijken tijdsdruk, personeelskrapte, planning rond vakanties en de beschikbaarheid van geschikte ruimtes belangrijke belemmeringen. Veel scholen behalen niet het volledige programma, maar komen gemiddeld tot circa vier vijfde van de geplande sessies. De ervaren pluspunten – zoals enthousiasme van leerlingen, zichtbare taalprogressie en professionele opbrengsten voor tutoren – wegen voor veel scholen zwaarder dan de knelpunten, maar de structurele inbedding vraagt om organisatorische aanpassingen. Genoemde structurele uitvoeringsknelpunten bleken in het so nadrukkelijker aanwezig, wat mogelijk de lagere effectiviteit daar kan verklaren. Daarnaast hebben leerlingen in het so gemiddeld complexere gedrags- en/of cognitieve problematiek, waardoor het groepsaanbod minder gemakkelijk bleek te “landen” en sommige onderdelen (met name narratief) minder goed uitvoerbaar bleken. Ook is de leeftijd van de kinderen in het so gemiddeld hoger; met name

voor oudere kinderen werd het programma qua taalgebruik (niveau en aansluiting) minder geschikt gevonden en soms aangepast. Daarnaast biedt het programma in de huidige vorm met name voor kinderen met cognitieve beperkingen te weinig ruimte voor herhaling.

(5) Uitvoerbaarheid RCT-design in het gespecialiseerd onderwijs. Tot slot wijzen de resultaten op mogelijkheden voor verdere optimalisatie van het programma en voor een toekomstige RCT-studie in het sbo en so. Op basis van de bevindingen kan worden geconcludeerd dat een RCT-studie naar de effectiviteit van NELI-S in het sbo in principe haalbaar is, maar dat inhoudelijke aanpassingen wenselijk zijn. Met name de thema's, woordkeuze en tempo verdienen aanpassing, zodat het programma beter aansluit bij oudere leerlingen en bij leerlingen met complexere problematiek. Hierdoor wordt voorkomen dat de inhoud als te 'kleuterachtig' of juist te compact (met te weinig ruimte voor herhaling) wordt ervaren.

9.2 Slotbeschouwing

De uitkomsten uit dit onderzoek kunnen we vergelijken met de uitkomsten uit de evaluatiestudie naar de effecten van NELI bij kinderen in het reguliere onderwijs (Van Driel et al., 2025) en studies naar de effecten van NELI in Groot-Brittannië (Hulme et al., 2025; Fricke et al. 2013). Bij een vergelijking tussen de verschilscores voor en na de NELI interventie tussen de drie studies, zien we dat de kinderen uit ons onderzoek minder sterk vooruit zijn gegaan dan de kinderen in het reguliere basisonderwijs in Nederland, met name op het gebied van betekenis uitdrukken en grammaticale complexiteit binnen een verteltaak. Bij de resultaten op de Actie Platen Test (zowel semantiek als grammatica) was er ongeveer 50% minder groei in het gemiddelde op de nameting van de interventiegroep. Ook de gemiddelde uitingslengte was minder sterk gegroeid dan gemiddeld genomen in het reguliere po, al was daar de groei meer vergelijkbaar (ongeveer 10% minder lange uitingen na de interventie bij de kinderen in het s(b)o vergeleken met het po). Ook vergeleken met de Engelse studies boekten de kinderen van ons onderzoek een minder grote vooruitgang op deze drie maten. Daarbij zagen we in ons onderzoek, dat kinderen in het sbo sterker vooruitgaan dan kinderen in het so.

Voor deze verschillen in de mate van voortuitgang gemeten na afloop van de interventie zijn meerdere verklaringen mogelijk.
Een factor is de *duur van de interventie*: de studies verschilden hierin: in Groot-Brittannië duurde NELI 30 weken, in plaats van 20 weken in Nederland.
Daarnaast speelt de combinatie van *leeftijd van de kinderen en type onderwijs* een rol. Met name het verschil in leeftijd, waarbij de leerlingpopulatie in het sbo en so ouder is dan de populatie in het reguliere basisonderwijs, en de context van het type onderwijs van de leerlingpopulatie lijkt belangrijk om het verschil in groei op de aanvullende productieve taalmaten (Actie Platen Test en Bus) te duiden. Het NELI programma is ontworpen voor jongere kinderen, waardoor een grotere groei in taalvaardigheid te verwachten is bij het jongere kind. De leerlingen in het so zijn daarnaast doorgaans ouder dan leerlingen in het sbo, wat verklaart dat we in het so de minste groei constateren. De populatie kinderen die gespecialiseerd onderwijs nodig heeft, heeft verder wellicht meer herhaling nodig dan kinderen in het reguliere PO om tot een grotere groei te kunnen komen in hun productieve taalvaardigheid.

Verder hebben we gezien dat wanneer de *interventie optimaal wordt uitgevoerd*, de effecten aanzienlijk groter zijn dan het gemiddelde effect in de totale interventiegroep. De toename in de score op de LanguageScreen is voor NELI-deelnemers bij optimalere uitvoeringskwaliteit ongeveer één standaarddeviatie hoger dan voor niet-deelnemers, wat aanzienlijk hoger is dan het gemiddelde effect voor de totale interventiegroep. Voor de subtest luisterbegrip is het effect zelfs groter.

Ook op de sociaal-emotionele ontwikkeling verwachtten we positieve effecten van NELI. Vooraf werd verondersteld dat NELI gunstige omstandigheden zou creëren voor het gedrag van leerlingen, door de kleinschalige setting, de nadruk op aandacht en luisteren, de succeservaringen en de versterkte relatie tussen tutor en kind. Dit wordt in dit onderzoek deels bevestigd. Leerlingen die aan NELI deelnamen, hadden een gunstigere ontwikkeling in opkomen voor zichzelf en rekening houden met anderen, twee domeinen die nauw aansluiten bij de sociale interactie die NELI stimuleert. De mediatieanalyses gaven geen aanwijzingen dat dit effect liep via een verbetering in taalvaardigheid; het lijkt daarmee om een zelfstandig effect van NELI te gaan. We vonden dus aanwijzingen voor effectiviteit op de meer sociale aspecten van gedrag, die direct worden aangesproken in de tutorsetting: leerlingen oefenen in de kleine groep met beurtnemen, luisteren naar anderen en het verwoorden van eigen gedachten. In de sociaal-emotionele uitkomstmaten werkhouding, impulscontrole, internaliserend en externaliserend probleemgedrag en emotioneel reactief gedrag gingen NELI-deelnemers gemiddeld niet meer vooruit dan de vergelijkingsgroep. Dit betreft meer gedragsregulerende aspecten die waarschijnlijk minder gevoelig zijn voor een interventie van deze omvang en duur, zeker in een populatie met vaak complexere gedrags- en ontwikkelingsproblematiek dan in het reguliere basisonderwijs. Dat de Engelse studie (West e.a., 2022) wel bredere gedragseffecten vond, kan samenhangen met het feit dat daar met een geaggregeerde gedragsmaat werd gewerkt, waarbinnen de sociale componenten mogelijk de doorslag gaven. De bevindingen uit de interviews nuanceren dit beeld verder: tutoren beschrijven zichtbare groei in mondigheid, zelfvertrouwen en sociaal initiatief bij individuele leerlingen, maar benadrukken tegelijkertijd dat deze ontwikkeling niet altijd exclusief aan NELI is toe te schrijven, omdat leerlingen doorgaans ook andere begeleiding ontvangen.

We concluderen dat er alle reden is om in het speciaal (basis)onderwijs in te zetten op taalondersteuning, met behulp van NELI of vergelijkbare interventies. Het is daarbij wel nodig dat een dergelijk programma goed aansluit bij de leeftijd en ontwikkelfase van leerlingen in het gespecialiseerd onderwijs. NELI was oorspronkelijk bedoeld voor kleuters in het reguliere basisonderwijs en sloot voor een deel van de leerlingen in het s(b)o qua leeftijd en tempo minder goed aan; een deel van de leerlingen had duidelijk baat bij meer tijd en herhaling. Voor NELI zijn aanpassingen in leeftijdsspecifieke thema's, voorbeelden en taalniveau, gecombineerd met ruimte voor vertraging en extra herhaling, nodig om het voor jongere en oudere leerlingen in het s(b)o passend en effectief te maken. Tegelijkertijd lijkt een programma als NELI qua organisatie goed aan te sluiten bij de context van het sbo en so, waar (vanwege specifieke zorgbehoeften van de kinderen) vaker individueel of in kleinere groepen wordt gewerkt en waar wat meer flexibiliteit mogelijk is.
Daarnaast is het nodig om veel explicieter te maken onder welke condities een programma werkt. De potentie van de interventie komt alleen tot uiting wanneer de implementatie voldoende stevig is georganiseerd. Dit betekent dat, ondanks de relatief beperkte investering in middelen (training en materialen), succesvolle toepassing in het sbo en so

afhankelijk is van de mate waarin scholen erin slagen de randvoorwaarden, met name op het gebied van organisatie en personele inzet, goed te realiseren. Onder deze gunstige condities wegen de baten van het programma op tegen de benodigde kosten (zowel in middelen als in organisatie-inspanning).

Literatuur

Beck, I. L., McKeown, M. G., Kucan, L. (2010). *Veel gestelde vragen over woordenschatonderwijs: volgens de aanpak van 'lezend woorden lezen'*. Helmond: Onderwijs Maak Je Samen.

Chi, W. E., Huang, S., Jeon, M., Park, E. S., Melguizo, T., & Kezar, A. (2022). A practical guide to causal mediation analysis: Illustration with a comprehensive college transition program and nonprogram peer and faculty interactions. *Frontiers in Education, 7*, 886722. https://doi.org/10.3389/feduc.2022.886722

Curtis, P. R., Estabrook, R., Kaiser, A. P., & Roberts, M. Y. (2019). The Longitudinal Effects of Early Language Intervention on Children's Problem Behaviors. *Child Development*, *90*(2), 576–592. https://doi.org/10.1111/cdev.12942

Driel, S. van, Meeter, M. & Wilt, F. van der (2025). *Het effect van de Nederlandse versie van de Nuffield Early Language Intervention (NELI) op de taalvaardigheid van jonge kinderen.* Amsterdam: Vrije Universiteit.

Fricke, S., Bowyer-Crane, C., Haley, A. J., Hulme, C., & Snowling, M. J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry*, 54(3), 280–290. https://doi.org/10.1111/jcpp.12010

Hargreaves, A. (2000). Mixed emotions: Teachers' perceptions of their interactions with students. *Teaching and teacher Education*, 16 (8), 811-826.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms Through Which Influential Early Childhood Program Boosted Adult Outcomes. *The American Economic Review*, *103*(6), 2052–2086. https://doi.org/10.1257/aer.103.6.2052

Henrichs, L. F. (2010). *Academic language in early childhood interactions: a longitudinal study of 3- to 6-year-old Dutch monolingual children*. [s.n.].

Hulme, C., McGrane, J., Duta, M., West, G., Cripps, D., Dasgupta, A., Snowling, M. J. (2024). LanguageScreen: The Development, Validation, and Standardization of an Automated Language Assessment App. *Language, Speech & Hearing Services in Schools*, *55*(3), 1–917. https://doi.org/10.1044/2024_LSHSS-24-00004

Hulme, C., West, G., Rios Diaz, M., Hearne, S., Korell, C., Duta, M., & Snowling, M. J. (2025). The Nuffield Early Language Intervention (NELI) programme is associated with lasting improvements in children's language and reading skills. *Journal of Child Psychology and Psychiatry*, *66*(9), 1357-1365.

Jansonius, K., Ketelaars, M., Borgers, M., Van Den Heuvel, E., Roeyers, H., Manders, E., & Zink, I. (2014). Renfrew Taalschalen Nederlandse Aanpassing. Garant: Antwerpen/ Apeldoorn.

Lee-Pang, L., Levickis, P., Murray, L., McFarland, L., & Quach, J. (2025). Applying Conceptualisations of Child Well-being to Early Childhood Education and Care: A Scoping Review of the Literature. *Early Childhood Education Journal*. https://doi.org/10.1007/s10643-025-01940-9

Leseman, P., & Veen, A. (red.). (2022). Het pre-COOL cohort tot en met groep 8: Ontwikkeling van kinderen en relatie met kwaliteit in de voor- en vroegschoolse periode (Rapport 1080). Amsterdam: Kohnstamm Instituut.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference. 2nd Edition.* Cambridge: Cambridge University Press.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: basic psychological needs in motivation, development, and wellness*. New York: Guilford Press.

Sibieta, L., Kotecha, M., & Skipp, A. (2016). Nuffield Early Language Intervention: Evaluation Report and Executive Summary. In *Education Endowment Foundation*. Education Endowment Foundation.

Siraj-Blatchford, I., & Sylva, K. (2004). Researching pedagogy in English pre-schools. *British Educational Research Journal*, *30*(5), 713–730.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531–543. https://doi.org/10.1016/j.jhealeco.2007.09.009

Van Driel, S., Meeter, M., & Van der Wilt, F. (2025). *Het effect van de Nederlandse versie van de Nuffield Early Language Intervention – NELI – op de taalvaardigheid van jonge kinderen.* Amsterdam: Vrije Universiteit van Amsterdam.

West, G., Snowling, M. J., Lervåg, A., Buchanan-Worster, E., Duta, M., Hall, A., McLachlan, H., & Hulme, C. (2021). Early language screening and intervention can be delivered successfully at scale: Evidence from a cluster randomized controlled trial. *Journal of Child Psychology and Psychiatry, 62*(12), 1425–1434. https://doi.org/10.1111/jcpp.13415

West, G., Lervåg, A., Snowling, M. J., Buchanan-Worster, E., Duta, M., & Hulme, C. (2022). Early language intervention improves behavioral adjustment in school: Evidence from a cluster randomized trial. *Journal of School Psychology*, *92*, 334–345. https://doi.org/10.1016/j.jsp.2022.04.006

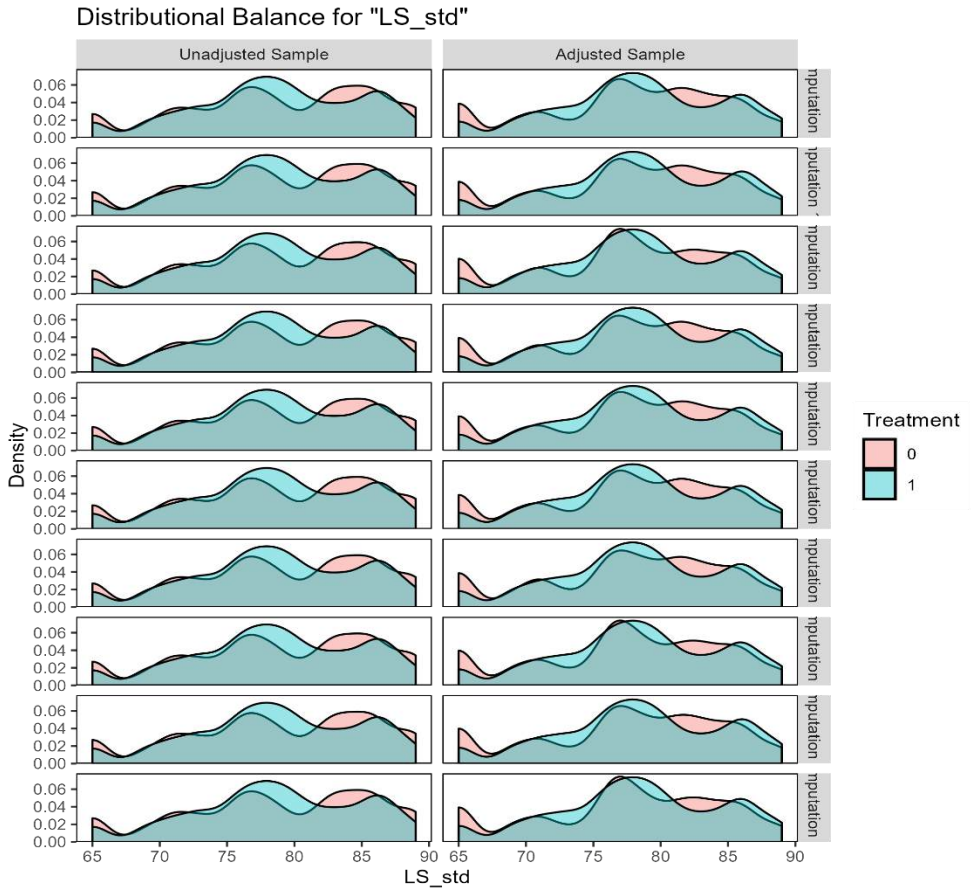
Whittle, R., Mansell, G., Jellema, P., & van der Windt, D. (2017). Applying causal mediation methods to clinical trial data: What can we learn about why our interventions (don't) work? *European Journal of Pain, 21*(4), 614–622. https://doi.org/10.1002/ejp.964

Wiig, E. H., Secord, W. A., Semel, E., & De Jong, J. (2012). CELF preschool-2-NL: clinical evaluation of language fundamentals: preschool-Nederlandstalige versie. Pierson.

Bijlagen

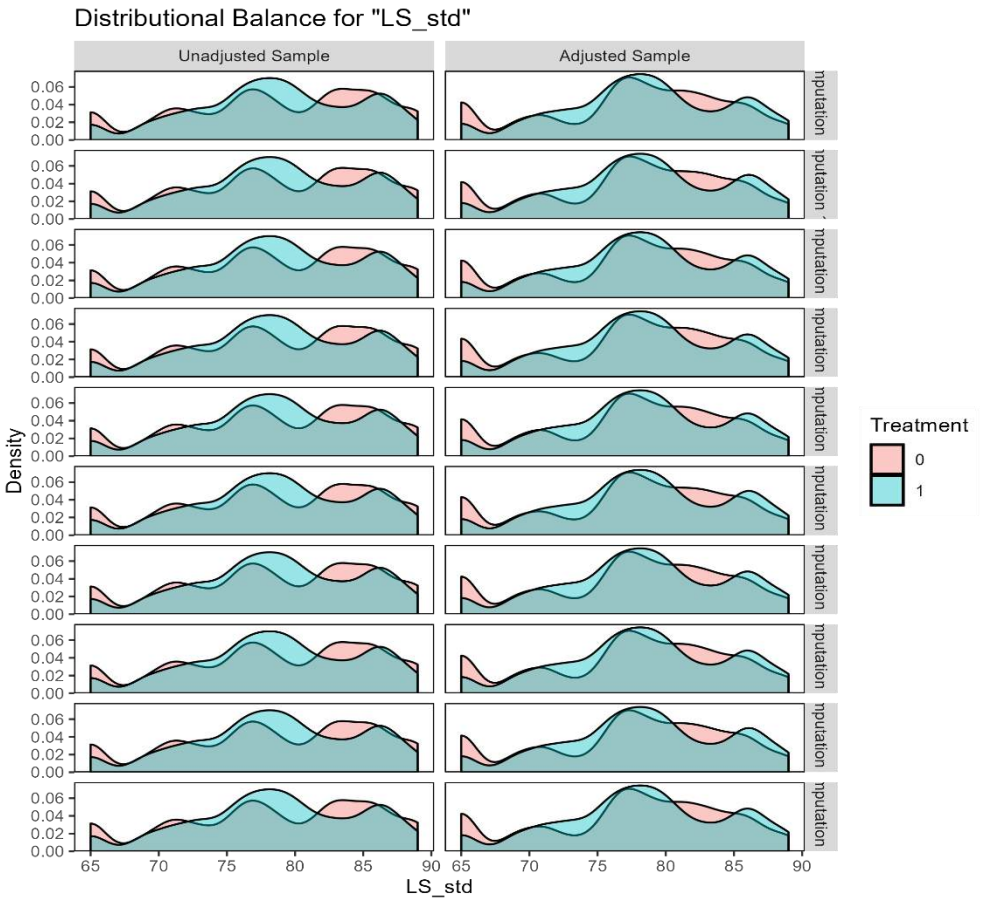
Bijlage I Aanvullende resultaten matching ten behoeve van LanguageScreen data

Vergelijking score-verdeling op de LanguageScreen voorafgaand en na matching voor de tien geïmputeerde datasets (dichtheidsplots). Dit betreft de matching ten behoeve van de vergelijking van de scores op de LanguageScreen



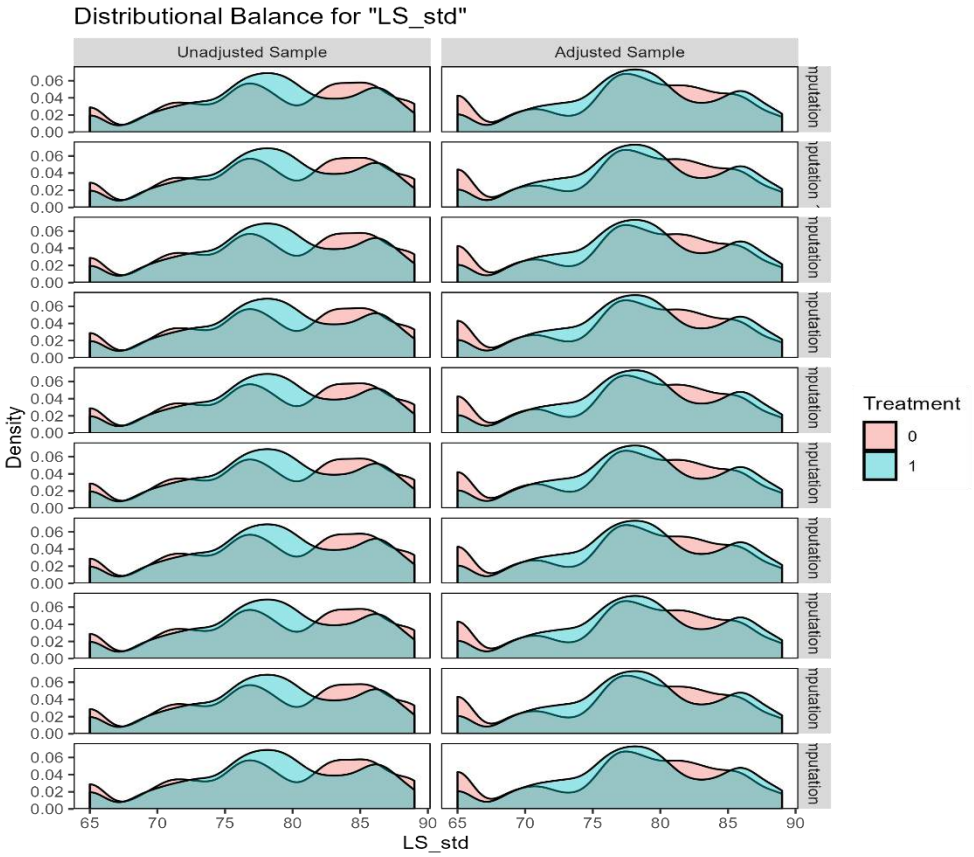
Bijlage II Resultaten matching ten behoeve van leerlingprofieldata

Bij de matching konden 7 leerlingen uit de interventiegroep niet worden gematcht met een vergelijkingsgroep-leerling, en 15 leerlingen uit de vergelijkingsgroep niet met een interventiegroep-leerling. Na weging is de effectieve steekproefgrootte (ESS) 112 leerlingen in de interventiegroep en 84 leerlingen in de vergelijkingsgroep. Door de matching zijn na weging de aandelen meisjes, leerlingen met een niet-Nederlandse thuistaal, leerlingen afkomstig uit sbo/so en met een ondersteuningsbehoefte op het gebied van externaliserend gedrag exact gelijk. De matching resulteerde dus in volledige balans op geslacht, thuistaal, schooltype (sbo/so) en ondersteuningsbehoefte met betrekking tot externaliserend gedrag. Wat betreft de vergelijking tussen de interventie- en vergelijkingsgroep met betrekking tot de LanguageScreen score zijn de resultaten als volgt. Voorafgaand aan matching was het gemiddelde verschil in score minimaal en niet significant (-0.011), na matching nog steeds heel klein en niet significant (0.12).
De gemiddelde afstand tussen de gematchte paren op de LanguageScreen bedraagt 0.403, wat acceptabel is (vuistregel: waarden < 0.25 zeer goed, 0.25-0.50 acceptabel).
De spreiding blijft na matching vergelijkbaar tussen beide groepen: de variantie ratio vóór matching was 0.767 en na matching 0.802 (de waarde 1 is optimaal). De verdeling van LanguageScreen scores was voorafgaand aan de matching al vergelijkbaar (*empirical cumulative distribution function* (e*CDF*) gemiddeld 0.037 en maximaal 0.099), na matching nog iets meer vergelijkbaar (gem. 0.027 en maximaal 0.075). Waarden onder 0.1 worden doorgaans gezien als teken van goede balans. Hieronder is een grafische weergave van de verdelingen weergegeven in de vorm van een dichtheidsplot voor de tien datasets.



Bijlage III Resultaten matching ten behoeve van mediatie-analyse

Bij de matching konden 7 leerlingen uit de interventiegroep niet worden gematcht met een vergelijkingsgroep-leerling, en 17 leerlingen uit de vergelijkingsgroep niet met een interventiegroep-leerling. Na weging is de effectieve steekproefgrootte (ESS) 114 leerlingen in de interventiegroep en 87 leerlingen in de vergelijkingsgroep. Door de matching zijn na weging de aandelen meisjes, leerlingen met een niet-Nederlandse thuistaal, leerlingen afkomstig uit sbo/so en met een ondersteuningsbehoefte op het gebied van externaliserend gedrag exact gelijk. De matching resulteerde dus in volledige balans op geslacht, thuistaal, schooltype (sbo/so) en ondersteuningsbehoefte met betrekking tot externaliserend gedrag. Wat betreft de vergelijking tussen de interventie- en vergelijkingsgroep met betrekking tot de LanguageScreen score zijn de resultaten als volgt. Voorafgaand aan matching was het gemiddelde verschil in score minimaal en niet significant (-0.024), na matching nog steeds heel klein en niet significant (0.10).
De gemiddelde afstand tussen de gematchte paren op de LanguageScreen bedraagt 0.390, wat acceptabel is (vuistregel: waarden < 0.25 zeer goed, 0.25-0.50 acceptabel).
De spreiding blijft na matching vergelijkbaar tussen beide groepen: de variantie ratio vóór matching was 0.808 en na matching 0.818 (de waarde 1 is optimaal). De verdeling van LanguageScreen scores was voorafgaand aan de matching al vergelijkbaar (*empirical cumulative distribution function* (e*CDF*) gemiddeld 0.035 en maximaal 0.102), na matching nog iets meer vergelijkbaar (gem. 0.024 en maximaal 0.068). Waarden onder 0.1 worden doorgaans gezien als teken van goede balans. Hieronder is een grafische weergave van de verdelingen weergegeven in de vorm van een dichtheidsplot voor de tien datasets.



Bijlage IV – Interventiekinderen: ontwikkeling taaltaken en LanguageScreen naar achtergrondkenmerken (waarden vetgedrukt wanneer p < .05)

	AWT			APT semantiek			APT grammatica			Zinnen Begrijpen			Busverhaal		
	est.	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**
(Intercept)	**0.483**	(0.199)	**0.016**	-0.088	(0.232)	0.705	-0.076	(0.219)	0.730	0.005	(0.209)	0.982	-0.279	(0.203)	0.172
nameting (tijd)	**0.566**	(0.193)	**0.005**	0.537	(0.456)	0.266	0.367	(0.328)	0.286	**0.376**	(0.169)	**0.029**	0.409	(0.425)	0.357
meisje (ref. jongen)	-0.161	(0.223)	0.472	0.049	(0.214)	0.820	-0.036	(0.234)	0.879	-0.139	(0.195)	0.476	-0.047	(0.263)	0.860
thuistaal niet alleen NL (ref alleen NL)	**-0.798**	(0.203)	**0.000**	-0.170	(0.221)	0.446	-0.062	(0.208)	0.768	**-0.392**	(0.183)	**0.034**	0.177	(0.262)	0.506
sbo (ref = so)	-0.070	(0.227)	0.758	0.337	(0.267)	0.209	0.247	(0.264)	0.352	0.471	(0.250)	0.062	0.381	(0.218)	0.083
leeftijd voormeting	**0.242**	(0.083)	**0.004**	**0.259**	(0.096)	**0.009**	**0.196**	(0.093)	**0.038**	**0.297**	(0.078)	**0.000**	**0.199**	(0.087)	**0.026**
externaliserend probleemgedrag (gecentreerd)	-0.037	(0.112)	0.743	0.016	(0.110)	0.886	0.044	(0.115)	0.701	-0.137	(0.098)	0.163	-0.027	(0.113)	0.813
tijd*meisje	-0.103	(0.195)	0.598	-0.063	(0.253)	0.805	-0.014	(0.302)	0.964	-0.068	(0.195)	0.727	-0.170	(0.300)	0.575
tijd*thuistaal	-0.021	(0.181)	0.907	-0.011	(0.274)	0.968	-0.053	(0.219)	0.809	0.268	(0.174)	0.127	-0.077	(0.275)	0.780
tijd*sbo	-0.004	(0.204)	0.985	-0.068	(0.253)	0.790	-0.129	(0.267)	0.632	-0.015	(0.192)	0.939	0.045	(0.354)	0.901
tijd*leeftijd voormeting	0.075	(0.078)	0.336	-0.047	(0.102)	0.650	0.017	(0.101)	0.865	-0.053	(0.071)	0.456	-0.133	(0.144)	0.371
tijd*externaliserend probleemgedrag	0.137	(0.098)	0.167	0.044	(0.129)	0.736	0.034	(0.118)	0.776	0.095	(0.096)	0.323	-0.068	(0.163)	0.681

	LanguageScreen (totaal)			**Expressieve woordenschat**			**Luisterbegrip**			**Receptieve woordenschat**			**Zinnen herhalen**		
	est.	s.e.	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**
(Intercept)	0.344	(0.219)	0.118	**0.648**	(0.189)	**0.001**	-0.035	(0.220)	0.873	0.320	(0.190)	0.094	0.076	(0.286)	0.792
nameting (tijd)	**0.688**	(0.127)	**0.000**	**0.324**	(0.131)	**0.014**	**0.789**	(0.185)	**0.000**	**0.468**	(0.155)	**0.003**	**0.658**	(0.191)	**0.001**
meisje (ref. jongen)	-0.216	(0.205)	0.295	**-0.419**	(0.211)	**0.049**	0.026	(0.209)	0.900	-0.363	(0.195)	0.064	0.106	(0.251)	0.672
thuistaal niet alleen NL (ref alleen NL)	**-0.713**	(0.187)	**0.000**	**-0.962**	(0.194)	**0.000**	-0.291	(0.190)	0.128	**-0.636**	(0.178)	**0.000**	-0.203	(0.227)	0.373
sbo (ref = so)	0.161	(0.260)	0.537	-0.074	(0.212)	0.726	0.291	(0.261)	0.266	0.218	(0.218)	0.318	0.061	(0.347)	0.860
leeftijd voormeting	**0.436**	(0.080)	**0.000**	**0.249**	(0.077)	**0.001**	**0.416**	(0.081)	**0.000**	**0.373**	(0.073)	**0.000**	**0.343**	(0.099)	**0.001**
externaliserend probleemgedrag (gecentreerd)	-0.007	(0.105)	0.946	0.046	(0.107)	0.667	0.001	(0.106)	0.992	-0.007	(0.099)	0.940	-0.105	(0.127)	0.413
tijd*meisje	0.027	(0.150)	0.857	0.112	(0.153)	0.466	-0.312	(0.217)	0.153	0.126	(0.180)	0.487	0.205	(0.224)	0.361
tijd*thuistaal	0.065	(0.136)	0.635	0.111	(0.140)	0.428	-0.036	(0.198)	0.855	0.098	(0.166)	0.554	0.010	(0.205)	0.961
tijd*sbo	**0.592**	(0.141)	**0.000**	**0.493**	(0.145)	**0.001**	**0.637**	(0.205)	**0.002**	0.167	(0.172)	0.333	**0.632**	(0.212)	**0.003**
tijd*leeftijd voormeting	**-0.115**	(0.053)	**0.032**	0.005	(0.055)	0.934	-0.136	(0.078)	0.080	-0.088	(0.065)	0.178	**-0.187**	(0.080)	**0.021**
tijd*externaliserend probleemgedrag	-0.039	(0.075)	0.602	-0.128	(0.077)	0.099	-0.007	(0.109)	0.947	-0.022	(0.091)	0.808	0.081	(0.113)	0.475

Bijlage V – Interventiekinderen: ontwikkeling taaltaken en LanguageScreen naar aanwezigheid en betrokkenheid bij deelname (waarden vetgedrukt wanneer p < .05)

	AWT			APT semantiek			APT grammatica			Zinnen Begrijpen			Busverhaal		
	est.	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**
(Intercept)	-0.044	(0.174)	0.801	-0.337	(0.236)	0.162	-0.093	(0.216)	0.668	**-0.381**	(0.187)	**0.043**	-0.301	(0.183)	0.105
nameting (tijd)	**0.463**	(0.098)	**0.000**	0.460	(0.287)	0.149	0.260	(0.197)	0.217	**0.485**	(0.097)	**0.000**	0.357	(0.279)	0.233
betrokkenheid (gec.)	0.280	(0.169)	0.098	0.329	(0.173)	0.060	0.415	(0.212)	0.059	0.081	(0.157)	0.608	0.361	(0.189)	0.059
aanwezigheid (gec.)	0.094	(0.204)	0.647	-0.189	(0.194)	0.332	-0.144	(0.218)	0.510	0.194	(0.190)	0.309	-0.121	(0.201)	0.548
sbo (ref = so)	-0.193	(0.224)	0.390	0.269	(0.249)	0.282	0.059	(0.272)	0.828	0.419	(0.241)	0.084	0.330	(0.243)	0.181
meisje (ref = jongen)	-0.185	(0.202)	0.361	-0.024	(0.177)	0.891	-0.082	(0.196)	0.676	-0.184	(0.180)	0.311	-0.161	(0.200)	0.424
leeftijd voormeting	**0.202**	(0.073)	**0.006**	**0.204**	(0.072)	**0.006**	**0.176**	(0.080)	**0.030**	**0.269**	(0.076)	**0.001**	0.149	(0.089)	0.105
tijd*betrokkenheid	0.074	(0.141)	0.602	-0.218	(0.213)	0.311	-0.081	(0.202)	0.689	0.043	(0.158)	0.787	-0.238	(0.256)	0.359
tijd*aanwezigheid	-0.084	(0.162)	0.604	0.025	(0.207)	0.905	0.083	(0.206)	0.689	-0.068	(0.168)	0.686	0.408	(0.251)	0.109

	LanguageScreen (totaal)			Expressieve woordenschat			Luisterbegrip			Receptieve woordenschat			Zinnen herhalen		
	est.	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**	**est.**	**s.e.**	**p**
(Intercept)	**-0.475**	(0.166)	**0.005**	-0.215	(0.167)	0.199	**-0.571**	(0.156)	**0.000**	-0.287	(0.153)	0.062	**-0.498**	(0.193)	**0.011**
nameting (tijd)	**0.873**	(0.061)	**0.000**	**0.608**	(0.062)	**0.000**	**0.855**	(0.088)	**0.000**	**0.629**	(0.078)	**0.000**	**0.801**	(0.081)	**0.000**
betrokkenheid (gec.)	**0.304**	(0.136)	**0.027**	0.255	(0.164)	0.121	**0.323**	(0.135)	**0.017**	0.235	(0.155)	0.130	0.141	(0.143)	0.327
aanwezigheid (gec.)	0.026	(0.169)	0.880	-0.078	(0.195)	0.691	0.024	(0.165)	0.886	0.030	(0.184)	0.871	0.183	(0.180)	0.311
sbo (ref = so)	0.207	(0.217)	0.341	0.013	(0.216)	0.953	0.337	(0.200)	0.094	0.145	(0.194)	0.455	0.161	(0.254)	0.528
meisje (ref = jongen)	-0.199	(0.167)	0.235	-0.291	(0.198)	0.144	-0.143	(0.146)	0.329	-0.321	(0.179)	0.074	0.138	(0.166)	0.406
leeftijd voormeting	**0.273**	(0.064)	**0.000**	**0.177**	(0.072)	**0.015**	**0.265**	(0.057)	**0.000**	**0.279**	(0.065)	**0.000**	**0.169**	(0.067)	**0.012**
tijd*betrokkenheid	0.199	(0.106)	0.062	0.135	(0.109)	0.217	0.099	(0.155)	0.522	0.215	(0.137)	0.118	0.230	(0.142)	0.107
tijd*aanwezigheid	0.187	(0.120)	0.119	**0.323**	(0.123)	**0.009**	0.211	(0.174)	0.226	0.142	(0.154)	0.358	-0.138	(0.160)	0.390