

# Bouwstenen van leesvaardigheid in het voortgezet onderwijs

Een *umbrella review* naar kenmerken van leesonderwijs en aspecten van leesgedrag en -houding die bijdragen aan het leesbegrip van leerlingen in het vo

Pim Woldendorp, Lesya Ganushchak en Roel van Steensel  
Erasmus Universiteit Rotterdam, augustus 2025

## Colofon

Titel: Bouwstenen van leesvaardigheid in het voortgezet onderwijs: Een  
umbrella review naar kenmerken van leesonderwijs en aspecten van  
leesgedrag en -houding die bijdragen aan het leesbegrip van leerlingen  
in het vo

Auteurs: Pim Woldendorp, Lesya Ganushchak en Roel van Steensel

Uitgevende instantie: Erasmus Universiteit Rotterdam

Jaar van uitgave: 2025

Opdrachtgever: Inspectie van het Onderwijs

Mogelijk gemaakt met financiële steun van het Nationaal Regieorgaan Onderwijsonderzoek.

## Inhoud

Colofon .....	2
Inleiding .....	5
Domeinspecifieke kenmerken van effectief leesonderwijs .....	5
Leesgedrag en -houding .....	8
Huidige reviewstudie .....	9
Methode.....	10
Zoekopdrachten .....	10
Selectie van studies.....	11
Data-extractie .....	15
Data-analyse.....	16
Resultaten.....	19
Meta-analyses: zich normaal ontwikkelende leerlingen .....	25
Kennis en vaardigheden.....	25
Leesgedrag en -houding.....	37
Kernmerken van de leeromgeving.....	42
Combinatie van variabelen .....	43
Meta-analyses: leerlingen met speciale onderwijsbehoeften.....	45
Kennis en vaardigheden.....	45
Kenmerken van de leeromgeving.....	46
Combinatie van variabelen .....	47
Systematische reviews: zich normaal ontwikkelende leerlingen .....	50
Kennis en vaardigheden.....	50
Leesgedrag en -houding.....	53
Kenmerken van de leeromgeving .....	54
Combinatie van variabelen .....	56
Systematische reviews: leerlingen met speciale onderwijsbehoeften.....	57
Kennis en vaardigheden.....	57
Kenmerken van de leeromgeving.....	61
Combinatie van variabelen .....	63

Gecombineerde meta-analyses en systematische reviews: leerlingen met speciale onderwijsbehoeften.....	67
Kennis en vaardigheden.....	67
Kenmerken van de leeromgeving.....	68
Overzichtstabel.....	69
Conceptuele overzichtsstudies .....	71
Discussie .....	76
Onderzoeksvraag 1: kenmerken van leesonderwijs, leeshouding en leesgedrag .....	77
Variabelen met systematische ondersteuning voor relaties/effecten .....	78
Variabelen met ondersteuning voor relaties/effecten voor leerlingen met speciale behoeften, maar niet voor zich normaal ontwikkelende leerlingen .....	78
Variabelen met ondersteuning voor relaties/effecten die maar voor een van beide doelgroepen zijn onderzocht .....	79
Variabelen zonder ondersteuning voor relaties/effecten of met wisselende ondersteuning .....	81
Variabelen waarnaar nog maar weinig literatuurstudies zijn gedaan .....	83
Consequenties voor het peilingsonderzoek.....	83
Onderzoeksvraag 2: definitie van concepten op het gebied van leesgedrag en -houding .....	85
Lacunes in de review.....	85
Conclusie .....	86
Referenties.....	87
Bijlage 1: Zoektermen per database.....	104
Bijlage 2: Codeerschema voor reviews van interventies .....	107
Bijlage 3: Codeerschema voor conceptuele studies .....	112
Bijlage 3: Bewijskracht en effecten per studie.....	115

## Inleiding

Lezen is een cruciale vaardigheid voor het schoolsucces en de maatschappelijke participatie van jongeren. Het stelt leerlingen in staat om informatie kritisch te verwerken, betekenis te construeren, verbanden te leggen tussen teksten en contexten, en actief deel te nemen aan een kennisintensieve samenleving (Kendeou et al., 2014). In dit licht hebben de Inspectie van het Onderwijs en het Nationaal Regieorgaan Onderwijsonderzoek (NRO) de Erasmus Universiteit Rotterdam opdracht gegeven tot een reviewstudie ter voorbereiding op een landelijke peiling naar leesvaardigheid aan het einde van leerjaar 2 van het voortgezet onderwijs (vo). De peiling bestrijkt alle onderwijssoorten binnen het vo: praktijkonderwijs, vmbo (bb, kb, gl/tl), havo en vwo.

De reviewstudie had als doel te bepalen welke factoren relevant zijn om mee te nemen in deze peiling. Deze factoren moeten inzicht geven in het onderwijsleerproces (in termen van schoolbeleid, onderwijsaanbod en/of lespraktijk) en in het leesgedrag en de leeshouding van leerlingen en leraren. Op basis van de uitkomsten moesten de volgende drie concrete producten worden opgeleverd:

1. Aanbevelingen voor het opnemen van factoren in vragenlijsten.
2. Een eventuele prioritering van deze factoren.
3. Suggesties voor factoren die het meest relevant zijn om op te nemen in een trendanalyse voor de periode 2023–2028.

De factoren binnen het onderwijsleerproces omvatten zowel domeinspecifieke aspecten—dat wil zeggen, kenmerken van effectief leesonderwijs (zie verder)—als meer algemene onderwijskundige principes en professionele kwaliteiten van leraren die van belang zijn in elk vakgebied. Het overzicht van deze variabelen moest een praktisch en inhoudelijk onderbouwd kader bieden voor de ontwikkeling van vragenlijsten, observatieprotocollen en andere onderzoeksinstrumenten.

### Domeinspecifieke kenmerken van effectief leesonderwijs

Met betrekking tot het onderwijsleerproces zijn in het Nederlandse taalgebied de afgelopen jaren verschillende evidence-based overzichten gepubliceerd van kenmerken van effectief leesonderwijs; het gaat daarbij met name om kenmerken die bijdragen aan het lezen met begrip (Geudens et al., 2022; Gobyn et al., 2019; Houtveen et al., 2019; Houtveen & Van

Steensel, 2022; Snel, 2024; Van den Broek et al., 2021). Deze richtten zich met uitzondering van Geudens et al. (2022) en Van den Broek et al. (2021) uitsluitend op het basisonderwijs. In twee recente scriptieonderzoeken onder begeleiding van de derde auteur zijn de onderwijskenmerken die in deze kaders worden benoemd, geïntegreerd tot acht kernelementen (Boute, 2024; Timmers, 2024). In de onderstaande tabel uit het onderzoek van Timmers worden deze elementen kort toegelicht.

### Tabel 1

*Overzicht van kernelementen uit recente overzichten van kenmerken van effectief leesonderwijs (naar Timmers, 2024, pp. 8-9)*

<b>Kernelement</b>	<b>Toelichting</b>
1. Monitoren en differentiëren	Aan de hand van (toets)gegevens wordt op effectieve en doelgerichte wijze gewerkt aan de leerprestaties van leerlingen. Leraren analyseren en evalueren voortgangsgegevens van leerlingen en stellen op basis daarvan concrete en uitdagende doelen. Er wordt formatief getoetst: leerlinggegevens zoals toets- en observatiedata worden gebruikt om het leerproces te ondersteunen, bijvoorbeeld door leerlingen inzicht te geven in de toepassing van (lees)strategieën om hun begrip te ondersteunen.
2. Achtergrondkennis	Om een tekst te kunnen begrijpen, heeft een leerling achtergrondkennis nodig. Die achtergrondkennis wordt met tekstuele informatie geïntegreerd in een rijke mentale representatie ( <i>situatiemodel</i> ). Goed begrip van een tekst zorgt ervoor dat er nieuwe kennis wordt toegevoegd aan het geheugen, die ingezet kan worden om volgende teksten te begrijpen. Effectief leesonderwijs integreert lees- en kennisontwikkeling, bijvoorbeeld door thematisch onderwijs.
3. Blootstelling aan en instructie over	Leerlingen moeten weten om te gaan met een grote variëteit aan tekstsoorten, zowel binnen als buiten de klas. Leesbegrip kan worden bevorderd door leerlingen

- verschillende teksttypen en genres teksten van verschillende typen en uit verschillende genres aan te bieden en ze daar kennis over te laten opdoen via tekststructuuronderwijs. Hierbij leren leerlingen onder andere basisstructuren herkennen.
4. Toepassen van leesstrategieën Leesstrategieën zijn cognitieve instrumenten die ingezet kunnen worden om het begrip van een tekst te monitoren, repareren of versterken wanneer het leesproces niet automatisch verloopt. Door middel van strategie-instructie leren leerlingen welke strategieën ze kunnen toepassen, hoe ze dat kunnen doen, waarom en in welke omstandigheden.
5. Interactie en discussiëren Open uitwisseling van ideeën over de tekst draagt bij aan leesbegrip, omdat het de actieve betrokkenheid van leerlingen bevordert en leerlingen geconfronteerd worden met andere interpretaties en perspectieven. Ook kunnen leerlingen zo begrip van de tekst co-construeren. Daarnaast kan discussie onderling zorgen voor het verwerven van kennis. Binnen het onderwijs is het belangrijk dat leerlingen ruimte krijgen om te interacteren en discussiëren over teksten.
6. Integreren van lezen en schrijven Schrijven kan het leesproces ondersteunen, aangezien lezen en schrijven verbonden processen zijn. Schrijven noodzaakt leerlingen hun eigen teksten kritisch te lezen, wat bijdraagt aan tekstbegrip. Dit maakt de integratie van lezen en schrijven binnen het leesonderwijs van toegevoegde waarde.
7. Leesmotivatie Intrinsieke motivatie heeft een positieve relatie met leesvaardigheid. Om de leesmotivatie van leerlingen te stimuleren, is een motiverende leesomgeving van belang. Een motiverende leesomgeving kan gecreëerd worden door in te spelen op de psychologische basisbehoeften van leerlingen: autonomie, competentie en verbondenheid.

8. Lezen in authentieke situaties	Doorgaans lezen mensen om een doel te bereiken, zoals leren voor een toets. In effectief leesonderwijs worden authentieke leessituaties gecreëerd, waarin leerlingen lezen om concrete doelen te behalen.
-----------------------------------	---

---

Eén van de vragen die in de huidige review beantwoord moest worden, was of deze acht elementen één-op-één te vertalen zijn naar het voortgezet onderwijs. Van den Broek et al. (2021) geven weliswaar aan dat de ontwikkeling van lezen met begrip gradueel verloopt en er dus geen sterke scheiding te verwachten is tussen het lezen op de basis- en middelbare school, maar de context in het voortgezet onderwijs is duidelijk anders. Dat komt onder meer door systemische verschillen tussen beide onderwijsvormen—de middelbare school kent een onderverdeling in onderwijstypen (pro, vmbo, havo, vwo) en een duidelijkere scheiding dan het basisonderwijs tussen vakken en vakdocenten—maar ook doordat leerlingen zich in een andere ontwikkelingsfase bevinden: de adolescentie.

### Leesgedrag en -houding

Met betrekking tot het leesgedrag en de leeshouding van leerlingen was het doel om duidelijk te maken hoe verschillende concepten die in dit domein gebruikt worden zich tot elkaar verhouden en welke relatie ze hebben met leesbegrip. Bij leesgedrag gaat het in principe om variabelen als leesfrequentie, leeshoeveelheid of blootstelling aan boeken binnen en buiten school (Mol & Bus, 2011) alsmede om de mate van betrokkenheid tijdens (schoolse) leestaken (Lee et al., 2021). Bij leeshouding waren we geïnteresseerd in de (intrinsieke en extrinsieke) leesmotivatie van leerlingen, hun leesplezier en leeszelfvertrouwen en in de houding van ouders ten opzichte van lezen (Petscher, 2010; Conradi et al., 2014; Schiefele et al., 2012).

Bij het leesgedrag van leraren waren we geïnteresseerd in de relatie tussen het leesbegrip van leerlingen en de eigen leesfrequentie van leraren in de vrije tijd, de activiteiten die zij ondernemen om hun kennis van jeugdliteratuur op peil te houden en zich te professionaliseren in leesonderwijs, evenals de visie die zij hebben op hun rol hierin. Ook wilden we informatie verzamelen over de vakinhoudelijke en didactische kennis van leraren over lezen en het vertrouwen dat ze daarin hebben (Dixon & Oakhill, 2024) alsmede over hun deelname aan bijscholing en het bredere schoolklimaat, met name de prestatiegerichtheid van de school.

## Huidige reviewstudie

In de reviewstudie stonden twee vragen centraal:

1. Welke kenmerken van het leesonderwijs (d.w.z. domeinspecifieke kenmerken en algemene onderwijskundige principes en professionele kwaliteiten van leraren) en welke aspecten van het leesgedrag en de leeshouding van leerlingen en leraren dragen bij aan het leesbegrip van adolescenten/leerlingen in het voortgezet onderwijs?
2. Hoe worden centrale concepten op het gebied van leesgedrag en -houding voor deze doelgroep gedefinieerd en geoperationaliseerd?

Om die onderzoeksvragen te beantwoorden, voerden we een zogenaamde *umbrella review* uit (Pollock et al., 2024). Een umbrella review is geen review van primaire, maar juist van secundaire studies: we vatten de uitkomsten samen van bestaande meta-analyses en systematische reviews. Een belangrijk voordeel van umbrella reviews is dat ze de uitkomsten van individuele overzichtsstudies overstijgen. Met zo'n review kan worden bepaald hoe consistent de antwoorden op bepaalde onderzoeksvragen zijn (Aromataris et al., 2015; Polanin et al., 2017): laten verschillende overzichtsstudies vergelijkbare resultaten zien of verschillen ze juist en, als dat laatste het geval is, waardoor kunnen zulke verschillen worden verklaard? Een ander voordeel is dat een umbrella review in relatief korte tijd kan worden uitgevoerd (Caird et al., 2015). De onderzoeksvragen waren immers breed: een review op basis van primaire studies zou te veel tijd hebben gekost.

De umbrella review bestond uit twee onderdelen, aansluitend bij de beide onderzoeksvragen: (1) een review van overzichtsstudies (d.w.z. meta-analyses en systematische reviews) naar de relatie tussen leesbegrip en kenmerken van het onderwijsleerproces, het leesgedrag en de leeshouding van leerlingen en leraren; en (2) een review van (conceptuele) overzichtsstudies naar definities en operationalisaties van centrale begrippen rond leesgedrag en leeshouding.

## Methode

In dit hoofdstuk beschrijven we hoe we de umbrella review hebben uitgevoerd. We gaan in op de vraag hoe relevante studies zijn gezocht en geselecteerd, en op het proces van data-extractie en -analyse.

### Zoekopdrachten

Op basis van de beide onderzoeksvragen onderscheidde we twee zoekopdrachten: (1) één om studies te vinden waarin is gekeken naar de relatie tussen leesbegrip en kenmerken van het onderwijsleerproces en/of het gedrag en/of de houding van leerlingen en/of leraren; en (2) één om studies te vinden waarin definities en operationalisaties van centrale concepten op het gebied van leesgedrag en -houding zijn uitgewerkt. We voerden beide zoekopdrachten uit in maart 2025 in samenwerking met een ervaren informatiespecialist van de bibliotheek van het Erasmus Medisch Centrum. We deden dat zowel in het Engels als in het Nederlands. In Tabel 2 is aangegeven welke databases we gebruikten voor de verschillende zoekopdrachten. Aangezien de gebruikte databases voor Zoekopdracht 1 vooral Engelstalige wetenschappelijke tijdschriften bevatten, is er een aanvullende Nederlandstalige Google-zoekopdracht uitgevoerd.

**Tabel 2**

*Gebruikte databases per zoekopdracht*

Zoekopdracht 1, Engelstalig	ERIC, PsycINFO, Web of Science, Scopus en Google Scholar
Zoekopdracht 1, Nederlandstalig	Google
Zoekopdracht 2, Engelstalig	Web of Science

In Zoekopdracht 1 werden drie categorieën zoektermen gecombineerd: zoektermen op het gebied van de centrale variabelen (bijv. *reading comprehension*, *reading instruction*, *reading behavior*, *reading attitude*), zoektermen op het gebied van de doelgroep (bijv. *adolescents*, *secondary education*) en zoektermen op het gebied van het type studie (bijv. *meta-analysis*, *systematic review*). Voor Zoekopdracht 2 werden twee categorieën gecombineerd: zoektermen op het gebied van de centrale variabelen (bijv. *reading motivation*, *reading behavior*, *reading engagement*) en zoektermen op het gebied van het type studie (bijv.

*concept, definition*). Bij deze tweede zoekopdracht zijn geen zoektermen op het gebied van de doelgroep gebruikt, omdat conceptuele studies zich (meestal) niet beperken tot een bepaalde doelgroep. De volledige zoekopdrachten zijn te vinden in Bijlage 1.

Op basis van Zoekopdracht 1 (Engelstalig) werden na duplicatie 1.109 unieke studies (uit in totaal 1.989 hits) geïdentificeerd, op basis van Zoekopdracht 2 waren dat er 152. Voor het Nederlandstalige deel van Zoekopdracht 1 zijn geen cijfers beschikbaar; hiervoor werd immers een Google-search uitgevoerd. Om het risico te beperken dat we op basis van onze selectie van zoektermen studies over het hoofd zouden zien, hebben we daarnaast 19 Nederlandse en Vlaamse leesexperts benaderd met de vraag welke overzichtsstudies volgens hen belangrijk waren om op te nemen in de umbrella review. Het ging om leesonderzoekers uit verschillende disciplines (literatuurwetenschap, onderwijstechnologie, onderwijswetenschappen, taalbeheersing, taalkunde en vakdidactiek) en medewerkers van relevante organisaties, zoals lerarenopleidingen, SLO, Stichting Lezen en de Inspectie van het Onderwijs. Van de 19 benaderde experts hebben er negen meegewerkt. Zij stuurden samen 50 titels in, waarvan er negen al uit onze eigen zoekopdrachten waren gekomen. Daarmee kwam het totale aantal unieke studies waaruit geselecteerd kon worden (dus voor beide zoekopdrachten samen) op 1.302.

## Selectie van studies

Voor de selectie van studies ter beantwoording van de eerste onderzoeksvraag maakten we gebruik van deze inclusiecriteria:

1. De studie is een meta-analyse of systematische review. Voor systematische reviews moest er ten minste sprake zijn van een systematische search en selectie van studies.
2. De taal van de publicatie is Engels of Nederlands.
3. De onafhankelijke variabelen liggen op het vlak van het onderwijsleerproces (schoolbeleid, onderwijsaanbod en/of lespraktijk) en/of op het vlak van het leesgedrag en/of de leeshouding van leerlingen en/of leraren.
4. De review presenteert gerichte conclusies voor het leesbegrip (of een verwant concept) van adolescenten/leerlingen in het voortgezet onderwijs.

Daarnaast hanteerden we de volgende exclusiecriteria:

1. De studie is een umbrella review.
2. De review is gericht op lezen in een vreemde taal. Reviews die betrekking hebben op tweede-taalleerders (doorgaans leerlingen met een migratieachtergrond die in het land

waarheen ze zijn verhuisd de taal van dat land leren) namen we overigens wel op, als onderdeel van de studies die zich richtten op speciale doelgroepen (zie verder).

3. Op basis van de review kunnen geen conclusies worden getrokken over het effect van specifieke kenmerken van het onderwijsleerproces op het leesbegrip van de doelgroep. We kwamen bijvoorbeeld studies tegen waarin het gemiddelde effect werd berekend van een breed scala aan interventies, maar waarin niet werd geanalyseerd wat de verschillen in effecten waren tussen die interventies voor vo-leerlingen (bijv. Filderman et al., 2022; Scammacca et al., 2007, 2015).

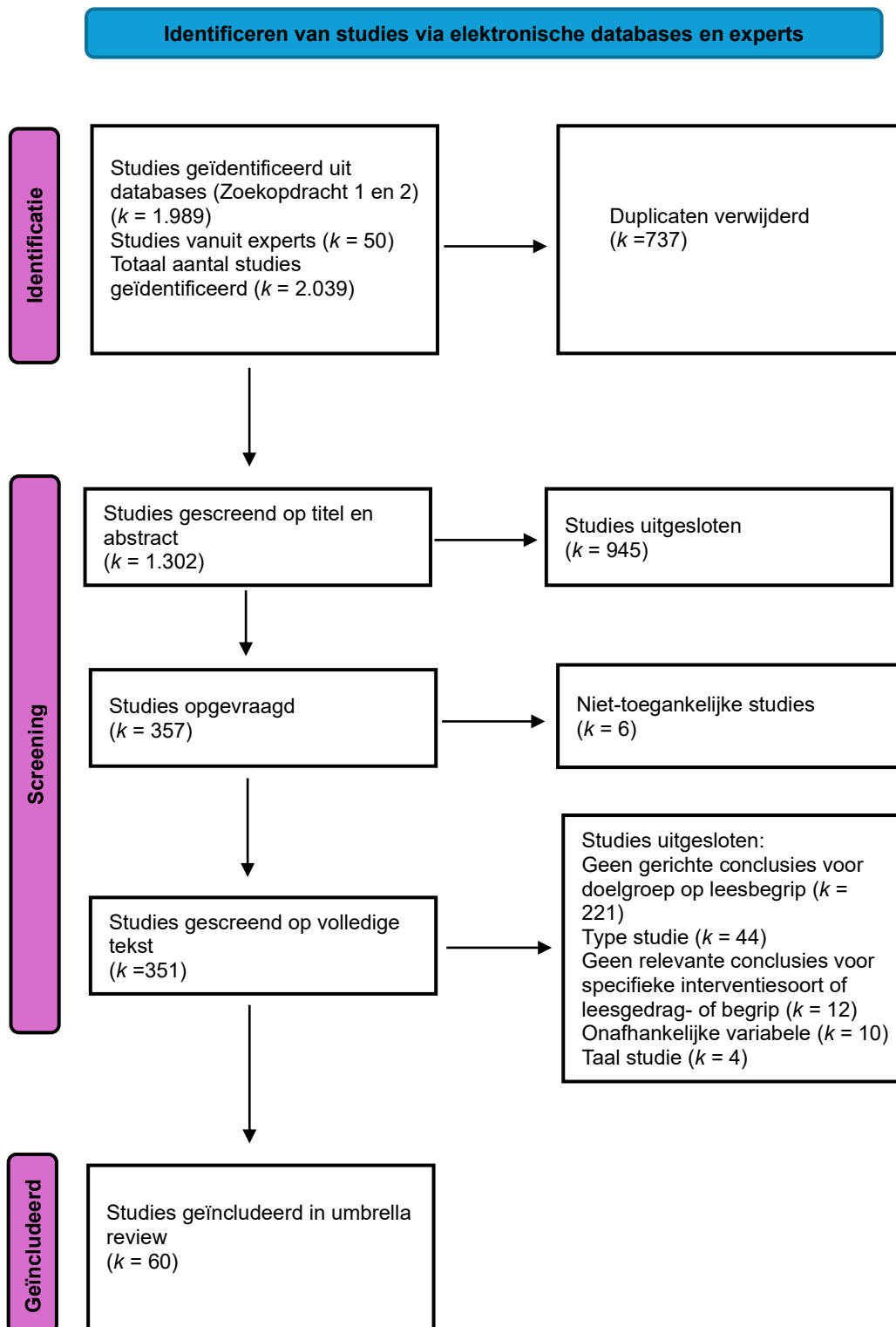
Voor de tweede onderzoeksvraag maakten we gebruik van deze inclusiecriteria:

1. De studie is een conceptuele overzichtsstudie: de studie maakt duidelijk hoe een bepaald centraal begrip wordt gedefinieerd en/of geoperationaliseerd.
2. De taal van de publicatie is Engels of Nederlands.
3. De centrale variabele is leesgedrag of leeshouding (of een verwant begrip).
4. De studie heeft (ook) aandacht voor leerlingen binnen het voortgezet onderwijs of adolescenten.

Alle 1.302 studies zijn op basis van deze criteria door minstens twee van de onderzoekers gescreend op titel en abstract. In een voorbereidende stap zijn 100 willekeurig geselecteerde titels door alle drie de onderzoekers bekeken. Op basis van een bespreking van de verschillen in selectie zijn de in- en exclusiecriteria, waar nodig, aangescherpt. In de eerste selectieronde zijn de resterende titels verdeeld: de eerste auteur bekeek de titels en abstracts van alle publicaties, de tweede en derde auteur deden dat ieder voor de helft van de publicaties. Verschillen in selectie werden opnieuw besproken, op basis waarvan een gezamenlijk besluit over in- of exclusie werd genomen. Voor deze selectieronde was de interbeoordelaarsbetrouwbaarheid (Fleiss's kappa) tussen de .50 en .67, wat op een gematigde tot goede overeenkomst duidt (Landis & Koch, 1977). Na de eerste selectieronde bleven er 357 publicaties over, waarvan in een tweede ronde de volledige tekst werd bekeken; opnieuw gebeurde dat telkens door twee onderzoekers. Een aantal van de opgevraagde publicaties was niet toegankelijk via de universiteitsbibliotheek. Voor deze studies is gekeken of ze bij andere bibliotheken te leen waren of aan te schaffen waren bij de uitgever. Het merendeel van deze studies was zo alsnog in te zien; zes studies waren echter niet te achterhalen. De interbeoordelaarsbetrouwbaarheid voor de tweede selectieronde was tussen de .42 en .56, wat op een gematigde overeenkomst duidt (Landis & Koch, 1977). Dat laatste had te maken met

het feit dat het met een eerste screening van de volledige tekst niet altijd gemakkelijk bleek om na te gaan of er gerichte conclusies voor vo-leerlingen werden getrokken. Weer zijn alle verschillen besproken en is telkens een gezamenlijk besluit genomen over in- of uitsluiting. Na de tweede selectieronde bleven er 134 studies over. Tijdens de data-extractie vielen er nog 74 studies uit, met name omdat ze geen gerichte conclusies rapporteerden voor adolescenten/leerlingen uit het voortgezet onderwijs. De uiteindelijke selectie bestond daarom uit 60 studies. Het volledige selectieproces is weergegeven in Figuur 1.

Figuur 1: PRISMA-diagram voor de selectieprocedure



## Data-extractie

We hebben afzonderlijke codeerschema's ontwikkeld voor de reviews die we analyseerden voor de eerste onderzoeksvraag (Bijlage 2) en voor de tweede onderzoeksvraag (Bijlage 3). Beide codeerschema's zijn gebaseerd op de aanbevelingen voor data-extractie bij umbrella reviews van Aromataris et al. (2015) en het codeerschema van de umbrella review van Geudens et al. (2022). In het eerste codeerschema werd de volgende informatie geregistreerd:

- Bibliografische informatie: titel, auteurs, publicatiejaar en bron.
- Type studie: meta-analyse of systematische review.
- Doel van de studie, onderzoeksvragen en/of hypothesen.
- Design van de primaire studies, bijvoorbeeld: experimenteel, quasi-experimenteel, correlatieel of single-group/-case.
- Kenmerken van de onderzochte populatie, bijvoorbeeld: leeftijd/onderwijsfase, aanwezigheid van lees- of leerproblemen.
- Uitvoering van een *quality appraisal*: controle van de kwaliteit van de primaire studies, al dan niet op basis van een standaardprocedure.
- Analyse van publicatiebias.<sup>1</sup>
- Definitie en operationalisatie van onafhankelijke variabelen (d.w.z. kenmerken van het onderwijsleerproces, leesgedrag en -houding) en afhankelijke variabelen (leesbegrip).
- Overzicht van resultaten, beperkingen en conclusies.

In het tweede codeerschema werden onder andere het doel van de studie, de methode (voor zover gerapporteerd), de gebruikte theorieën, de gegeven definities van het concept en de relatie die de auteurs (wel of niet) legden met leesbegrip beschreven.

Om het eerste codeerschema te valideren, zijn eerst twee studies (een meta-analyse en een systematische review) door alle drie de auteurs gecodeerd. Hun coderingen zijn vergeleken en verschillen in coderingen zijn besproken. Op basis daarvan zijn de codeerschema's aangescherpt en vastgesteld. Vervolgens is de eerste helft van de studies gecodeerd door de eerste auteur. Daarna zijn nogmaals twee studies door alle drie de auteurs gecodeerd om de stabiliteit van de coderingen van de eerste auteur te optimaliseren.

---

<sup>1</sup> Publicatiebias verwijst naar de situatie dat wetenschappelijke tijdschriften vaker studies publiceren die (positieve) relaties/effecten laten zien, waardoor in een meta-analyse of systematische review de gemiddelde relaties/effecten worden overschat.

Vervolgens heeft de eerste auteur de resterende studies gecodeerd. Om het tweede codeerschema te valideren, is één studie door alle drie de auteurs gecodeerd en zijn verschillen besproken. Voor dit codeerschema is er geen tweede gezamenlijke codeerronde uitgevoerd, omdat er uit de search maar vier conceptuele overzichtsstudies kwamen. Vragen en twijfels die tijdens het codeerproces naar voren kwamen, zijn tijdens gezamenlijke overleggen besproken.

## Data-analyse

Op basis van de ingevulde codeerschema's hebben we per meta-analyse, systematische review en conceptuele overzichtsstudie korte beschrijvingen gemaakt die de basis vormden voor het resultatenhoofdstuk. Om de eerste onderzoeksvraag te kunnen beantwoorden, hebben we de betreffende literatuurstudies gecategoriseerd. Allereerst maakten we een onderscheid tussen meta-analyses en systematische reviews en tussen studies die betrekking hebben op zich normaal ontwikkelende leerlingen en leerlingen met speciale onderwijsbehoeften; in het laatste geval gaat het om leerlingen met lees- en/of leerproblemen, stoornissen en beperkingen of om tweede-taalleerders. Binnen deze categorieën hebben we verder onderscheid gemaakt tussen reviews die zich richtten op:

1. **Kennis en vaardigheden.** We onderscheidden zeven subcategorieën: (a) inferenties maken/achtergrondkennis, (b) integreren van meerdere teksten, (c) kennis van tekststructuur, (d) (mondelinge) taalvaardigheid, (e) schrijfvaardigheid, (f) technisch lezen (decoderen, vloeiend lezen) en (g) zelfregulatie, metacognitie en leesstrategieën.
2. **Leesgedrag en -houding.**
3. **Kenmerken van de leeromgeving.** We onderscheidden zes subcategorieën: (a) digitale technologie, (b) *graphic organizers*, (c) *peer mediation*/coöperatief leren, (d) personalisatie, (e) tutoring en (f) vakintegratie.
4. **Kenmerken van de schoolorganisatie.**
5. **Combinaties van variabelen.**

Categorie (1) heeft betrekking op de in de inleiding genoemde domeinspecifieke onderwijskenmerken, terwijl categorie (3) betrekking heeft op meer algemene onderwijskundige principes. De verschillende (sub)categorieën lichten we nader toe in het volgende hoofdstuk.

Vervolgens hebben we per categorie bepaald welke literatuurstudies over het algemeen positieve relaties of effecten lieten zien, welke studies geen of eventueel negatieve

relaties/effecten lieten zien en voor welke studies het beeld gemengd was. We beschouwden effecten pas als positief als die statistisch significant waren en boven de drempel van een kleine relatie of effect (bijv. Cohen's  $d = 0.20$ ) uitkwamen. Ook hebben we voor elke studie een indicatie gegeven van de bewijskracht. Dat deden we door te kijken naar wat de auteurs rapporteerden over de kwaliteit van de primaire studies en over mogelijke publicatiebias. In Tabel 3 is aangegeven hoe we voor elke studie tot een score voor bewijskracht zijn gekomen.

**Tabel 3**

*Scoring van bewijskracht*

Quality appraisal	0	De auteurs voerden geen controle uit op de kwaliteit van de primaire studies.
	0.5	De auteurs voerden wel een kwaliteitscontrole uit, maar maakten daarbij geen gebruik van een gestandaardiseerde procedure.
	1	De auteurs maakten gebruik van een gestandaardiseerde procedure voor kwaliteitscontrole.
In-/exclusie o.b.v. kwaliteit	0	Er werden geen studies uitgesloten op basis van (enige indicator van) studiekwaliteit.
	1	Er werden wel studies uitgesloten op basis van (enige indicator van) studiekwaliteit <i>of</i> alle geïnccludeerde studies waren van voldoende kwaliteit.
Overall studiekwaliteit	0	De overall kwaliteit van de studies is slecht of er wordt niet over kwaliteit gerapporteerd.
	0.5	De kwaliteit van de studies varieert.
	1	De overall kwaliteit van de studies is goed.
Relatie kwaliteit-effecten*	0	Veel aspecten van studiekwaliteit zijn gerelateerd aan effecten of er wordt niet gerapporteerd over relaties.
	0.5	Sommige aspecten zijn wel gerelateerd aan effecten, andere niet.
	1	Er is geen relatie tussen studiekwaliteit en effecten.
Publicatiebias	0	Er is duidelijke publicatiebias of er wordt niet gerapporteerd over publicatiebias.
	0.5	Er is enige publicatiebias.
	1	Er is geen publicatiebias/er zijn geen aanwijzingen dat publicatiebias uitkomsten heeft bepaald.

**Noot.** \* Omdat daarin geen moderatoranalyses worden uitgevoerd, is deze categorie niet beoordeeld voor systematische reviews.

Op basis van deze scoring kenden we elke literatuurstudie een kwalificatie “zwakke bewijskracht” (–), “redelijke bewijskracht” (~) of “sterke bewijskracht” (+) toe (meta-analyses: score 0 of 1 = –; score 2 of 3 = ~; score 4 of 5 = +; systematische reviews: score 0-1 = –; score 1,5-2,5 = ~; score 3-4 = +). De eerste en derde auteur scoorden de studies onafhankelijk van elkaar; in drie gevallen verschilde hun oordeel. De Fleiss’s kappa was 0.81, wat op een zeer goede overeenkomst duidt (Landis & Koch, 1977). De auteurs bespraken de verschillen en kwamen op basis daarvan tot een definitieve scoring.

## Resultaten

In dit hoofdstuk geven we korte beschrijvingen van ieder van de geselecteerde literatuurstudies. Voor Onderzoeksvraag 1 doen we dat op basis van de categorisering die we in het vorige hoofdstuk hebben geïntroduceerd; de verschillende (sub)categorieën lichten we verder toe in Tabel 4. Bij de beschrijvingen van de meta-analyses en systematische reviews benoemen we telkens het doel van de studie en lichten we toe hoeveel en wat voor studies werden geïnccludeerd, op welke onafhankelijke variabelen de literatuurstudie zich richtte (d.w.z. op welke kenmerken van het onderwijsleerproces en/of welke aspecten van het leesgedrag of de leeshouding), hoe de afhankelijke variabele (leesbegrip) is vastgesteld, wat de belangrijkste resultaten waren in het licht van de onderzoeksvraag en wat de auteurs rapporteerden over de kwaliteit van de primaire studies en over publicatiebias. Op basis van deze beschrijvingen wordt vervolgens een overzichtstabel gepresenteerd, waarin de informatie over relaties/effecten en bewijskracht per (sub)categorie wordt samengevat. Bij de beschrijving van de conceptuele overzichtsstudies (Onderzoeksvraag 2) geven we aan op hoeveel en wat voor studies ze zich baseerden en lichten we de definitie en operationalisatie van de centrale concepten toe.

**Tabel 4**

*Toelichting op de categorisering*

---

<b>Kennis en vaardigheden</b>	
Inferenties maken/achtergrondkennis	Lezen wordt vaak voorgesteld als een proces waarbij lezers een mentale representatie van een tekst construeren (Kintsch, 1998). Omdat niet alle informatie in een tekst expliciet wordt beschreven, vraagt de constructie van zo'n mentale presentatie het maken van afleidingen of inferenties. Om inferenties te kunnen maken, is de integratie van tekstuele informatie en achtergrondkennis essentieel. Deze categorie omvat interventies die leerlingen helpen om inferenties te maken en/of die hen helpen om achtergrondkennis op te bouwen of te activeren die bij het maken van inferenties van belang is.
Integreren van meerdere teksten	<i>Multiple text/document comprehension</i> verwijst naar de processen en gedragingen waarbij leerlingen betekenis geven aan complexe inhoudens verspreid over meerdere teksten (List & Alexander, 2017). Daarvoor moeten ze keuzes moeten maken over welke (delen van) teksten ze wel/niet lezen en in welke volgorde ze die lezen, de relevantie en betrouwbaarheid van teksten bepalen, uit geselecteerde teksten informatie ophalen of afleiden die relevant is voor hun leesdoel, en die informatie vergelijken, combineren en integreren met hun achtergrondkennis (McCrudden et al., 2022; Salmerón et al., 2018).
Kennis van tekststructuur	Lezers zijn beter in staat om te bepalen welke betekenisrelaties er bestaan tussen de verschillende onderdelen van een tekst wanneer ze kennis hebben van de manier waarop teksten zijn georganiseerd en van het vocabulaire dat wordt gebruikt om die structuur tot uitdrukking te brengen (Bogaerds-Hazenberg et al., 2022). Bij informatieve of zakelijke teksten wordt vaak onderscheid gemaakt tussen een aantal basisstructuren: beschrijving, chronologie, vergelijking, oorzaak-gevolg en probleem-oplossing (Hebert et al., 2016). Bij narratieve teksten wordt gesproken van <i>story grammar</i> (Dymock, 2007).

---

(Mondelinge) taalvaardigheid	Een bekend theoretisch model van begrijpend lezen is de <i>Simple View of Reading</i> (SVR; Hoover & Gough, 1990). Het model stelt dat begrijpend lezen het product (d.w.z. de vermenigvuldiging) is van taalbegrip (mondelinge taalvaardigheid) en decoderen; het idee van vermenigvuldiging drukt uit dat beide vaardigheden essentieel zijn voor begrip en dat naarmate beide groeien, begrip exponentieel toeneemt. Mondelinge taalvaardigheid omvat onder andere woordenschat, luisterbegrip en morfologische kennis.
Schrijfvaardigheid	We beschouwen schrijfvaardigheid als het vermogen om op coherente wijze gedachten uit te drukken in geschreven tekst door middel van een proces van plannen, formuleren en reviseren (Flower & Hayes, 1981). Er zijn drie theoretische verklaringen voor de veronderstelde relatie tussen schrijfvaardigheid en leesbegrip (Van der Hoeven et al., 2022): (1) lezen en schrijven doen een beroep op dezelfde cognitieve processen en systemen (achtergrondkennis, kennis over teksten, [meta]kennis over geschreven taal en procedurele kennis, met name over effectieve strategieën); (2) lezen en schrijven zijn beide vormen van schriftelijke communicatie: door afwisselend de rol van lezer en schrijver op zich te nemen, leren leerlingen hoe schriftelijke communicatie werkt; (3) lezen en schrijven kunnen ingezet worden om eenzelfde doel te bereiken: om een werkstuk te kunnen schrijven moeten leerlingen bijvoorbeeld vaak eerst teksten lezen.
Technisch lezen (decoderen, vloeiend lezen)	Decoderen kan worden omschreven als het leren “kraken van de schriftcode”: het gaat om het leren koppelen van geschreven symbolen (letters en lettercombinaties) aan klanken en het combineren van die klanken tot betekenisvolle eenheden (Castles et al., 2018). Als kinderen die vaardigheid hebben verworven, leren ze vloeiend lezen. Bij vloeiend lezen gaat het om een combinatie van accuratesse, automatisering en <i>prosodie</i> (lezen met passende intonatie, duur, accenten en pauzes) (Kuhn et al., 2010).

	Vloeiend lezen faciliteert begrijpend lezen: als leerlingen niet meer “struikelen over woorden”, kunnen ze hun mentale inspanning richten op begrip.
Zelfregulatie, metacognitie en leesstrategieën	<p>Zelfregulatie is het vermogen om de eigen gedachten, gevoelens en gedragingen aan te sturen en aan te passen om persoonlijke doelen te bereiken (Zimmerman, 2000). Metacognitie omvat de kennis en regulatie van de eigen cognitieve activiteiten tijdens leerprocessen (Veenman et al., 2006).</p> <p>Leesstrategieën kunnen worden omschreven als de “mentale hulpmiddelen die een lezer actief en doelbewust inzet om tekstbegrip op te bouwen, te bewaken en bij te sturen” (Rogiers et al., 2022, pp. 85-86). Er is een relatie tussen leesstrategieën en een begrip als metacognitie: leesstrategieën kunnen worden gezien als de cognitieve <i>tools</i> waarmee bepaalde metacognitieve activiteiten (plannen, oriënteren, monitoren, evalueren) kunnen worden uitgevoerd: jezelf vragen stellen tijdens het lezen is bijvoorbeeld een manier om het leesproces te monitoren.</p>
<b>Leesgedrag en -houding</b>	<p>Bij leesgedrag keken we vooral naar <i>print exposure</i> (de blootstelling aan boeken) en leesfrequentie. Bij leeshouding gaat het om variabelen als leesmotivatie, leesattitude en leesbetrokkenheid. Leesmotivatie kan worden gedefinieerd als de neiging om te lezen die voortkomt uit de overtuigingen, gevoelens en doelen die iemand heeft ten aanzien van lezen (Conradi et al., 2014). Leesmotivatie is een complex concept en omvat bijvoorbeeld intrinsieke en extrinsieke motivatie, leeszelfvertrouwen, de waarde die wordt gehecht aan lezen, interesse en beheersings- en prestatiedoelen (Van Steensel, 2022). Leesattitude wordt gezien als een combinatie van overtuigingen, gevoelens en gedragsintenties (McKenna et al., 1995) en vertoont dus sterke overlap met motivatie (Schiefele et al., 2012). Bij leesbetrokkenheid gaat het om een combinatie van gedrag (bijv. iemands inzet tijdens leesactiviteiten), emoties tijdens het lezen,</p>

	cognitieve activiteiten (bijv. het gebruik van leesstrategieën) en/of sociale handelingen (bijv. uitwisseling over gelezen teksten) (Lee et al., 2021; Unrau & Quirk, 2014).
<b>Kenmerken van de leeromgeving</b>	
Digitale technologie	Hierbij kan het gaan om de inzet van uiteenlopende digitale interventies om het leesbegrip te ondersteunen, maar ook om het gebruik van digitale apparaten als medium voor het lezen (in vergelijking met het lezen vanaf papier).
Graphic organizers	Graphic organizers zijn visuele presentaties van tekstuele informatie, waarin met behulp van lijnen, pijltjes of andere ruimtelijke weergaven de inhoud en structuur van een tekst worden weergegeven en belangrijke conceptuele relaties worden verduidelijkt (Darch & Eaves, 1986). Voorbeelden van graphic organizers zijn <i>concept maps</i> , <i>story maps</i> en Venn-diagrammen.
Peer mediation/coöperatief leren	Bij peer-mediated instructie en coöperatief leren gaat het om situaties waarin leerlingen in duo's of kleine groepjes samenwerken aan een (lees)taak (Maheady et al., 1988; O'Donnell, 2011; Slavin, 2011). De gunstige werking van zulke vormen van samenwerking wordt op verschillende manieren verklaard (O'Donnell, 2006; Slavin, 2011). Die werking kan voortkomen uit sociale processen, bijvoorbeeld doordat de wederzijdse afhankelijkheid tijdens gezamenlijke taken de taakmotivatie bevordert. Die werking kan ook voortkomen uit cognitieve processen, bijvoorbeeld doordat de coöperatieve context diepere verwerking stimuleert: zo profiteren leerlingen ervan als ze leerstof aan elkaar uitleggen.
Personalisatie	Personalisatie verwijst naar benaderingen waarbij aanbod en instructie worden aangepast aan het niveau, de interesses of andere persoonlijke eigenschappen van leerlingen (Baye et al., 2019).
Tutoring	Tutoring verwijst naar situaties waarin leerlingen individuele begeleiding of begeleiding in kleine groepjes krijgen bij verwerking van de leerstof (Baye et al., 2019). Net als bij personalisatie worden

	<p>aanbod en instructie afgestemd op de behoeften van leerlingen. Een tutor kan een leraar, <i>paraprofessional</i>, vrijwilliger of medeleerling zijn. Er is dus een gedeeltelijke overlap tussen deze categorie en peer mediation/coöperatief leren; in de meta-analyse waarin de effecten van tutoring zijn onderzocht (Baye et al. 2019), werd echter geen onderscheid gemaakt tussen type tutor.</p>
Vakintegratie	<p>Bij vakintegratie worden de inhouden van het lees- en vakonderwijs verbonden. Dat kan de vorm krijgen van taalgericht vakonderwijs, door Hajer (2005) gedefinieerd als “vakonderwijs waarin expliciete taaldoelen worden gesteld, dat contextrijk is, vol interactiemogelijkheden zit en waarbinnen benodigde taalsteun wordt geboden” (p. 4). Toegepast op lezen kan het dan bijvoorbeeld gaan om vaklessen waarin expliciete aandacht wordt besteed aan de structuur van informatieve teksten of leerlingen worden geholpen bij het samenvatten. Vakintegratie kan ook inhouden dat in de lessen Nederlands aandacht wordt besteed aan vakteksten.</p>
<b>Kenmerken van de schoolorganisatie</b>	<p>Hierbij gaat het om organisatiekenmerken die het leesonderwijs op een school faciliteren, zoals de aanwezigheid van een taal- of leesbeleid of de aandacht voor gezamenlijke afstemming en professionalisering.</p>

## Meta-analyses: zich normaal ontwikkelende leerlingen

### Kennis en vaardigheden

**Inferenties maken/achtergrondkennis.** Elleman (2017) testte het effect op leesbegrip van instructie waarmee leerlingen worden geholpen om inferenties te maken, door de auteur gedefinieerd als “the process by which a reader integrates information within or across texts using his or her background knowledge to fill in information not explicitly stated” (p. 761). Ze onderscheidde zeven instructieve activiteiten: (1) beantwoorden/genereren van inferentievragen, (2) gebruik van aanwijzingen in de tekst om tot een coherente tekstrepresentatie te komen, (3) gebruik van achtergrondkennis, (4) gebruik van graphic organizers om tekststructuur te verhelderen of de relatie met achtergrondkennis te leggen, (5) onderscheiden van en oefenen met verschillende typen inferenties, (6) zelf uitleggen of uitbreiden van ideeën in de tekst om zo de relatie tussen tekstuele informatie en achtergrondkennis te kunnen leggen, en (7) *perspective taking* (zich verplaatsen in personages om motieven af te leiden en zo voorspellingen te kunnen doen of oorzaak-gevolgrelaties te kunnen afleiden). Elleman analyseerde 33 (quasi-)experimentele studies afkomstig uit 25 publicaties, met in totaal 1.752 leerlingen uit basis- en voortgezet onderwijs (Grade 2-9). Van die 33 studies gingen er 14 over leerlingen uit Grade 6 tot en met 9. Er werden effecten getest op drie typen leesbegrip: letterlijk begrip (vragen gericht op informatie die expliciet in de tekst staat), inferentieel begrip (vragen gericht op informatie die niet expliciet in de tekst staat) en algemeen begrip (een combinatie van de twee). De effectgrootten waren in alle gevallen significant en positief: voor letterlijk begrip was het effect klein ( $d = 0.28$ ,  $p = .04$ ), voor inferentieel en algemeen begrip middelgroot (resp.  $d = 0.68$ ,  $p < .01$  en  $d = 0.58$ ,  $p < .01$ ). Voor algemeen begrip was er geen sprake van heterogeniteit in effectgrootten, voor de andere twee effectmaten wel; leerjaar bleek in die laatste gevallen echter geen significante moderator. Inferentie-instructie leek dus ook voor middelbare scholieren een gunstige bijdrage te leveren aan leesbegrip. Welk type instructie voor hen het beste werkte, is niet nagegaan: er werden geen moderatoranalyses uitgevoerd binnen onderwijsfasen. Elleman volgde geen formele procedure voor quality appraisal en sloot geen studies uit op andere gronden dan design, maar codeerde wel een aantal kwaliteitskenmerken (wel/geen random toewijzing, aard van de controlegroep, correctie voor aanvangsverschillen). Er was variatie in studiekwaliteit, hoewel die wel verschilde per kenmerk: zo was er vaak sprake van random toewijzing, maar werd er weinig aandacht

besteed aan implementatiekwaliteit. Studiekwaliteit bleek niet gerelateerd aan effectgrootte. Ook vond de auteur geen aanwijzingen voor publicatiebias.

Ook Rice en Wijekumar (2024) onderzochten het effect van inferentie-instructie op leesbegrip, maar maakten een wat minder fijnmazige vergelijking dan Elleman (2017): ze onderscheidden alleen strategie-instructie en oefenen met inferentievragen. De manier waarop leesbegrip werd gemeten, werd niet nader gespecificeerd dan door het onderscheid tussen gestandaardiseerde en door onderzoekers ontwikkelde toetsen. De auteurs analyseerden 81 (quasi-)experimentele vergelijkingen uit 56 studies, die tezamen 5.088 participanten omvatten, variërend van jonge kinderen tot volwassen leerders. Van alle vergelijkingen had 20 procent betrekking op *middle school*-leerlingen.<sup>2</sup> Voor deze doelgroep was het effect van inferentie-instructie op leesbegrip significant en positief; de effectgrootte overschreed net niet de drempel van een middelgroot effect ( $g = 0.49, p < .001$ ). Over het geheel vonden de auteurs geen verschil in effecten tussen de twee instructietypen; het is echter niet duidelijk of dat ook voor *middle school*-leerlingen gold. Rice en Wijekumar onderzochten op basis van de zogenaamde *What Works Clearinghouse (WWC)*-richtlijnen en procedures uit eerdere meta-analyses verschillende aspecten van studiekwaliteit (design, uitval, aard van de controlegroep, implementatiekwaliteit, betrouwbaarheid van effectmaten en aanvangsverschillen tussen condities). Ze gebruikten die aspecten overigens niet om studies uit te sluiten. Wel onderzochten ze moderatoreffecten van de geobserveerde variatie in kwaliteit, maar vonden er geen. De auteurs vonden beperkte aanwijzingen voor publicatiebias.

**Kennis van tekststructuur.** Hebert et al. (2016) gingen het effect na van tekststructuurinstructie op het leesbegrip van leerlingen in basis- en voortgezet onderwijs (Grade 1-12). De auteurs omschreven tekststructuurinstructie als lees- of schrijfinstructie waarin leerlingen leren over het gebruik van tekststructuren (met name, beschrijving, chronologie, vergelijking, oorzaak-gevolg en probleem-oplossing) met als doel hun begrip van teksten te vergroten. Die instructie kan bijvoorbeeld bestaan uit het leren identificeren van tekststructuren, het leren beantwoorden van vragen over tekststructuren en het leren schrijven van teksten volgens bepaalde structuren. Bij het meten van leesbegrip werd een onderscheid gemaakt tussen begrip van informatieve teksten (vaak gemeten met niet-

---

<sup>2</sup> Meestal omvat *middle school* drie leerjaren, doorgaans Grade 6 tot en met 8. Sommige studies in de meta-analyse van Rice en Wijekumar (2024) hadden ook betrekking op *high-school*-leerlingen (Grade 9-12), maar over hen kon geen aparte uitspraak worden gedaan, omdat ze in één categorie werden samengenomen met volwassen leerders.

gestandaardiseerde toetsen) en algemeen leesbegrip (vooral gemeten met gestandaardiseerde toetsen). Hebert et al. analyseerden 45 (quasi-)experimenten, waarvan er 22 betrekking hadden op leerlingen in het voortgezet onderwijs; het totale aantal leerlingen werd niet genoemd. Het gemiddelde effect op het begrip van informatieve teksten was op zowel korte als langere termijn significant, positief en middelgroot (resp.  $g = 0.57$ , 95% CI [0.39, 0.76]<sup>3</sup> en  $g = 0.57$ , 95% CI [0.26, 0.87]). Schoolniveau (basisschool versus middelbare school) was geen significante moderator van effecten op korte termijn en op langere termijn was er geen heterogeniteit in effectgrootten. In beide gevallen gold dus dat het gevonden effect ook opging voor vo-leerlingen. Op het begrip van teksten met een andere structuur dan de aangeleerde werd overigens geen significant effect gevonden. Op algemeen leesbegrip was dat er wel ( $g = 0.57$ , 95% CI [0.03, 0.25]), maar Hebert et al. twijfelden aan de geldigheid van dit resultaat, omdat maar vier van de acht effectgrootten positief waren. Conclusies over de effecten van specifieke typen tekststructuurinstructie voor vo-leerlingen konden niet worden getrokken, omdat er geen moderatoranalyses binnen onderwijsfasen werden uitgevoerd. Hoewel niet volgens een formele systematiek, voerden de auteurs wel een evaluatie uit van de kwaliteit van de geïncludeerde studies: ze keken naar design (wel/geen random toewijzing), uitval, aard van de controleconditie, controle voor leraareffecten, implementatiekwaliteit en kwaliteit van de effectmaten. Ze sloten geen studies uit op basis van kwaliteit, maar constateerden wel dat er aanzienlijke variatie was, die overigens niet bleek samen te hangen met effectgrootte. Analyses duiden op enige publicatiebias, maar de auteurs concludeerden dat die geen consequenties had voor de interpretatie van de resultaten.

Ook Pyle et al. (2017) onderzochten het effect van tekststructuurinstructie (meer specifiek: instructie in het gebruik van kennis over de structuur van informatieve teksten) op het leesbegrip van leerlingen uit basis- en voortgezet onderwijs (kleuterbouw-Grade 12). Leesbegrip werd gemeten via gestandaardiseerde en niet-gestandaardiseerde toetsen en met uiteenlopende toetswijzen: begripsvragen, graphic organizers, samenvattingen, *recall*-taken (“herinneringsvragen” zonder de tekst erbij) en tekststructuuridentificatie. De meta-analyse bestond uit 21 (quasi-)experimentele studies met in totaal 4.254 leerlingen. Van die studies hadden er 16 betrekking op leerlingen uit het basisonderwijs en vijf op middelbare scholieren. Voor die laatste groep werd een significant, positief en middelgroot effect gevonden ( $d = 0.63$ , 95% CI [0.50, 0.77]). Over alle studies heen bleken effecten gerelateerd aan het aantal

---

<sup>3</sup> Studies verschillen in het type statistische informatie. Als er betrouwbaarheidsintervallen werden gepresenteerd, hebben we daar de voorkeur aan gegeven; als dat niet het geval was, presenteren we *p*-waarden.

aangeleerde tekststructuren en interventieduur, maar het is niet duidelijk of dat ook voor vo-leerlingen gold; voor hen werden effecten niet gespecificeerd naar type interventie. Pyle et al. gebruikten een formele procedure voor het vaststellen van studiekwaliteit (Gersten et al., 2005), waarbij ze keken naar design (wel/geen random toewijzing), helderheid van de interventiebeschrijving, implementatiekwaliteit, evaluatiewijze en berekening van de effectgrootten. Op basis hiervan onderscheidde de auteurs zes studies van hogere en 13 van lagere kwaliteit, hoewel ze niet specificeerden tot welke categorie de studies in het voortgezet onderwijs hoorden. Studies werden alleen uitgesloten als ze niet waren gepubliceerd in een *peer-reviewed* tijdschrift. Studiekwaliteit bleek gerelateerd aan interventie-effecten: de effecten waren kleiner bij studies van hogere kwaliteit. Er was enige indicatie voor publicatiebias, maar analyses lieten zien dat die de resultaten niet vertekende.

**(Mondelinge) taalvaardigheid.** Elleman et al. (2009) onderzochten het effect van woordenschatinstructie op het leesbegrip van kinderen in de voorschoolse periode<sup>4</sup>, het basis- en voortgezet onderwijs. Woordenschatinstructie nam verschillende vormen aan. In de meta-analyse werd bijvoorbeeld onderscheid gemaakt tussen benaderingen waarin leerlingen woordbetekenissen leren, waarin ze nieuwe woorden verbinden aan bekende woorden (bijv. door antoniemen te geven) en waarin ze nieuwe woorden actief gebruiken, bijvoorbeeld in een te schrijven tekst. Leesbegrip werd doorgaans gemeten met meerkeuze- of open vragen over *passages* (gedefinieerd als tekstfragmenten met ten minste vier zinnen) en met zowel gestandaardiseerde als niet-gestandaardiseerde tests. De meta-analyse omvatte 44 studies afkomstig uit 37 publicaties, met in totaal 3.063 kinderen. Elleman et al. vonden geen effect van woordenschatinstructie op gestandaardiseerde toetsen, maar wel op niet-gestandaardiseerde toetsen; dat effect was positief en middelgroot ( $d = 0.50, p < .01$ ). Leerjaar was geen significante moderator, wat betekent dat het gunstige effect ook gold voor vo-leerlingen. Over alle studies heen waren er geen effecten van instructiekenmerken; voor de vo-leerlingen werd dat niet afzonderlijk bepaald. Elleman et al. gebruikten geen formele procedure voor quality appraisal, maar codeerden wel bepaalde kenmerken van studiekwaliteit. Studies werden niet uitgesloten op basis van kwaliteit. Die liet dan ook variatie zien: zo ontbrak vaak informatie over implementatiekwaliteit, training van uitvoerders en betrouwbaarheid van effectmaten. Geen van de kwaliteitskenmerken bleek echter gerelateerd aan effectgrootte. De auteurs vonden nauwelijks aanwijzingen voor publicatiebias.

---

<sup>4</sup> Bij kinderen in de voorschoolse periode gaat het om het begrip van voorgelezen tekst.

Goodwin en Ahn (2013) testten het effect van morfologische instructie op verschillende taalvaardigheidsmaten, waaronder leesbegrip. Door middel van morfologische instructie leren leerlingen over hoe woorden zijn opgebouwd uit kleinere, betekenisvolle delen, zogenaamde *morfemen*.<sup>5</sup> De auteurs onderscheidden vier instructiebenaderingen: (1) het verbeteren van het bewustzijn van de morfologische structuur van woorden, (2) het bevorderen van de kennis van de betekenis van morfemen, (3) het ondersteunen van *morphological problem solving* (het ophelderen van woordbetekenissen door het gebruik van morfologische informatie) en (4) het formuleren van hypothesen over de betekenis van onbekende woorden door morfologische analyse. Bij de meting van leesbegrip werd geen nadere toelichting gegeven dan dat er onderscheid werd gemaakt tussen gestandaardiseerde toetsen en toetsen ontwikkeld door onderzoekers. Goodwin en Ahn analyseerden 29 (quasi-)experimentele studies en één niet-experimentele studie, afkomstig uit 27 verschillende publicaties; over het totale aantal participanten ontbrak informatie. Van de 92 gerapporteerde effectgrootten hadden er 14 betrekking op leesbegrip, maar onduidelijk is hoeveel daarvan gingen over vo-leerlingen (hoewel die wel als aparte categorie werden geanalyseerd). Over alle leeftijdsgroepen heen vonden de auteurs geen significant interventie-effect op leesbegrip ( $d = 0.09$ , 95% CI [-0.01, 0.20]). Dat interventie-effect leek er ook niet te zijn voor middelbare scholieren, maar wel voor jongere kinderen (voorschool/basisschool) en voor steekproeven met een brede leeftijdsgroep. Ze volgden geen vaste procedure voor quality appraisal, maar codeerden wel enkele kwaliteitskenmerken (design, type controlegroep, implementatiecheck en publicatietype). Studies werden niet uitgesloten op grond van kwaliteit. De kwaliteit van de geïncludeerde studies liet dan ook variatie zien en één van de kwaliteitsaspecten (design) bleek gerelateerd aan interventie-effecten: die waren groter bij quasi- en niet-experimentele studies dan bij experimentele studies. De auteurs vonden geen aanwijzingen voor publicatiebias.

Rogde et al. (2019) onderzochten de effecten van wat ze *taalbegripsinterventies* noemden: interventies gericht op het stimuleren van vaardigheden als woordenschat, grammatica en narratief begrip. Tegen de achtergrond van theoretische modellen als de Simple View of Reading (zie Tabel 4), veronderstelden de auteurs een mediatie-effect (taalbegripsinterventies leiden via het vergroten van taalbegrip tot meer leesbegrip), dat met name zou gelden voor oudere leerlingen. In de eerdere fasen van de leesontwikkeling wordt

---

<sup>5</sup> Het woord *onbegrijpelijk* heeft bijvoorbeeld drie morfemen: het voorvoegsel *on-*, de stam *begrijp* en het achtervoegsel *-elijk*.

leesbegrip immers vooral bepaald door decodeervaardigheid. De meta-analyse van Rogde et al. omvatte 43 (quasi-)experimentele studies, waarvan er 16 (met in totaal 17.803 leerlingen) leesbegrip als uitkomst hadden, alle gemeten met gestandaardiseerde tests. Zes van deze studies zijn uitgevoerd met vo-leerlingen (Grade 6 tot 8); de overige tien studies hadden betrekking op jongere kinderen. Voor leerlingen in het voortgezet onderwijs was er een significant, positief effect op leesbegrip, maar de effectgrootte was marginaal ( $g = 0.06, p < .01$ ). Omdat er bovendien geen effect was op taalbegrip ( $g = 0.07, p > .05$ ), concludeerden de auteurs dat er van het veronderstelde mediatie-effect geen sprake was. Omdat Rogde et al. geen moderatoranalyses uitvoerden binnen onderwijsfasen kan ook geen uitspraak worden gedaan over de vraag of bepaalde typen interventies wel effectief waren voor vo-leerlingen. De auteurs voerden geen formele quality appraisal uit, maar codeerden wel verschillende aspecten van studiekwaliteit (bijv. wel/geen random toewijzing, implementatiekwaliteit, uitval). Ze sloten geen studies uit anders dan op grond van de aan- of afwezigheid van een controlegroep. Wel vonden ze een relatie tussen effectgrootte en variatie in de genoemde kwaliteitskenmerken: effecten waren groter bij quasi-experimenten, negatief bij gerapporteerde implementatieproblemen en groter wanneer er minder uitval was. Ze vonden geen aanwijzingen voor publicatiebias.

**Schrijfvaardigheid.** Eén van de doelen van de meta-analyse van Graham en Hebert (2011) was om na te gaan of schrijven over gelezen teksten het leesbegrip bevordert.<sup>6</sup> Ze maakten onderscheid tussen vijf categorieën schrijfactiviteiten: (1) het beantwoorden of genereren van vragen, (2) aantekeningen maken, (3) samenvatten, (4) het schrijven van korte reacties op gelezen tekst en (5) het schrijven van een langere tekst naar aanleiding van gelezen tekst. De manier waarop leesbegrip werd gemeten, werd niet nader toegelicht: de auteurs maakten alleen onderscheid tussen gestandaardiseerde en niet-gestandaardiseerde toetsen. De meta-analyse omvatte 95 (quasi-)experimentele studies, waarvan er 63 betrekking hadden op leerlingen in het voortgezet onderwijs. Over alle leeftijdsgroepen heen rapporteerden Graham en Hebert significante, positieve, kleine tot middelgrote effecten (gestandaardiseerde toetsen:  $d = 0.37, p < .001$ ; niet-gestandaardiseerde toetsen:  $d = 0.50, p < .001$ ). Alleen bij de niet-gestandaardiseerde toetsen was er sprake van heterogeniteit in effectgrootten. Onderwijsfase was voor die toetsen een significante moderator, maar er was alleen een verschil tussen middle school- en high school-leerlingen (voor de eerste groep was

---

<sup>6</sup> Graham en Hebert (2011) gingen daarnaast de effecten na van schrijfstructuur en het vergroten van de schrijffrequentie op leesbegrip, maar daarbij splitsten ze effecten niet uit naar verschillende leeftijdsgroepen.

het effect 0.33 SD groter dan voor de tweede groep), niet tussen basisschoolleerlingen en middelbare scholieren. Als leerlingen echter expliciete training kregen in het uitvoeren van schrijfactiviteiten waren juist de effecten voor high school-leerlingen groter (het verschil was 0.27 SD). Hoewel niet op basis van een vaste procedure voor quality appraisal, voerden Graham en Hebert een uitgebreide kwaliteitscheck uit, waarbij ze keken naar: wel/geen random toewijzing, uitval, beschrijving van de controleconditie, implementatiekwaliteit, controle van leraareffecten, betrouwbaarheid van effectmaten en *baseline equivalence*. Ze sloten geen studies uit anders dan op grond van design. Er was dan ook variatie in kwaliteit: over de kwaliteitsindicatoren heen varieerde de proportie studies die eraan voldeed van .30 tot .84. De totaalscore voor studiekwaliteit was gerelateerd aan effectgrootte: hogere kwaliteit ging samen met kleinere effecten. De auteurs maakten geen analyse van mogelijke publicatiebias.

Graham et al. (2023) onderzochten hoe schrijfinstructie in het voortgezet onderwijs effectief kan worden vormgegeven en keken daarbij ook welk type schrijfinstructie gunstig is voor leesbegrip. Dat laatste werd op verschillende manieren gemeten: met onder meer traditionele begripstoetsen (vragen over teksten), *cloze-tests*<sup>7</sup> en naverteltaken.

Schrijfinstructie kon bestaan uit verschillende typen activiteiten:

1. Oefenen, dat wil zeggen: vergroten van de schrijftijd.
2. De inzet van procedures of hulpmiddelen om leerlingen te ondersteunen tijdens het schrijven (bijv. doelen stellen, voorbereidende activiteiten, het gebruik van digitale tools, feedback en hulp van medeleerlingen).
3. Het aanleren van specifieke vaardigheden (bijv. spelling, grammatica, zinsopbouw), processen (bijv. de toepassing van schrijfstrategieën, reviseren, samenvatten of het inzetten van *mental imagery*, d.w.z. het gebruik van mentale visualisaties om het begrip te ondersteunen) of kennis (bijv. kennis over tekststructuur, woordenschat of kennis over tekstkwaliteit op basis van goede voorbeelden).
4. De inzet van specifieke methoden om schrijfvaardigheid te verbeteren (bijv. observatie en computergestuurde instructie).
5. Uitgebreide schrijfprogramma's, bijvoorbeeld programma's waarin een procesgerichte benadering van schrijven (schrijven als een cyclisch proces van plannen, formuleren en reviseren) wordt aangeleerd.

---

<sup>7</sup> Een test waarbij in een tekst met een zekere regelmaat woorden zijn weggelaten, die de leerling moet invullen.

De meta-analyse omvatte 352 (quasi-)experimentele studies met in totaal 52.984 leerlingen; in 37 van de studies werd ook gekeken naar effecten op leesbegrip. Schrijfinstructie had een significant, positief en klein effect op leesbegrip ( $g = 0.26$ , 95% CI [0.11, 0.41]): leerlingen op bovenstaande manieren ondersteunen bij het schrijven zorgde ervoor dat ze niet alleen beter leerden schrijven (de gemiddelde effectgrootte op schrijfwatkomsten was  $g = 0.49$ , 95% CI [0.43, 0.55]), maar ook dat ze teksten beter begrepen. Graham et al. rapporteerden geen aparte effecten voor de hierboven genoemde instructietypen. Hoewel de auteurs geen standaardprocedure voor quality appraisal toepasten, keken ze wel naar bepaalde aspecten van studiekwaliteit. Ze sloten studies uit op basis van steekproefgrootte (studies met minder dan 10 participanten werden niet opgenomen), design (alleen quasi-experimenten werden geïncludeerd) en pre-testverschillen. Toch was er ook onder de geïncludeerde studies variatie in kwaliteit; wel verschilde die per indicator: zo was uitval in vrijwel geen van de studies een probleem, maar besteedden maar weinig onderzoekers aandacht aan implementatiekwaliteit. Geen van de indicatoren was gerelateerd aan interventie-effecten. Verder vonden de auteurs op basis van een analyse van alle effectgrootten (d.w.z. niet alleen die voor leesbegrip) enige indicatie voor publicatiebias.

Hebert et al. (2013) vergeleken het effect van vier typen schrijfin interventies op leesbegrip: samenvattingen schrijven, vragen beantwoorden, notities maken en *extended writing*; bij dat laatste moesten leerlingen bijvoorbeeld op basis van de inhoud van een gelezen tekst een nieuwe tekst schrijven, op basis van tekstuele informatie een standpunt onderbouwen of een persoonlijke reactie op een gelezen tekst schrijven. Leesbegrip werd op verschillende manieren gemeten: via gestandaardiseerde tests en tests die door onderzoekers waren ontwikkeld, via taken die wel of niet leken op de interventie-activiteiten en via uiteenlopende toetswijzen (o.a. meerkeuze- of open vragen over teksten, samenvattingen, recall-taken, essays en associatietaken). De meta-analyse omvatte 27 (quasi-)experimentele studies, waarvan er 23 betrekking hadden op leerlingen in het voortgezet onderwijs en vier op basisschoolleerlingen; het totale aantal participanten werd niet vermeld. Anders dan in de meta-analyses van Graham en Hebert (2011) en Graham et al. (2023) werd in de meta-analyse van Hebert et al. (2013) geen vergelijking gemaakt met een controleconditie, maar werden de interventies alleen onderling vergeleken. Uit die vergelijkingen kwamen echter geen significante verschillen naar voren: het ene type interventie leek dus niet meer bij te dragen aan leesbegrip dan het andere, ook niet voor vo-leerlingen. De auteurs sloten geen studies uit anders dan op grond van design (alle ingesloten studies waren experimenteel of quasi-experimenteel), maar brachten wel verschillende aspecten van studiekwaliteit in kaart

en drukten die uit in een cijfer van 0 tot 1. Over de aspecten heen varieerde het gemiddelde kwaliteitscijfer van 0.45 tot 0.64, waarbij Hebert et al. aantekenden dat er met name weinig informatie was over de implementatiekwaliteit van de interventies. De auteurs maakten geen analyse van de relatie tussen studiekwaliteit en effectgrootte. Ook keken ze niet naar mogelijke publicatiebias.

**Zelfregulatie, metacognitie en leesstrategieën.** Vijf meta-analyses richtten zich op zelfregulatie, metacognitie en/of leesstrategieën: in één geval bevatte de meta-analyse uitsluitend correlatieve studies (Sun et al., 2021), de overige studies waren meta-analyses van interventiestudies (Davis, 2010; Fan, 1993; Haller et al., 1988; Okkinga et al., 2018).

Sun et al. (2021) onderzochten bij verschillende leeftijdsgroepen (kleuters, basisschoolleerlingen, middelbare scholieren, universitaire studenten) de correlaties tussen leesbegrip en het gebruik van vier typen leesstrategieën: (1) *affectieve strategieën*: de “mentale kracht” die lezers in staat stelt om negatieve gevoelens gerelateerd aan leeservaringen (bijv. leesangst) te overwinnen; (2) *elaboratiestrategieën*: activiteiten zoals parafraseren en samenvatten, die lezers helpen om informatie uit verschillende tekstdelen met elkaar te integreren en tekstuele informatie te integreren met bestaande kennis; (3) *monitoringsstrategieën*: het vermogen tot zelfregulatie, *self-questioning* en het bepalen van het verloop (bijv. de snelheid) van het leesproces; en (4) *organisatiestrategieën*: het vermogen om tekstuele informatie te ordenen aan de hand van kennis over tekststructuren. Hoe leesbegrip werd gemeten, werd niet nader toegelicht. De meta-analyse van Sun et al. omvatte 57 studies met in totaal 21.548 participanten, waarvan er 16 zijn uitgevoerd met leerlingen in het voortgezet onderwijs ( $n = 13.506$ ). Over alle leeftijdsgroepen heen varieerde de samenhang tussen leesbegrip en het gebruik van de vier strategieën tussen Fisher's  $z = 0.43$  (middelgroot effect) en Fisher's  $z = 0.58$  (groot effect). Geen van de verschillen in effectgrootten tussen de vier typen strategieën was significant. Bovendien bleek leeftijdsgroep voor geen van de vier verbanden een significante moderator: de samenhang tussen leesbegrip en het gebruik van de vier typen strategieën was voor vo-leerlingen dus even sterk als voor de andere groepen. Sun et al. maakten geen analyse van studiekwaliteit en ze sloten geen studies uit op grond van studiekwaliteit anders dan steekproefgrootte (het minimum aantal participanten was 30). De auteurs onderzochten mogelijke publicatiebias, maar vonden daar geen aanwijzingen voor.

Davis (2010) testte het effect van strategie-instructie op het leesbegrip van leerlingen in de bovenbouw van de basisschool en de onderbouw van de middelbare school (Grade 4-8). In de onderzochte interventies kregen leerlingen twee of meer van de volgende

leesstrategieën aangeboden: strategieën waarmee ze hun leesbegrip konden monitoren (bijv. tijdens het lezen samenvatten), strategieën waarmee ze begripsproblemen konden herstellen (bijv. herlezen of de zinscontext gebruiken om de betekenis van een onbekend woord te achterhalen) en strategieën waarmee ze hun tekstbegrip konden versterken of uitbreiden (bijv. inferenties maken). Davis onderzocht 90 studies met ongeveer 6.180 leerlingen in de experimentele groepen. Leesbegrip werd op uiteenlopende manieren gemeten (met vragen over teksten, cloze- of *maze-tests*<sup>8</sup>, zinsverificatietaken en naverteltaken). In de analyses werd overigens alleen onderscheid gemaakt tussen gestandaardiseerde en niet-gestandaardiseerde toetsen. Het effect van strategie-instructie bleek in beide gevallen significant en positief: op gestandaardiseerde toetsen was er sprake van een klein effect ( $g = 0.36$ , 95% CI [0.21, 0.51]), op niet-gestandaardiseerde toetsen was het effect middelgroot ( $g = 0.55$ , 95% CI [0.36, 0.73]). Davis testte ook langere-termijneffecten: op gestandaardiseerde toetsen was het retentie-effect niet significant ( $g = 0.27$ , 95% CI [-0.04, 0.58]), maar dat resultaat was gebaseerd op maar drie effectgrootten; op niet-gestandaardiseerde toetsen (13 effectgrootten) was het retentie-effect wel significant ( $g = 1.06$ , 95% CI [0.65, 1.46]). Hoewel de auteur geen aparte effectgrootten presenteerde voor de verschillende leerjaren, kan er uit de metaregressies die ze uitvoerde, worden afgeleid dat er voor vo-leerlingen sprake was van positieve effecten: effecten op gestandaardiseerde toetsen waren namelijk kleiner als de studie was gericht op leerlingen in Grade 4 en 5 (en dus groter voor vo-leerlingen;  $\beta = -0.28$ ) en effecten op niet-gestandaardiseerde toetsen waren kleiner voor leerlingen in Grade 5 en 8 (en dus groter voor leerlingen in Grade 6 en 7;  $\beta = -0.30$ ). Op basis van de analyses kunnen geen uitspraken worden gedaan over de effecten voor vo-leerlingen van instructie in specifieke strategieën. Davis volgde geen standaardprocedure voor quality appraisal, maar registreerde wel een aantal aspecten van studiekwaliteit: design (wel/geen random toewijzing), aandacht voor implementatiekwaliteit, aard van de controleconditie, baseline equivalence en correctie voor aanvangsverschillen. Ze vond aanzienlijke verschillen in kwaliteit, maar sloot op basis daarvan geen studies uit. Twee kwaliteitsaspecten bleken gerelateerd aan effectgrootte: wanneer implementatiekwaliteit werd gemeten en wanneer werd gecorrigeerd voor aanvangsverschillen tussen condities waren effecten kleiner. Verder vond Davis enige indicatie voor publicatiebias, maar ze constateerde dat die te beperkt was om van invloed te kunnen zijn op de conclusies.

---

<sup>8</sup> Een variant op een cloze-test, waarbij eveneens woorden zijn weggelaten uit een tekst die een leerling moet aanvullen, maar waarbij de leerling moet kiezen uit opties.

Fan (1993) onderzocht het effect op leesbegrip van instructie in wat zij *metacognitieve strategieën* noemde: activiteiten zoals (jenzelf) vragen stellen tijdens het lezen, samenvatten, parafraseren, ophelderen van onduidelijkheden, hoofdgedachten identificeren, details aan hoofdgedachten koppelen, tekststructuur herkennen, inferenties maken, voorspellen van tekstinhoud, het leesproces evalueren en reguleren, en het detecteren van leesfouten. Leesbegrip werd op diverse manieren gemeten, via gestandaardiseerde toetsen en tests ontwikkeld door onderzoekers en met verschillende toetswijzen (meerkeuzevragen, open vragen, samenvattingen, cloze tests, identificeren/onderstrepen van relevante informatie en recall-taken). In de meta-analyse werden 41 interventiestudies met 212 effectgrootten voor leesbegrip opgenomen. De dataset omvatte in totaal 3.219 leerlingen/studenten uit basis-, voortgezet en hoger onderwijs, maar niet in alle gevallen werd de leeftijdsgroep gespecificeerd. Bij 14 van de 46 effectgrootten waarvoor die informatie wel bekend was, kwamen de leerlingen uit Grade 7 tot en met 9, bij vijf uit Grade 10 tot en met 12. Fan vond over alle studies heen een significant, positief en middelgroot interventie-effect op leesbegrip ( $g = 0.56$ ,  $SE = 0.05$ ). Leerjaar bleek geen significante moderator: het effect voor vo-leerlingen verschilde dus niet van dat voor leerlingen uit andere leeftijdsgroepen. De auteur analyseerde geen moderator-effecten van instructiekenmerken voor vo-leerlingen afzonderlijk. Hoewel Fan geen formele quality appraisal uitvoerde, codeerde ze wel enkele aspecten van studiekwaliteit (type posttest, selectie van participanten, toewijzing aan condities, aard van de controleconditie en steekproefomvang). Ze gebruikte studiekwaliteit (met name de afwezigheid van een vergelijkbare controlegroep) om studies uit te sluiten. Toch was er ook onder geïnccludeerde studies variatie in kwaliteit; die variatie bleek echter niet gerelateerd aan effectgrootte. Fan maakte geen analyse van mogelijke publicatiebias.

Ook Haller et al. (1988) testten het effect op leesbegrip van instructie in metacognitieve strategieën, door hen gelabeld als *metacognitieve vaardigheden*. Ze verdeelden die onder in drie clusters: *bewustzijn*, *monitoring* en *regulatie*. Bij bewustzijn gaat het zowel om het bewustzijn van het eigen niveau van begrip als om het bewustzijn van opvallende tekstkenmerken: welke informatie staat expliciet in de tekst en welke moet worden afgeleid, bevat de tekst ambigue of zelfs inaccurate informatie? Monitoring heeft betrekking op het controleren van begrip, bijvoorbeeld door activiteiten voorafgaand aan het lezen (een leesdoel stellen, voorspellingen doen), activiteiten tijdens het lezen (jenzelf vragen stellen, samenvatten, parafraseren), activiteiten na het lezen (de geldigheid van voorspellingen over tekstinhoud nagaan) en *self-directing activities* (tekstuele informatie integreren met achtergrondkennis, hoofdgedachten vergelijken). Bij regulatie gaat het om

*compensatiestrategieën*: activiteiten zoals herlezen, waarmee begripsproblemen kunnen worden gerepareerd. De meta-analyse bestond uit 20 (quasi-)experimentele studies, waaraan in totaal 1.553 leerlingen uit basis- en voortgezet onderwijs (Grade 2-12) deelnamen. De auteurs vonden een “substantieel” effect: de gemiddelde effectgrootte was 0.71, met een standaardafwijking van 0.81. Hoewel Haller et al. geen formele moderatoranalyse uitvoerden, concludeerden ze wel dat de instructie van metacognitieve vaardigheden het meest effectief was voor leerlingen in Grade 7 en 8.<sup>9</sup> Daarnaast suggereerden ze dat er verschillen waren in de effecten van typen instructie-activiteiten (opnieuw zonder te verwijzen naar een moderatoranalyse), maar ze presenteerden geen afzonderlijke conclusies voor vo-leerlingen. De auteurs gebruikten geen gestandaardiseerde procedure voor quality appraisal, maar registreerden wel enkele aspecten van studiekwaliteit, zoals design (wel/geen random toewijzing), de vraag of de controlegroep zich bewust was van de aanwezigheid van een experimentele groep en hoe de onderzoekers omgingen met een mogelijk *Hawthorne-effect*<sup>10</sup>. Ze sloten studies uit op basis van het ontbreken van een controlegroep (zie eerder). Hoewel ze geen verdere informatie gaven over de variatie in de kwaliteit van de geïncludeerde studies, analyseerden ze wel of er een relatie was tussen kwaliteit en effectgrootte, maar vonden die niet. Verder onderzochten Haller et al. niet of er sprake was van publicatiebias.

Okkinga et al. (2018) gingen het effect na van de toepassing van strategie-instructie in natuurlijke klassensituaties op het leesbegrip van leerlingen in basis- en voortgezet onderwijs (Grade 3-12). Leesstrategieën werden gedefinieerd als de mentale hulpmiddelen die een lezer doelbewust inzet om tekstbegrip te monitoren, repareren of versterken. De auteurs verdeelden ze onder in drie categorieën. Onder strategieën *voor het lezen* vallen activiteiten als voorspellen, het activeren van voorkennis en het stellen van leesdoelen. Strategieën *tijdens het lezen* bestaan uit jezelf vragen stellen tijdens het lezen, parafraseren, samenvatten en inferenties maken. Bij strategieën *na het lezen* gaat het om samenvatten en memoriseren. Behalve deze strategieën werd gekeken naar de inzet van graphic organizers, kennis van tekststructuur, mental imagery en expliciete monitoringstrategieën (bijv. reparatiestrategieën als herlezen). Hoe leesbegrip werd gemeten, werd—behalve dat er onderscheid werd gemaakt tussen gestandaardiseerde en niet-gestandaardiseerde toetsen—niet nader toegelicht. Okkinga et al. includeerden 52 (quasi-)experimentele studies die in totaal 152 effectgrootten omvatten;

---

<sup>9</sup> Hoewel ze geen precieze effectgrootten voor de verschillende leerjaren presenteerden, gaan we er op basis van de constatering dat de grootste effecten werden gevonden voor leerlingen in Grade 7 en 8 vanuit dat die het overall gemiddelde overstegen.

<sup>10</sup> Het *Hawthorne-effect* houdt in dat het gegeven *dat* mensen deelnemen aan onderzoek er al voor zorgt dat zij hun gedrag aanpassen.

het totaal aantal participanten werd niet gerapporteerd. De gemiddelde effecten waren significant en positief. Op gestandaardiseerde toetsen bereikten ze echter niet de drempel van een klein effect, noch op korte termijn ( $d = 0.19, p < .001$ ) noch op langere termijn ( $d = 0.17, p = .001$ ). Op niet-gestandaardiseerde tests was het effect op korte termijn klein ( $d = 0.43, p < .001$ ) en op langere termijn groot ( $d = 0.95, p < .001$ ). Voor gestandaardiseerde toetsen was leeftijdsgroep geen significante moderator, voor niet-gestandaardiseerde toetsen wel. In dat laatste geval werd het grootste effect gevonden bij leerlingen in Grade 6 tot en met 8 ( $d = 0.62$ ); in Grade 3-5 ( $d = 0.39$ ) en Grade 9-12 ( $d = 0.09$ ) waren de effecten kleiner. Okkinga et al. vonden daarnaast moderatoreffecten van interventiekenmerken (ze zagen bijvoorbeeld dat interventies waarin leerlingen leerden leesdoelen te stellen effectiever waren dan interventies waarin dat niet gebeurde), maar ze maakten geen uitsplitsing naar onderwijsfase. Ze voerden geen systematische quality appraisal uit, maar codeerden wel een aantal studiekenmerken, zoals de vraag of er sprake was van random toewijzing aan condities, de aard van de controlegroep en de vraag of interventie-effecten werden getest met multilevel-analyses.<sup>11</sup> Alleen de vraag of studies gepubliceerd waren, was een reden voor exclusie. Onder geïnccludeerde studies was er variatie in kwaliteit; in twee gevallen was die variatie gerelateerd aan effectgrootte: voor gestandaardiseerde toetsen bleken effecten kleiner wanneer er gebruik werd gemaakt van multilevel-analyses; voor niet-gestandaardiseerde toetsen bleken effecten groter wanneer er sprake was van een *business-as-usual*-controlegroep. Okkinga et al. vonden beperkte aanwijzingen voor publicatiebias, overigens zonder dat dit consequenties had voor hun conclusies.

## Leesgedrag en -houding

In drie meta-analyses werd gekeken naar leesgedrag: in twee van die meta-analyses werd de samenhang tussen leesgedrag en leesbegrip onderzocht (Altamura et al., 2025; Mol & Bus, 2011), terwijl één meta-analyse zich richtte op interventies om via de stimulering van leesgedrag het leesbegrip te bevorderen (Bus et al., 2024).

Altamura et al. (2025) testten het verband tussen leesbegrip en digitale leesgewoonten in de vrije tijd: de tijd die mensen besteden aan het lezen van digitale teksten voor sociale communicatie (bijv. e-mailen of gebruik van sociale media) of informatieverwerking (bijv. informatie opzoeken, een blog of een e-boek lezen). Leesbegrip werd gedefinieerd als letterlijk en/of inferentieel begrip en op verschillende manieren gemeten (bijv. met

---

<sup>11</sup> Vanwege de hiërarchische structuur van data verzameld in scholen.

meerkeuze-/open vragen over teksten en cloze tests). Altamura et al. analyseerden 26 correlatieve studies met in totaal 469.564 participanten tussen de 9 en 92 jaar oud. Van die participanten was bijna driekwart leerling in het voortgezet onderwijs: 36 procent was afkomstig van een middle school (Grade 6-8) en nog eens 36 procent van een high school (Grade 9-12). De auteurs vonden over leeftijdsgroepen heen een significante, positieve, maar marginale correlatie tussen leesbegrip en digitale leesgewoonten ( $r = .01, p = .03$ ). Onderwijsfase was een significante moderator: bij high school-leerlingen werd een positieve, eveneens marginale correlatie gevonden ( $r = .09$ ), terwijl die correlatie bij middle school-leerlingen negatief en marginaal was ( $r = -.03$ ). Altamura et al. voerden vervolgens een analyse uit waarbij alleen studies met basisschoolleerlingen en middelbare scholieren werden geïnccludeerd: ze vonden dat effectgrootten toenamen naarmate de gemiddelde leeftijd van de participanten steeg ( $\beta = 0.12, p = .03$ ). De auteurs concludeerden dat digitale leesgewoonten in de vrije tijd een beperkt positief verband hebben met leesbegrip en dat dit verband alleen geldt voor oudere en niet voor jongere leerlingen. Ze constateerden echter ook dat de gevonden correlaties aanzienlijk kleiner waren dan die tussen leesbegrip en papieren leesgewoonten in de vrije tijd (zie verder) en beschouwden die uitkomst als ondersteuning voor de zogenaamde *Shallowing Hypothesis* (digitaal lezen leidt tot oppervlakkiger lezen). Hoewel de auteurs geen standaardprocedure voor quality appraisal volgden, registreerden ze wel enkele kwaliteitsindicatoren (wijze van steekproeftrekking, betrouwbaarheid van de effectmaten). Studiekwaliteit was geen exclusiecriteria en, hoewel ze variatie lieten zien, bleken de kwaliteitsindicatoren niet gerelateerd aan effectgrootten. Ook vonden de auteurs geen aanwijzingen voor publicatiebias.

De overzichtsstudie van Mol en Bus (2011) bestaat uit drie meta-analyses naar het verband tussen lezen in de vrije tijd (print exposure) en leesvaardigheid in verschillende leeftijdsgroepen, variërend van kinderen in de voorschoolse periode tot studenten in het hoger onderwijs. Een van de meta-analyses had betrekking op leerlingen in basis- en voortgezet onderwijs (Grade 1-12). De 40 studies in deze meta-analyse omvatten in totaal 2.792 leerlingen; onduidelijk is echter hoe deze zijn verdeeld over de beide onderwijsfasen. Lezen in de vrije tijd werd gemeten met titel- of auteursherkenningstests: leerlingen moesten in lijsten met echte en verzonden boektitels of auteursnamen aangeven welke ze herkenden. De aanname was dat leerlingen die meer juiste titels en namen herkenden vaker lazen dan leerlingen die minder juiste titels en namen herkenden. Van de 40 studies hadden er 21 leesbegrip als uitkomst, dat meestal werd gemeten met gestandaardiseerde toetsen waarin leerlingen meerkeuze- of open vragen moesten beantwoorden over korte teksten of door

middel van cloze tests. Over het algemeen had lezen in de vrije tijd een gematigd positief verband met leesbegrip ( $r = .36, p < .001$ ): leerlingen die meer lasen in hun vrije tijd scoorden beter op leesbegripstests. Leeftijd en onderwijsfase (basisschool, middle school of high school) waren geen significante moderatoren: het positieve verband gold dus ook voor leerlingen in het voortgezet onderwijs. Op basis van de drie meta-analyses samen suggereerden Mol en Bus een proces van wederkerige beïnvloeding, waarbij lezen in de vrije tijd en leesvaardigheid elkaar over tijd versterken: de sterkte van de correlatie was namelijk groter wanneer participanten ouder waren. De auteurs volgden geen standaardprocedure voor quality appraisal en sloten geen studies uit op grond van mindere methodologische kwaliteit. Wel codeerden ze enkele kwaliteitskenmerken (design, betrouwbaarheid van instrumenten), maar ze rapporteerden die niet en voerden er ook geen moderatoranalyses mee uit. Mol en Bus analyseerden mogelijke publicatiebias en maakten op basis daarvan enkele aanpassingen in de schatting van effectgrootten.

Bus et al. (2024) onderzochten de effecten van *vrij lezen* op verschillende aspecten van de leesvaardigheid: behalve naar leesbegrip (op uiteenlopende manieren gemeten, maar vaak met een gestandaardiseerde test) keken ze naar technisch lezen (*word recognition*) en leesattitude. Ze omschreven vrij lezen als een korte activiteit (10-15 minuten per dag), aangeboden tijdens schooltijd, waarbij leerlingen in aanvulling op het reguliere leesonderwijs de gelegenheid krijgen te lezen, zonder dat ze worden beperkt in hun boekenkeuze of verantwoording hoeven af te leggen (bijv. doordat ze getoetst worden over wat ze hebben gelezen). De meta-analyse van Bus et al. omvatte 47 (quasi-)experimentele of *ex-post-facto*<sup>12</sup>-studies, afkomstig van 29 publicaties met in totaal 7.439 leerlingen. De meeste studies ( $k^{13} = 30$ ) hadden betrekking op basisschoolleerlingen; 17 studies zijn uitgevoerd bij leerlingen in het voortgezet onderwijs. Voor die laatste leerlingen was er geen significant effect van vrij lezen op leesbegrip ( $g = 0.12, p = .13$ ).<sup>14</sup> Bus et al. noemden drie mogelijke verklaringen voor het ontbreken van effecten. Allereerst is begrijpend lezen een complexe vaardigheid. De tijd die in geïncludeerde studies werd besteed aan vrij lezen (doorgaans zo'n 25 minuten per dag over een periode van drie maanden) was mogelijk te beperkt om leerlingen genoeg extra leeservaring op te laten doen. Daarnaast kan de keuze voor begripsmaten een rol hebben

---

<sup>12</sup> Studies waarin een vergelijking wordt gemaakt tussen bestaande groepen die wel en niet zijn blootgesteld aan een interventie zonder dat door middel van een pre-test hun vergelijkbaarheid bij aanvang van de interventie is vastgesteld.

<sup>13</sup> Met de letter  $k$  verwijzen we naar het aantal studies.

<sup>14</sup> Datzelfde gold overigens voor leerlingen in het basisonderwijs, hoewel het effect in het laatste geval wel significant kleiner was.

gespeeld: juist het complexe karakter van begrijpend lezen maakt het moeilijk om tests te kiezen die gevoelig zijn voor het meten van interventie-effecten. Ten slotte was in 12 studies vrij lezen niet de kern van de interventie: deze studies lieten een negatief gemiddeld effect zien ( $g = -0.24$ ). Bus et al. voerden een kwaliteitscontrole uit aan de hand van de *Cochrane*-indicatoren (ontbreken van aanvangsverschillen tussen condities, controle voor verstorende variabelen, meting van implementatiekwaliteit, hoeveelheid ontbrekende gegevens en gebruik van gevalideerde en betrouwbare meetinstrumenten). Ze vonden variatie in kwaliteit (hoewel die wel verschilde per indicator), maar sloten op basis daarvan geen studies uit. Ook was die variatie niet gerelateerd aan effectgrootte. Verder vonden de auteurs aanwijzingen voor publicatiebias, die duiden op een mogelijke onderschatting van interventie-effecten.

Leesmotivatie was het onderwerp van twee meta-analyses. Toste et al. (2020) onderzochten de correlatie tussen leesmotivatie en verschillende aspecten van leesvaardigheid. Eén van die aspecten was wat ze *meaning-focused skills* noemden, “the ability to comprehend and gain knowledge from reading text” (p. 424): een combinatie van leesbegrip en woordenschat. Hoe leesbegrip werd gemeten, werd niet verder toegelicht. Motivatie werd gedefinieerd als een construct met drie dimensies (Conradi et al., 2014):

1. *Goal orientation*. Hierbij gaat het om de doelen die lezers nastreven met lezen: ligt de focus op individuele ontwikkeling en beheersing (*mastery orientation*) of op prestatie in vergelijking met anderen (*performance orientation*)?
2. *Beliefs*. Hierbij gaat het om de overtuigingen die lezers hebben over hun leesvaardigheid of over het lezen zelf: hoeveel vertrouwen hebben zij in de eigen leesvaardigheid en hoe waarderen zij lezen als activiteit en de eigen ervaringen met lezen?
3. *Dispositions*. Hierbij gaat het om begrippen als *attitude* en *interesse*: welke gevoelens hebben lezers ten aanzien van lezen en zijn zij gemotiveerd om te lezen vanuit een persoonlijke belangstelling voor bepaalde onderwerpen?

Motivatie werd gemeten met zelfrapportage, beoordelingen door leraren of observaties door onderzoekers. De meta-analyse van Toste et al. (2020) bevatte 132 studies, met meer dan 690.000 participanten uit basis- en voortgezet onderwijs (Grade 1-12). Van de 1.154 correlaties waren er 338 afkomstig uit studies in middle schools (Grade 6-8) en 118 uit studies in high schools (Grade 9-12). De overall correlatie tussen motivatie en meaning-focused skills was significant en positief, maar klein ( $r = .19$ , 95% CI [.15, .23]). Die correlatie kon nog verder worden uitgesplitst naar leesbegrip; ook dat verband was klein ( $r =$

.20, 95% CI [.17, .23]). Onderwijsfase (basisschool, middle school, high school) bleek geen significante moderator van de correlatie tussen leesmotivatie en meaning-focused skills: het gevonden verband was dus ook geldig voor leerlingen in het voortgezet onderwijs.

Correlaties met beliefs en dispositions waren over het algemeen sterker dan die met goal orientations, maar of dat ook voor vo-leerlingen gold, werd niet berekend. Toste et al. maakten geen analyse van studiekwaliteit en ze sloten geen studies uit op basis van andere criteria dan publicatiestatus: ongepubliceerde studies werden buiten beschouwing gelaten. De auteurs vonden geen aanwijzingen dat publicatiebias de uitkomsten van de meta-analyse heeft beïnvloed.

Van der Sande et al. (2023)<sup>15</sup> onderzochten het effect van leesmotivatie-interventies op leesmotivatie en leesbegrip bij leerlingen in basis- en voortgezet onderwijs. Net als Toste et al. (2020) vertrokken de auteurs van Conradi et al.'s (2014) definitie van leesmotivatie als een multi-dimensioneel construct, bestaande uit doelen, overtuigingen en disposities. Ze maakten bovendien een onderverdeling in *affirming motivations* (motivaties die het lezen bevorderen, zoals intrinsieke motivatie en *self-efficacy*), *undermining motivations* (motivaties die het lezen belemmeren, zoals de neiging om lezen te vermijden en ervaren moeilijkheid) en extrinsieke motivatie (bijv. lezen om erkenning te krijgen van relevante anderen). Van der Sande et al. (2023) includeerden alleen studies die gekaderd waren in motivatietheorie en categoriseerden interventies op basis van het theoretische mechanisme dat ze aanspraken. Zo maakten ze onderscheid tussen interventies die leerlingen beoogden te motiveren door hun interesses aan te spreken, door gevoelens van autonomie, competentie en verbondenheid te stimuleren of door *mastery goals* (d.w.z. een focus op individuele ontwikkeling en beheersing) aan te wakkeren. De uiteindelijke dataset bestond uit 39 studies met 55 (quasi-)experimentele vergelijkingen; het totale aantal leerlingen werd niet gerapporteerd. Onderwijsfase bleek een significante moderator van interventie-effecten op leesbegrip: voor basisschoolleerlingen was het gemiddelde effect groter dan voor middelbare scholieren. Dat laatste effect was weliswaar significant, maar marginaal ( $d = 0.12$ ,  $SE = 0.03$ ). Een vergelijkbaar moderatoreffect was er overigens niet voor affirming motivations: hier waren effecten voor basisschoolleerlingen en middelbare scholieren nagenoeg gelijk (resp.  $d = 0.38$ ,  $SE = 0.08$  en  $d = 0.36$ ,  $SE = 0.09$ ). Van der Sande et al. vonden over alle studies heen

---

<sup>15</sup> In de expertraadpleging werd ook nog verwezen naar de meta-analyse van Van Steensel, Van der Sande et al. (2016). We hebben ervoor gekozen alleen de meta-analyse van Van der Sande et al. (2023) op te nemen: deze meta-analyse is uitgevoerd door dezelfde auteurs en vormt een ge-update en gepubliceerde versie van de meta-analyse uit 2016.

moderatoreffecten van interventiekenmerken (programma's sorteerden bijv. grotere effecten als ze aandacht besteedden aan de interesses van leerlingen), maar splitsten moderatoranalyses niet uit naar onderwijsfase. De auteurs volgden geen gestandaardiseerde procedure voor quality appraisal, maar rapporteerden wel over bepaalde kwaliteitsindicatoren zoals design (experimenteel versus quasi-experimenteel) en aard van de controlegroep. Ze excludeerden studies zonder controlegroep en studies die niet in peer-reviewed tijdschriften waren gepubliceerd. Ze constateerden variatie in studiekwaliteit, maar geen van de gebruikte indicatoren bleek gerelateerd aan effecten op leesbegrip. Er was enige indicatie voor publicatiebias, die duidde op een onderschatting van interventie-effecten.

## Kernmerken van de leeromgeving

**Digitale technologie.** Berkeley et al. (2015) analyseerden 27 studies die de effecten van digitale tekstinterventies op leesbegrip testten. De interventies werden onderverdeeld in twee categorieën. Bij de eerste categorie (*Basic access to text*) ging het om digitale bewerkingen waarmee de toegankelijkheid van teksten werd vergroot. Die bewerkingen hadden betrekking op de presentatie van teksten (bijv. de mogelijkheid om het font aan te passen) en op de mogelijkheid om te “vertalen” (bijv. tekst om te zetten in spraak of die te vereenvoudigen). De tweede categorie (*Instructional enhancements*) omvatte digitale aanpassingen waarmee het begrip van de tekst werd ondersteund. Het ging hierbij om aanpassingen waarmee de betekenis van tekstonderdelen werd geïllustreerd (denk aan het gebruik van geluid, video en animaties), om toevoegingen die hielpen om een overzicht van de belangrijkste elementen van een tekst te krijgen (bijv. inhoudsopgaven en *concept maps*) en om “instructieve” toevoegingen, dat wil zeggen: toevoegingen die leerlingen aanzetten tot actieve verwerking (bijv. prompts en vragen). Leesbegrip werd op verschillende manieren gemeten (bijv. via tests waarin leerlingen meerkeuze- of open vragen moesten beantwoorden over teksten, maze-tests en naverteltaken). De meta-analyse bevatte gegevens van in totaal 16.513 leerlingen, variërend van kleuters tot en met leerlingen in het voortgezet onderwijs; die laatste groep was het onderwerp van 12 studies. Van de 27 studies was ongeveer de helft ( $k = 13$ ) gericht op zich normaal ontwikkelende leerlingen; 14 studies waren uitsluitend gericht op leerlingen met leerproblemen.<sup>16</sup> Alle studies hadden een quasi-experimenteel design. Berkeley et al. vonden voor vo-leerlingen een klein negatief effect van Basic access to text-interventies ( $g = -0.39$ ), maar het aantal studies waarop ze zich baseerden was klein ( $k = 4$ ). Instructional

---

<sup>16</sup> In wezen past deze studie dus ook in de categorie “Meta-analyses: leerlingen met speciale onderwijsbehoeften”.

enhancement-interventies hadden bij dezelfde leeftijdsgroep juist een klein positief effect ( $g = 0.43$ ,  $k = 8$ ). De auteurs voerden een uitgebreide kwaliteitscheck uit, zij het niet gebaseerd op een gestandaardiseerde procedure. Ze sloten alleen studies uit op basis van design en, ondanks aanzienlijke variatie in studiekwaliteit, maakten ze geen analyse van de relatie tussen kwaliteit en effectgrootte. Evenmin onderzochten ze mogelijke publicatiebias.

Delgado et al. (2018) onderzochten de vraag of de aard van het medium waarvan gelezen wordt—papier of een (digitaal) scherm—effect had op leesbegrip, waarbij ze onder digitaal lezen het lezen op apparaten zoals computers, tablets, mobiele telefoons en e-readers verstonden. Tegen de achtergrond van de eerdergenoemde Shallowing Hypothesis waren de auteurs geïnteresseerd in de vraag of digitaal lezen samengaat met minder diepgaande verwerking van teksten. Delgado et al. analyseerden 53 studies met in totaal 169.524 participanten; 53 procent daarvan waren leerlingen uit het voortgezet onderwijs. Bij leesbegrip werd een onderscheid gemaakt tussen *textual comprehension* (oppervlakkig begrip waarbij de nadruk ligt op het ophalen van letterlijke [detail]informatie uit teksten) en *inferential, high-level comprehension* (diepgaand begrip waarbij de nadruk ligt op het maken van afleidingen), en tussen gestandaardiseerde en door onderzoekers ontwikkelde tests. Bij de analyse van de effecten voor middelbare scholieren werd daar echter geen nadere vergelijking tussen gemaakt. Onderwijsfase (basisonderwijs, voortgezet onderwijs, universiteit) bleek geen significante moderator: voor alle groepen was er een significant negatief effect van digitaal lezen op leesbegrip, hoewel dat voor vo-leerlingen marginaal was ( $g = -0.15$ , 95% CI [-0.29, -0.02]). Delgado et al. volgden geen vaste procedure voor quality appraisal. Wel codeerden ze drie aspecten van studiekwaliteit: steekproeftrekking, wijze van toewijzing aan condities en steekproefgrootte. Ze sloten geen studies uit op grond van kwaliteit, maar onderzochten wel of de geobserveerde variatie in wijze van toewijzing en steekproefgrootte interventie-effecten modereerden; dat bleek niet het geval. De auteurs vonden geen aanwijzingen voor publicatiebias.

## Combinatie van variabelen

Baye et al. (2019) onderzochten de effecten van tien verschillende leesinterventies op het leesbegrip en/of de woordenschatontwikkeling van leerlingen uit het voortgezet onderwijs (Grade 6-12):

1. Tutoring: frequente extra begeleiding in één-op-één-situaties of kleine groepjes door een leraar, paraprofessional, vrijwilliger of oudere medeleerling.

2. Coöperatief leren: leerlingen werken dagelijks samen in heterogene groepjes van vier of vijf, met als doel elkaar te helpen bij het leren, met name bij het toepassen van (metacognitieve) begripsstrategieën.
3. Schoolbrede interventies, waarbij het gaat om programma's waarin (cross-disciplinaire) lerarenteams worden geprofessionaliseerd en/of organisatorische aanpassingen worden gemaakt (bijv. het kiezen voor generieke lesmethoden).
4. Schrijfinterventies: instructie in specifieke schrijfvaardigheden of over het schrijfproces.
5. Inhoudsgerichte interventies: instructie in metacognitieve strategieën gericht op kennisverwerving binnen bepaalde vakken (bijv. mens- en maatschappij- of natuurwetenschappelijke vakken).
6. Woordenschatinterventies: aanleren van specifieke woorden of van strategieën om de betekenis van onbekende woorden te achterhalen.
7. Strategie-instructie, waarbij leerlingen stapsgewijze aanpakken krijgen aangeleerd om leesbegripsproblemen op te lossen. Het gaat bijvoorbeeld om instructie van strategieën zoals samenvatten, vragen stellen, hoofdgedachte achterhalen en het gebruik van graphic organizers.
8. Personalisatie, waarbij instructie wordt aangepast aan het niveau, de interesses of andere kenmerken van individuele leerlingen.
9. Afwisseling van leraargestuurde en gepersonaliseerde instructie.
10. Intensieve groepsaanpakken: remediëring, vaak gericht op decoderen of vloeiend lezen.

Uitsluitend effecten gemeten met gestandaardiseerde maten werden meegenomen. Baye et al. analyseerden de uitkomsten van 69 (quasi-)experimentele studies. Over het algemeen vonden ze een significant, positief maar marginaal interventie-effect ( $g = 0.09, p < .001$ ). Een vergelijking van de verschillende interventietypen liet zien dat er alleen voor tutoring (en dan uitsluitend in één-op-één-situaties) een significant effect optrad dat ten minste klein was ( $g = 0.24, p < .001$ ). De auteurs codeerden aspecten van de kwaliteit van de primaire studies, maar maakten daarvoor geen gebruik van een gestandaardiseerde procedure. Studiekwaliteit was een inclusie criterium: studies die niet (quasi-)experimenteel waren en studies waarvan de pretestverschillen te groot waren, werden geëxcludeerd. De studiekwaliteit was daardoor sterk: zo werd in 90 procent van de studies random toewijzing toegepast; die variabele was

overigens niet gerelateerd aan interventie-effecten. Hoewel ongepubliceerde studies werden meegenomen in de selectie, besteedden de auteurs geen aandacht aan publicatiebias.

## Meta-analyses: leerlingen met speciale onderwijsbehoeften

### Kennis en vaardigheden

**Zelfregulatie, metacognitie en leesstrategieën.** Cure et al. (2021) onderzochten het effect van één specifieke leesstrategie (*story mapping*) op het begrip van narratieve teksten bij 8- tot 14-jarigen met een beperking. De auteurs definieerden *story mapping* als een grafische organisatiestrategie waarbij onderdelen van een verhaal (de zogenaamde *story grammar*) en de verbanden daartussen visueel gepresenteerd worden. Leerlingen leerden deze strategie aan door middel van een combinatie van expert-modeling en begeleide inoefening. De meting van narratief begrip bestond in alle gevallen uit het beantwoorden van vragen over teksten, hoewel de aard van de vragen verschilde (meerkeuze versus open, schriftelijk versus mondeling). De meta-analyse omvatte zeven single-case studies met in totaal 30 participanten, waarvan er negen afkomstig waren van een middle school; de overige waren basisschoolleerlingen. De beperkingen die de participanten hadden, liepen uiteen: ze varieerden van autisme-spectrumstoornissen tot specifieke leerbeperkingen, gedragsstoornissen, spraak-/taalstoornissen en gezondheidsbeperkingen. Voor middle schoolleerlingen was het effect van *story mapping* significant, positief en groot (Tau-U = 1.0, 95% CI [82.8%, 100%]). Cure et al. voerden een quality appraisal uit op basis van standaarden ontwikkeld door Kratochwill et al. (2013) en Horner et al. (2005), die onder meer betrekking hadden op design, betrouwbaarheid van metingen en het aantal beschikbare datapunten. Studies die niet aan de standaarden voldeden, werden uitgesloten. De geïncludeerde studies varieerden in kwaliteit (bij 5 van de geïncludeerde studies was enige reserve), maar de auteurs gingen niet na of die variatie gerelateerd was aan interventie-effecten. Ook maakten ze geen analyse van mogelijke publicatiebias.

Peng et al. (2023) testten het effect van verschillende typen leesstrategie-interventies op het leesbegrip van zwakke lezers in basis- en voortgezet onderwijs (Grade 3-12). Ze maakten onderscheid tussen interventies die zich richtten op (1) het achterhalen van de hoofdgedachte van een tekst of samenvatten, (2) het gebruik van tekststructuurkenmerken, (3) het doen van voorspellingen over tekstinhoud op basis van achtergrondkennis en het evalueren van die voorspellingen, (4) het navertellen van teksten, (5) het maken van inferenties door tekstonderdelen te integreren of tekstuele informatie te integreren met achtergrondkennis, (5) het gebruik van graphic organizers en (6) *self-monitoring*: het

monitoren van begrip, bijvoorbeeld door jezelf tijdens het lezen vragen te stellen. Leesbegrip werd vastgesteld met zowel gestandaardiseerde toetsen als door onderzoekers ontwikkelde toetsen en door middel van verschillende toetswijzen (*passage comprehension*, naverteltaken, samenvattingen, inferentievragen en woordenschattaken). De meta-analyse omvatte 52 (quasi-)experimenten; het totaal aantal participanten werd niet genoemd. Op basis van een zogenaamde *Bayesian Network Meta-Analysis*<sup>17</sup> concludeerden Peng et al. dat de combinatie van drie strategieën (hoofdgedachte achterhalen, tekststructuurkenmerken gebruiken en navertellen) de sterkste bijdrage leverde aan leesbegrip en dat dit effect alleen gold als er in interventies aandacht was voor achtergrondkennis. Leeftijd en leerjaar bleken geen significante moderatoren, wat betekent dat de overall conclusie ook geldig is voor middelbare scholieren. De auteurs maakten geen gebruik van een vaste procedure voor quality appraisal. Wel evalueerden ze studiekwaliteit aan de hand van design, aard van de contolegroep en implementatiekwaliteit. Design was een in-/exclusiecriteria: alleen (quasi-)experimenten werden opgenomen. De geobserveerde variatie in studiekwaliteit bleek niet gerelateerd aan interventie-effecten. Bij een *sensitivity*-analyse zonder studies met lage kwaliteit convergeerde het statistische model niet: in de uiteindelijke analyse zijn dus ook studies met lagere kwaliteit opgenomen. Peng et al. vonden geen aanwijzingen voor publicatiebias.

## Kenmerken van de leeromgeving

**Digitale technologie.** Wood et al. (2018) onderzochten het effect van *text-to-speech*-technologie en vergelijkbare hardop-leesinterventies op het leesbegrip van leerlingen in basis- en voortgezet onderwijs (Grade 3-12). Bij *text-to-speech*-technologie gaat het om interventies waarbij een tekst, ter ondersteuning van het begrip, wordt voorgelezen door een computerstem of een opgenomen menselijke stem. Leesbegrip werd gedefinieerd als betekenisconstructie op het niveau van zinnen, alinea's of grotere tekstonderdelen en gemeten met zowel gestandaardiseerde toetsen als toetsen ontwikkeld door onderzoekers. De meta-analyse omvatte 22 studies met in totaal 2.942 leerlingen. Alle leerlingen hadden dyslexie, een andere leesstoornis of een leerbeperking op het gebied van lezen. Over alle leeftijdsgroepen heen had het gebruik van de onderzochte hulpmiddelen een significant, positief en klein effect op leesbegrip ( $g = 0.35, p < .01$ ). Leerjaar was geen moderator van effectgrootte: het positieve effect gold dus ook voor leerlingen in het voortgezet onderwijs.

---

<sup>17</sup> Een voordeel van *Bayesian Network Meta-Analysis* is dat de interactie tussen verschillende interventiecomponenten ermee kan worden geanalyseerd, zodat kan worden vastgesteld welke *combinatie* van kenmerken het beste werkt.

Hoewel Wood et al. geen systematische procedure voor quality appraisal volgden, registreerden ze wel enkele kenmerken van studiekwaliteit (publicatiestatus, random toewijzing, uitval en het ontbreken van aanvangsverschillen tussen condities). De auteurs gebruikten studiekwaliteit niet als exclusie criterium. Ook rapporteerden ze niet of nauwelijks over variatie in studiekwaliteit of de relatie ervan met effectgrootten. Wel vonden ze enige indicatie voor publicatiebias.

## Combinatie van variabelen

Edmonds et al. (2009) onderzochten de effecten van diverse typen interventies op het leesbegrip van zwakke lezers in het voortgezet onderwijs (Grade 6-12). De interventies hadden betrekking op verschillende niveaus van tekstverwerking: het decoderen van woorden (*word study*), vloeiend lezen, woordenschat, (tekst)begrip of een combinatie daarvan; bij begripsinterventies ging het in veel gevallen om leesstrategie-instructie. In de meeste studies werd leesbegrip op verschillende manieren gemeten (o.a. via meerkeuze-/open vragen over teksten, samenvatten, recall-taken en naverteltaken). In de meta-analyse werden die verschillende maten samengenomen; wel werd er een onderscheid gemaakt tussen gestandaardiseerde en niet-gestandaardiseerde toetsen. Edmonds et al. includeerden 29 studies met in totaal 976 participanten, waarvan er 19 betrekking hadden op middle school-leerlingen, vijf op high school-leerlingen en twee op een combinatie daarvan. Alle steekproeven bestonden volledig uit zwakke lezers, maar de kenmerken van die lezers verschilden: in 17 studies hadden leerlingen een gediagnosticeerde leer- of leesstoornis, in de overige 12 studies hadden leerlingen een zwakke leesvaardigheid, maar was er geen sprake van een diagnose of ging het om een combinatie van gediagnosticeerde en niet-gediagnosticeerde leerlingen. Het gezamenlijke effect van de interventies was significant, positief en groot ( $d = 0.89$ , 95% CI [0.42, 1.36]). Wanneer alleen naar gestandaardiseerde toetsen werd gekeken, was het effect nog altijd significant, maar kleiner ( $d = 0.47$ , 95% CI [0.12, 0.82]). Daarnaast vonden de auteurs alleen significante effecten van begripsinterventies en gecombineerde interventies (respectievelijk  $d = 1.23$ , 95% CI [0.96, 1.50] en  $d = 0.72$ , 95% CI [0.45, 0.99]), niet van *word study*- en *fluency*-interventies. Twee moderatoreffecten waren significant: effecten bleken groter wanneer de interventie werd uitgevoerd door een onderzoeker ( $d = 1.15$ ) dan wanneer die werd uitgevoerd door een leraar ( $d = 0.77$ ) en wanneer de steekproef uitsluitend bestond uit leerlingen met een gediagnosticeerde leer- of leesbeperking ( $d = 1.50$ ) dan wanneer die bestond uit zwakke lezers zonder diagnose ( $d = 0.45$ ) of uit een combinatie van leerlingen ( $d = 0.68$ ). Bij de quality appraisal volgden de

auteurs de eerdergenoemde WWC-richtlijnen. Maar twee studies voldeden aan alle richtlijnen (random toewijzing, rapportage van implementatiekwaliteit en gebruik van gestandaardiseerde instrumenten). Er werden geen studies uitgesloten op grond van studiekwaliteit. Ook maakten Edmonds et al. geen analyse van de relatie tussen studiekwaliteit en effectgrootte of testten ze mogelijke publicatiebias.

Kaldenberg et al. (2015) onderzochten het effect van woordenschat- en leesbegripsinterventies op het begrip van natuurwetenschappelijke teksten bij leerlingen uit Grade 5 tot en met 11. Bij woordenschatinterventies ging het om de expliciete instructie van woordbetekenissen voorafgaand aan het lezen van een tekst met als doel die tekst beter te begrijpen. Bij begripsinterventies ging het om het aanleren van strategieën die losstonden van specifieke teksten, zoals het gebruik van een graphic organizer of het herkennen van de manier waarop alinea's zijn gestructureerd. Leesbegrip werd gemeten met begripsvragen over teksten, cloze- en recall-taken. De auteurs analyseerden 20 studies uit 12 publicaties; met een gemiddelde steekproefgrootte van 76 was het totaal aantal leerlingen ongeveer 1.500. Eén van de inclusiecriteria was dat de steekproef ten minste één leerling bevatte met een gediagnosticeerde leerstoornis, maar in de meeste gevallen (9 van de 12) bestond de steekproef volledig uit zulke leerlingen. Over alle interventies heen was het effect significant, positief en groot ( $g = 0.98$ ,  $SE = 0.15$ ). Kaldenberg et al. testten ook retentie-effecten: die waren significant, positief en middelgroot ( $g = 0.69$ ,  $SE = 0.21$ ). Er was een moderatoreffect van interventietype: de woordenschatinterventies sorteerden significant grotere effecten dan de begripsinterventies ( $g = 1.25$ ), hoewel de effectgrootte ook in het tweede geval nog altijd substantieel was ( $g = 0.64$ ). De auteurs voerden een quality appraisal uit volgens de *Council for Exceptional Children (CEC) Standards for Evidence-Based Practices in Special Education* (Cook et al., 2014). Studiekwaliteit varieerde per standaard: vaak was die goed, maar slechts weinig studies rapporteerden bijvoorbeeld implementatiekwaliteit. Studies werden alleen uitgesloten op grond van design: er werden uitsluitend (quasi-)experimenten geïncludeerd. Geen van de indicatoren voor studiekwaliteit bleek gerelateerd aan effectgrootte. De auteurs rapporteerden geen analyse van publicatiebias.

Sohn et al. (2023) onderzochten de effecten van leesbegripsinterventies bij middelbare scholieren (Grade 6-12) met leesproblemen (d.w.z. leerlingen met een gediagnosticeerde leesstoornis of leerlingen zonder een diagnose, maar met achterblijvende leesvaardigheid). Ze maakten een onderscheid tussen drie typen interventies:

1. (Meta)cognitieve strategie-interventies, waarbij het ofwel gaat om instructie in (metacognitieve) strategieën waarmee leerlingen bewust worden gemaakt van hun

denkprocessen tijdens het lezen en waarmee ze leren die processen te monitoren ofwel om instructie in specifieke (cognitieve) leesstrategieën zoals samenvatten.

2. Meervoudige interventies: een combinatie van instructie gericht op *phonics* (aanleren van klank-lettercombinaties), vloeiend lezen, kennis van tekststructuur, woordenschat en/of achtergrondkennis.
3. Overige interventies: interventies die zich richtten op uitsluitend phonics-instructie of op peer mediation of *computer-adapted instruction*; die laatste twee interventietypen werden overigens niet verder beschreven.

De meting van leesbegrip werd niet nader toegelicht dan dat er onderscheid werd gemaakt tussen gestandaardiseerde toetsen en toetsen die door onderzoekers waren ontwikkeld. Sohn et al. analyseerden 37 (quasi-)experimentele studies met in totaal 4.136 leerlingen. Van deze leerlingen was 61 procent afkomstig van een middle school en 34 procent van een high school; de overige leerlingen maakten deel uit van een gemengde steekproef. Tezamen hadden de leesinterventies een significant, positief en middelgroot effect op leesbegrip ( $g = 0.63, p < .001$ ). De auteurs vonden een moderatoreffect van interventietype: de (meta)cognitieve strategie-interventies hadden een significant groter effect dan de andere interventies samen ( $g = 1.17, p < .001; \beta = .72, p = .001$ ). Voor de meervoudige interventies was het effect juist kleiner dan dat van de andere interventies samen ( $\beta = -.57, p = .002$ ), hoewel ook het effect van de meervoudige interventies nog altijd significant en positief was ( $g = 0.31, p < .001$ ). Omdat de categorie “overig” maar vier studies bevatte, kon daarvoor het verschil met de andere interventies niet betrouwbaar worden vastgesteld. Daarnaast stelden Sohn et al. moderatoreffecten vast van:

- Tekstinhoud: als die was gericht op één vak (hier: Engelse taal en literatuur, mens- en maatschappijvakken) waren de effecten groter dan wanneer die was gericht op meerdere vakken.
- Interventieduur: kortere interventies waren effectiever dan langere.
- Uitvoerder: interventies uitgevoerd door onderzoekers waren effectiever dan die uitgevoerd door leraren.
- Type leerlingen: studies gericht op leerlingen met een diagnose lieten grotere effecten zien dan die gericht op leerlingen zonder diagnose.
- Type effectmaat: gestandaardiseerde toetsen lieten kleinere effecten zien dan toetsen ontwikkeld door onderzoekers.

Net als Kaldenberg et al. (2015) voerden Sohn et al. (2023) een quality appraisal uit op basis van de CEC-criteria. Daarnaast registreerden ze of studies gebruikmaakten van een *randomized controlled trial* (RCT), of implementatiekwaliteit werd gemeten en of er in de effectanalyse gecorrigeerd werd voor aanvangsverschillen tussen condities. Sohn et al. sloten studies uit wanneer ze geen (quasi-)experimentele opzet hadden of wanneer ze ongepubliceerd waren. Van de 37 geïncludeerde studies werden er 19 gekwalificeerd als van hoge kwaliteit; die studies lieten kleinere effecten zien dan studies met een lagere kwaliteit. De auteurs vonden geen aanwijzingen voor publicatiebias.

## Systematische reviews: zich normaal ontwikkelende leerlingen

### Kennis en vaardigheden

**Inferenties maken/achtergrondkennis.** Hattan et al. (2024) onderzochten de effecten van verschillende manieren om achtergrondkennis te activeren op het leesbegrip van leerlingen in het basis- en voortgezet onderwijs en studenten in het hoger onderwijs. Ze onderscheidde acht technieken:

1. *Open-ended prompts*, waarbij het meestal ging om vragen die worden gesteld voordat leerlingen een tekst lezen en waarmee ze worden aangemoedigd om alles te delen wat ze al over het onderwerp van die tekst weten.
2. *Procedurele/strategische ondersteuning tijdens het lezen*, waarbij leerlingen tijdens het lezen bijvoorbeeld vragen of prompts krijgen (van hun leraar of via de computer) om hen aan te moedigen een tekst diepgaander te verwerken.
3. *Visuele representaties*, waarbij leerlingen wordt gevraagd een grafische weergave (bijv. een concept map) te maken van hun kennis over het onderwerp van een tekst.
4. *Analogical reasoning*, waarbij vergelijkingen worden gepresenteerd tussen nieuwe en bestaande kennis of waarbij leerlingen wordt gevraagd om zelf zulke vergelijkingen te maken.
5. *Tekstuele aanpassingen*, waarbij het bijvoorbeeld gaat om het toevoegen van informatie of *advance organizers*, die leerlingen kunnen helpen hun achtergrondkennis te activeren en zo de tekst beter te begrijpen.
6. *Augmented activation*, waarbij leerlingen erop worden geattendeerd dat wat ze gaan lezen over een onderwerp mogelijk verschilt van wat ze al weten over dat onderwerp.
7. *Extratextual activities*, een brede categorie waaronder bijvoorbeeld activiteiten vallen waarmee het onderwerp van een tekst tastbaar wordt gemaakt (denk aan proefjes),

maar ook om het gebruik van verhelderende illustraties voorafgaand aan het lezen van een tekst of van een pre-test om achtergrondkennis te activeren.

8. *Spontaneous activation*, waarbij het ging om de inzet van activiteiten waarbij “spontaan” achtergrondkennis wordt geactiveerd, zoals hardop denken.

Leesbegrip werd op verschillende manieren gemeten (bijv. met meerkeuze-/open vragen over teksten, zinsverificatietaken, woordenschattaken, essays, recall-taken en leessnelheid). Hattan et al. analyseerden 88 technieken afkomstig uit 54 studies; het totaal aantal participanten werd niet genoemd. Van de 88 technieken werden er 38 ingezet voor leerlingen in het voortgezet onderwijs (middle school:  $k = 18$ ; high school:  $k = 20$ ). Voor 51 van de 88 technieken vonden de auteurs positieve effecten op leesbegrip. Voor middelbare scholieren leken drie technieken met name effectief: de inzet van open-ended prompts, visuele representaties en tekstuele aanpassingen. De auteurs maakten geen analyse van studiekwaliteit en sloten dus ook geen studies uit op basis van kwaliteit. Ook maakten ze geen analyse van mogelijke publicatiebias.

Sulaiman et al. (2023) onderzochten het effect van een specifieke methode—*Listen-Read-Discuss* (LRD)—op het leesbegrip van leerlingen in (senior) high schools. LRD is een techniek die er onder andere uit bestaat dat leerlingen voorafgaand aan het lezen van een tekst kennis opbouwen. De aanpak begint met een brainstorm over het onderwerp van de tekst, waarna de leraar uitleg geeft over de tekst (Sulaiman et al. spreken van een *lecture*, overigens zonder aan te geven waaruit die bestaat), de leerlingen de tekst lezen en er daarna in groepjes over in gesprek gaan. De les eindigt met een begripscheck en reflectie. Hoe leesbegrip werd gemeten, werd niet toegelicht. Sulaiman et al. analyseerden vijf studies: twee (quasi-)experimenten, twee *Classroom Action Research*-studies (niet nader toegelicht) en een case study. De auteurs constateerden dat in alle studies het leesbegrip tijdens de interventieperiode vooruit was gegaan, hoewel die conclusie niet in alle gevallen was gebaseerd op kwantitatieve data. Ook werd maar in de beschrijving van één experiment duidelijk dat de groei in de experimentele conditie sterker was dan die in de controleconditie. Suleiman et al. maakten geen analyse van studiekwaliteit of van mogelijke publicatiebias. Ze sloten alleen ongepubliceerde studies uit.

**Integreren van meerdere teksten.** Barzilai et al. (2018) onderzochten of interventies gericht op intertekstuele integratie—het verbinden, combineren of organiseren van informatie uit verschillende teksten—bijdroegen aan leesbegrip. Ze onderscheidden negen typen interventies: (1) expliciete instructie, (2) modeling, (3) de inzet van prompts (vragen,

aanwijzingen of reminders om leerlingen aan te moedigen teksten te integreren), (4) annoteren/samenvatten van individuele teksten, (5) het gebruik van graphic organizers (visuele representaties zoals tabellen of grafieken), (6) *collaborative discussions and practice* (in duo's of groepjes discussiëren over of taken maken n.a.v. meerdere teksten), (7) individuele oefening, (8) feedback en (9) "overige" interventies. De meting van leesbegrip had betrekking op het integreren van meerdere teksten; dat werd op verschillende manieren gedaan: met meerkeuze- of open vragen over meerdere teksten, essays, het beoordelen van stellingen gebaseerd op meerdere teksten, hardop denken, observaties of de analyse van gesprekken tijdens de taakuitvoering. Barzilai et al. analyseerden 61 studies: het merendeel waren (quasi-)experimenten (74%), maar de review bevatte ook descriptieve studies en een ontwerpstudie. Bijna 60 procent van de studies werd uitgevoerd in het voortgezet onderwijs (onderbouw: 31%, bovenbouw: 28%). Van de tien interventies in de basisschool en de onderbouw van het vo (samengenomen als één categorie) waren de meeste (7) effectief, de drie interventies in de bovenbouw van het vo waren allemaal effectief. De auteurs concludeerden dat de meeste effectieve interventies werden gekenmerkt door expliciete instructie, gezamenlijke discussies en oefening, graphic organizers, modeling en individuele oefening, hoewel er geen uitsplitsing werd gemaakt voor vo-leerlingen. De auteurs maakten geen analyse van studiekwaliteit of van mogelijke publicatiebias en sloten alleen ongepubliceerde studies uit.

**Zelfregulatie, metacognitie en leesstrategieën.** Lan et al. (2014) onderzochten het effect van metacognitieve (strategie-)instructie aangeboden via de computer op het leesbegrip van leerlingen in basis- en voortgezet onderwijs en studenten in het hoger onderwijs. De auteurs verdeelden de onderzochte strategieën onder in drie categorieën:

1. Regulatie: het kunnen plannen, monitoren en evalueren van het eigen leesproces.
2. Strategie-aanwijzingen (bijv. annotaties) in combinatie met hardop denken.
3. Woordenschat- en begripsondersteuning.

Daarnaast werd in een aantal studies het effect van medium onderzocht: Lan et al. gingen na of het voor het begrip uitmaakte of teksten op papier of op een computer werden aangeboden. De auteurs gingen niet in op de meting van leesbegrip; wel leken ze een onderscheid te maken tussen gestandaardiseerde en door onderzoekers ontwikkelde toetsen. De review omvatte 17 (quasi-)experimentele studies, afkomstig uit 14 publicaties met in totaal 1.210 participanten. In de zeven studies die zich richtten op regulatie werden zowel voor vo-leerlingen (Grade 6 en 7) als voor universitaire studenten positieve resultaten gevonden. Ook

in de studies naar woordenschat- en begripsondersteuning werden over het algemeen gunstige uitkomsten gevonden, nu bij leerlingen uit Grade 4 en 5 en bij high-school-leerlingen. Er waren geen studies met middelbare scholieren waarin strategie-aanwijzingen en hardop denken werden gecombineerd. Daarnaast was er geen effect van medium. Lan et al. maakten geen systematische analyse van studiekwaliteit. Wel sloten ze studies uit op basis van design en publicatiestatus: alleen (quasi-)experimentele studies uit peer-reviewed tijdschriften werden geïnccludeerd. Verder werd er niet naar mogelijke publicatiebias gekeken.

## Leesgedrag en -houding

Anderson et al. (2023) gingen na (a) in hoeverre leesinterventies voor leerlingen in het voortgezet onderwijs (Grade 6-12) instructieve activiteiten bevatten die hun engagement of (lees)betrokkenheid beoogden te stimuleren; en (b) of er, in de studies die de leesbetrokkenheid van leerlingen maten, een relatie was tussen die betrokkenheid en hun leesbegrip.<sup>18</sup> De review van Anderson et al. bestond uit 95 (quasi-)experimentele studies afkomstig uit 91 publicaties. De auteurs maakten onderscheid tussen 11 typen instructieve activiteiten<sup>19</sup> en vonden dat strategie-instructie de meest gebruikte manier was om betrokkenheid te bevorderen: die werd in 72 studies (76%) toegepast. In maar 16 van de studies werd betrokkenheid bekeken: gedragsmatige betrokkenheid (vaak gemeten met observaties) werd het vaakst onderzocht ( $k = 13$ ), cognitieve en academische betrokkenheid beduidend minder (resp.  $k = 3$  en  $2$ ). Van de 16 studies waren er acht waarin het verband tussen betrokkenheid en cognitieve uitkomsten werd geanalyseerd. Van deze studies vonden er zeven een verband tussen betrokkenheid en leesbegrip; onduidelijk is overigens om welke dimensie(s) van betrokkenheid het ging. De auteurs maakten geen systematische analyse van studiekwaliteit of mogelijke publicatiebias. Wel werden studies uitgesloten wanneer ze niet in een peer-reviewed tijdschrift waren gepubliceerd en op basis van design: alleen (quasi-)experimenten werden geïnccludeerd.

---

<sup>18</sup> In deze studie werd dus niet onderzocht of instructie gericht op leesbetrokkenheid bijdroeg aan leesbegrip, maar of er een relatie was tussen betrokkenheid en begrip.

<sup>19</sup> Die 11 categorieën waren: (1) formuleren van leer-/kennisdoelen, (2) *real-world interactions* (activiteiten waarmee thema's uit een tekst tastbaar worden gemaakt, zoals excursies), (3) ondersteunen van de autonomie (leerlingen keuzes bieden), (4) bieden van interessante teksten, (5) bieden van relevantie (bijv. door teksten te kiezen die aansluiten bij de ervaringen van leerlingen), (6) strategie-instructie, (7) samenwerking stimuleren, (8) complimenteren, belonen en feedback geven, (9) evaluatie (het werk van leerlingen beoordelen op een manier die aansluit bij hun leerdoelen), (10) inzet van technologie en (11) "overige" activiteiten (leerlingen aanmoedigen om met expressie te lezen).

## Kenmerken van de leeromgeving

**Digitale technologie.** Hare et al. (2024) onderzochten voor kinderen en adolescenten (leeftijdssrange: 1-17 jaar) het effect van medium op leesbegrip, leesbetrokkenheid en andere variabelen, zoals woordenschat en leessnelheid. In de 121 geïncludeerde studies werd een vergelijking gemaakt tussen lezen van papier en in een digitaal format (d.w.z. via apparaten zoals computers, tablets, telefoons en e-readers). Hoe leesbegrip werd gemeten, werd niet toegelicht. Ook gaven de auteurs geen informatie over de aard van de gebruikte designs. Over het algemeen vonden Hare et al. geen duidelijke aanwijzingen voor een mediumeffect op leesbegrip. Een uitzondering waren leerlingen in Grade 1-3: daar ging digitaal lezen doorgaans samen met beter begrip. In maar tien studies werd gekeken naar effecten op het leesbegrip van middelbare scholieren; in alle gevallen ging het om leerlingen van high schools. In vier van die studies werd een positief effect gevonden van lezen van papier, in vijf studies waren er geen verschillen tussen de beide typen media en in één studie was het effect afhankelijk van een andere variabele (overigens zonder dat werd aangegeven welke). Hare et al. sloten geen studies uit op grond van kwaliteit. Wel maakten ze gebruik van een gestandaardiseerde procedure voor quality appraisal, de *Quality Assessment Tool for Studies with Diverse Designs* (QATSDD), op basis waarvan ze een onderscheid maakten tussen studies van lage, middelmatige en hoge kwaliteit (resp.  $k = 21$ , 20 en 23). De auteurs maakten geen analyse van mogelijke publicatiebias.

Ter Beek et al. (2018) onderzochten de effecten van digitale *scaffolds* op het begrip van informatieve teksten bij middelbare scholieren (Grade 6-11). De auteurs definieerden deze scaffolds als hulpmiddelen, strategieën en handreikingen die leerlingen kunnen ondersteunen bij het reguleren van hun leren. Die ondersteuning had betrekking op drie typen processen:

1. Cognitieve processen. Scaffolds zijn dan vooral bedoeld om ervoor te zorgen dat leerlingen een leestaak zelfstandig tot een goed einde kunnen brengen, bijvoorbeeld door ze extra informatie te geven over de te lezen materialen.
2. Metacognitieve processen. Scaffolds zijn dan vooral bedoeld om ervoor te zorgen dat leerlingen hun leesproces kunnen plannen, monitoren en/of evalueren.
3. Motivationale processen. Scaffolds zijn dan vooral bedoeld om de interesse van leerlingen voor en hun emoties of gevoel van controle over de leestaak te vergroten.

Leesbegrip werd gemeten met zowel gestandaardiseerde als door onderzoekers ontwikkelde taken. De review was gebaseerd op vijf (quasi-)experimentele studies met in totaal 962 leerlingen. De effecten varieerden van marginaal tot groot ( $d = 0.02-1.23$ ); de auteurs concludeerden dan ook dat het niet goed mogelijk is om een duidelijk antwoord te geven op de vraag welke scaffolds precies bijdragen aan het leesbegrip. Ter Beek et al. maakten geen analyse van studiekwaliteit of mogelijke publicatiebias. Wel werden studies uitgesloten op basis van design: alleen (quasi-)experimenten werden geïnccludeerd.

**Vakintegratie.** McCulley en Osman (2015) onderzochten het effect van leesinstructie tijdens *social studies* op het leesbegrip en de vakkennis van leerlingen in middle en high school; *social studies* is vergelijkbaar met een vak als *mens en maatschappij* of *maatschappijleer* in Nederland. De wijze waarop de leesinstructie werd vormgegeven, varieerde: de auteurs onderscheidden woordenschatinstructie, strategie-instructie, instructie in kritisch lezen, ondersteuning bij het ordenen van achtergrondkennis en nieuwe kennis, *reciprocal teaching* en samenwerkend lezen. Bij de meting van leesbegrip werd alleen onderscheid gemaakt tussen gestandaardiseerde en door onderzoekers ontwikkelde toetsen. De review bestond uit 12 (quasi-)experimentele studies met in totaal 3.448 leerlingen uit Grade 6-12. Vier van de studies hadden leesbegrip als uitkomstmaat, zeven studies onderzochten (ook) effecten op vakkennis en twee studies keken (ook) naar andere vaardigheden (algemene leesvaardigheid en schrijfvaardigheid). McCulley en Osman concludeerden dat de integratie van leesinstructie in *social studies* gunstige effecten had voor zowel vakkennis als leesbegrip. In het laatste geval vonden de auteurs maar één marginaal effect ( $d = -0.02$ ); alle andere effecten passeerden ten minste de grens van een klein effect, met een uitschieter naar een groot effect ( $d = 1.14$ ). Voor de beoordeling van studiekwaliteit maakten de auteurs gebruik van onderdelen van een gestandaardiseerde procedure (de WWC-richtlijnen) en codeerschema's uit eerdere meta-analyses, waarbij ze onder meer keken naar design, toewijzing aan condities en de registratie van implementatiekwaliteit. Studies werden uitgesloten als een controlegroep ontbrak. Onder de geïnccludeerde studies was er variatie in kwaliteit: zo maakten vijf studies geen gebruik van random toewijzing en ontbrak in veel gevallen informatie over implementatiekwaliteit. De auteurs maakten geen analyse van mogelijke publicatiebias.

Ook Wade (1983) testte het effect van leesinstructie tijdens *social studies* in het voortgezet onderwijs (Grade 6-12), maar dan uitsluitend op het begrip van teksten die binnen dat vak werden gebruikt. De auteur sprak van een *functional approach*, waarin vakleerkrachten expliciete leesinstructie geven tijdens vaklessen op de momenten dat die

instructie noodzakelijk is. Ze maakte onderscheid tussen meervoudige interventies, waarin diverse doelen en lesmethoden werden gecombineerd, en enkelvoudige interventies, waarin maar één type instructie werd aangeboden. Ook de inhoud van dat laatste type varieerde; het ging om *comprehension aids* (zoals advance organizers), uitleg over de structuur van alinea's, het aanleren van concepten die belangrijk zijn voor het begrip van een tekst, *programmed materials* (d.w.z. uitgewerkte voorbeelden) en gesproken opnamen van teksten die leerlingen kunnen meelesen. Leesbegrip werd gemeten met gestandaardiseerde en door onderzoekers ontwikkelde toetsen. Wade analyseerde 29 studies: 18 (quasi-)experimenten (met in totaal 11.793 leerlingen) en 11 *pre-experimenten* (d.w.z. studies zonder controlegroep). Niet alle interventies waren succesvol. De auteur concludeerde dat met name in die studies effecten werden gevonden waarin meervoudige interventies werden ingezet, waarin leraren betrokken waren bij het ontwerp van de interventie en waarin leraren een actieve rol speelden bij de uitvoering van de interventie. Wade gebruikte (onderdelen van) systematische procedures voor een quality appraisal, maar gaf hierover geen andere informatie dan dat ze onderscheid maakte tussen de eerdergenoemde designs. Hoewel studies niet werden uitgesloten op grond van design, werden interventie-effecten wel per design besproken; er leek een relatie tussen design en de aanwezigheid van effecten. Wade maakte geen analyse van mogelijke publicatiebias.

## Combinatie van variabelen

De review van Paul en Clarke (2016) bestond uit acht (quasi-)experimenten met 2.431 middelbare scholieren (leeftijdswaarde: 11-18 jaar). Ze testten de effecten van twee typen leesinterventies op hun leesvaardigheid, waaronder hun leesbegrip:

1. Meervoudige interventies ( $k = 4$ ), die zich richtten op verschillende dimensies van de leesvaardigheid (bijv. decoderen, vloeiend lezen, leessnelheid, leesbegrip) en in sommige gevallen ook op andere taalvaardigheden (bijv. spelling, luisterbegrip). In drie gevallen ging het om *computer-aided instruction*, in het vierde geval om een interventie waarbij een logopedist leraren hielp om hun mondelinge en schriftelijke instructies toegankelijker te maken.
2. Interventies die zich uitsluitend richtten op leesbegrip ( $k = 4$ ). In deze categorie ging het in drie gevallen om strategie-interventies en in één geval om een interventie waarbij de nadruk lag op het ontwikkelen van vakkennis.

Leesbegrip werd gemeten met gestandaardiseerde en door onderzoekers ontwikkelde toetsen.

Geen van de interventies uit de eerste categorie bleek positieve effecten op leesbegrip te sorteren. Voor de interventies in de tweede categorie was dit wel het geval; over het algemeen was er sprake van kleine effecten. De auteurs gebruikten geen gestandaardiseerde procedure voor quality appraisal, maar registreerden wel enkele kwaliteitskenmerken, zoals design en de vraag of onderzoekers bij de afname van toetsen blind waren voor de conditie waaraan leerlingen waren toegewezen. Paul en Clarke hanteerden een strenge selectieprocedure: alleen RCT's en studies gepubliceerd in peer-reviewed tijdschriften werden geïncludeerd. De auteurs maakten echter geen analyse van mogelijke publicatiebias.

## Systematische reviews: leerlingen met speciale onderwijsbehoeften

### Kennis en vaardigheden

**Kennis van tekststructuur.** Hall-Mills et al. (2022) onderzochten het effect van tekststructuurinstructie (d.w.z. expliciete instructie over de structuur en organisatie van teksten in verschillende genres) op het leesbegrip van leerlingen met leerproblemen in het voortgezet onderwijs (Grade 7-12). Bij de onderzochte interventies ging het in de meeste gevallen om het aanleren van strategieën waarmee leerlingen de structuur van een tekst kunnen identificeren of waarmee ze structuurkennis kunnen inzetten om relevante informatie in een tekst te lokaliseren, dan wel om de inzet van hulpmiddelen zoals graphic organizers, waarmee ze centrale elementen uit een tekst kunnen ophalen. Bij de meting van leesbegrip werd onderscheid gemaakt tussen gestandaardiseerde tests en door onderzoekers ontwikkelde tests. Die laatste lagen dicht bij de inhoud van de interventies: denk aan een meting van de hoeveelheid informatie die leerlingen zich konden herinneren uit teksten die in een interventie werden gebruikt. In alle gevallen ging het om het begrip van informatieve teksten. De steekproeven bestonden uit leerlingen met gediagnosticeerde lees- of leerproblemen of leerlingen van wie de school aangaf dat ze het risico liepen op een leesachterstand. De review omvatte zeven studies. De aard van die studies was divers: twee meta-analyses (beide ook opgenomen in de huidige umbrella review: Hebert et al., 2016; Pyle et al., 2017), twee RCT's, vier quasi-experimenten en één single-case study. Alle studies lieten een positief effect zien en in alle gevallen op één na waren de effecten significant. De effectgrootten varieerden van klein tot groot ( $d = 0.31-2.17$ ); effecten werden niet gedifferentieerd naar interventietype. Hall-Mills et al. (2022) maakten gebruik van een gestandaardiseerde procedure voor quality appraisal: de *Quality of Evidence Summary Table* (QuEST) van de *American Speech-Language-Hearing Association* (Wolters et al., 2011). De aard van het design was een van de inclusiecriteria. Er was variatie in studiekwaliteit, bijvoorbeeld op het

gebied van de rapportage van implementatiekwaliteit. De auteurs maakten geen analyse van mogelijke publicatiebias.

**Technisch lezen (decoderen, vloeiend lezen).** Steinle et al. (2022) onderzochten het effect van fluency-interventies (interventies die tot doel hadden bij te dragen aan leessnelheid, accuratesse en expressie) op het vloeiend lezen en leesbegrip van zwakke lezers in het voortgezet onderwijs (Grade 6-12). De auteurs onderscheidde drie typen interventies:

1. Herhaald lezen (*repeated reading*) met een model: een volwassene of meer vaardige medeleerling leest de tekst eerst voor en de leerling herhaalt de tekst hardop, waarbij de leerling indien nodig gecorrigeerd wordt en/of feedback krijgt.
2. Herhaald lezen zonder model: de leerling leest een tekst meerdere keren hardop, waarbij de leerling indien nodig gecorrigeerd wordt en/of feedback krijgt van een leraar.
3. Fluency-interventies zonder repeated reading.

Leesbegrip werd gemeten met zowel gestandaardiseerde als door onderzoekers ontwikkelde toetsen en zwakke lezers werden gedefinieerd als leerlingen met “reading difficulties, reading disabilities, learning disabilities, dyslexia, or at-risk for reading or learning disabilities” (p. 5). De review omvatte in eerste instantie 17 studies met in totaal 337 leerlingen: vier (quasi-)experimentele studies en 13 studies met een single-case design. Van die laatste studies werden er uiteindelijk maar zeven meegenomen in de analyse, omdat alleen die voldeden aan minimale kwaliteitseisen (zie verder). De auteurs concludeerden dat de interventies wel bijdroegen aan vloeiend lezen, maar dat er geen duidelijk bewijs was voor een effect op leesbegrip: van de interventies in de quasi-experimentele studies resulteerde er maar één (een fluency-interventie zonder herhaald lezen) in positieve uitkomsten voor leesbegrip en hoewel ook in vier single-case studies gunstige resultaten voor leesbegrip werden gevonden, voldeed maar één van die studies aan de hoogste kwaliteitsstandaard (een studie naar herhaald lezen zonder model). De auteurs volgden een gestandaardiseerde procedure voor quality appraisal: de WWC-richtlijnen in combinatie met een codeerschema ontwikkeld in een eerdere studie. Voor de single-case studies speelde die quality appraisal een rol in de bepaling van interventie-effecten (zie hiervoor). Steinle et al. rapporteerden geen analyse van mogelijke publicatiebias.

Ook Wexler et al. (2008) gingen het effect na van fluency-interventies op het vloeiend en begrijpend lezen van zwakke lezers in het voortgezet onderwijs (Grade 6-12). De interventies bestonden met name uit herhaald hardop lezen met/zonder een model (zie Steinle

et al., 2022), waarbij het model een volwassene, een meer vaardige medeleerling of een audio-opname kon zijn. Leesbegrip werd gemeten met gestandaardiseerde en door onderzoekers ontwikkelde toetsen. Zwakke lezers waren leerlingen met dyslexie, een andere lees-, leer- of spraakstoornis, leerlingen met niet-gediagnosticeerde leesproblemen en leerlingen met een lage leesvaardigheid. De review omvatte 19 studies met in totaal 340 leerlingen. De geïncludeerde designs waren divers: single-subject studies ( $k = 11$ ), (quasi-)experimentele studies ( $k = 6$ ) en studies zonder controlegroep ( $k = 2$ ). Net als Steinle et al. stelden Wexler et al. (2008) vast dat fluency-interventies wel bijdroegen aan vloeiend lezen, maar geen effect hadden op leesbegrip. De auteurs gebruikten een gestandaardiseerde procedure om de kwaliteit van de studies vast te stellen: de WWC-richtlijnen. Ze includeerden alleen gepubliceerde studies met bepaalde designs (zie eerder), maar gaven geen verdere informatie over de variatie in studiekwaliteit. Ook maakten ze geen analyse van mogelijke publicatiebias.

**Zelfregulatie, metacognitie en leesstrategieën.** Daniel et al. (2021) onderzochten de effecten van interventies uitgevoerd in kleine groepjes en gericht op verschillende aspecten van de leesvaardigheid van zwakke lezers in het voortgezet onderwijs (Grade 6-12), waaronder hun leesbegrip. Zwakke lezers werden gedefinieerd als leerlingen met dyslexie of een andere leerstoornis, of leerlingen met een lage leesvaardigheid. De review bestond uit tien studies met 865 leerlingen (directe post-tests) en 693 leerlingen (uitgestelde post-tests). De auteurs verdeelden de interventies die in de studies werden getest onder in vier categorieën: metacognitieve-strategie-interventies, woordenschatinterventies, interventies gericht op het (vloeiend) lezen van meerlettergrepige woorden en meervoudige interventies. De studies waarin effecten op leesbegrip werden gemeten, behoorden alle vijf tot de eerste categorie. Het gemiddelde effect van deze interventies op leesbegrip (directe post-tests) was positief, significant en middelgroot ( $g = 0.67$ , 95% CI [0.10, 1.25]). Op de uitgestelde post-tests werd geen significant effect gevonden. Daniel et al. maakten gebruik van een gestandaardiseerde procedure om studiekwaliteit vast te stellen: die uit de *Guide for Education-Related Intervention Study Syntheses*, waarin onder meer gelet wordt op design, steekproeftrekking en -grootte, validiteit en betrouwbaarheid van effectmaten, en uitval. Hoewel twee van de studies niet aan de standaarden van die procedure voldeden, werden ze niet uitgesloten; wel werden alleen (quasi-)experimenten geïncludeerd. De auteurs maakten geen analyse van mogelijke publicatiebias.

Daniel en Williams (2021) gingen het effect na van instructie in één specifieke leesstrategie (self-questioning) op de leesvaardigheid (waaronder het leesbegrip) van zwakke

lezers in basis- en voortgezet onderwijs (kleuters-Grade 12). De instructie bestond eruit dat leerlingen uitleg kregen over hoe ze tijdens het lezen van een tekst vragen kunnen formuleren die hen helpen hun begrip van de tekst te monitoren. Vervolgens werden ze aangemoedigd om die strategie tijdens of na het lezen toe te passen. In de meeste gevallen werd leesbegrip gemeten met een door onderzoekers ontwikkelde test. De review omvatte tien studies (RCT's, quasi-experimenten en single-case studies) met 266 zwakke lezers, gedefinieerd als leerlingen met een lees- of leerstoornis, of een lage leesvaardigheid. Vijf van de studies hadden betrekking op middelbare scholieren (middle en high school); maar één van die studies liet een positief effect op leesbegrip zien. Daniel en Williams gebruikten een gestandaardiseerde procedure voor quality appraisal (die van de WWC; zie eerder), maar uitkomsten daarvan vormden geen grond voor uitsluiting. Alle RCT's en quasi-experimenten voldeden aan de standaarden van de WWC, de single-case studies deden dat niet. De auteurs maakten geen analyse van mogelijke publicatiebias, maar ze volgden wel een procedure om zulke bias te voorkomen, namelijk door contact op te nemen met de auteurs van primaire studies om na te gaan of zij onderzoek hadden gedaan naar hetzelfde thema dat (nog) niet gepubliceerd was.

Solis et al. (2012) testten het effect van verschillende leesstrategie-interventies op het leesbegrip van leerlingen met leerproblemen in het voortgezet onderwijs (Grade 6-8). Ze maakten onderscheid tussen interventies (1) waarin leerlingen leerden de hoofdgedachte van een tekst samen te vatten, (2) waarin ze leerden hun eigen begripsproces te monitoren of (3) waarin meerdere strategieën werden gecombineerd. Meestal werden voor de meting van leesbegrip door de onderzoekers ontwikkelde maten gebruikt, die door Solis et al. niet verder werden gespecificeerd. Alleen studies waarin ten minste de helft van de steekproef bestond uit leerlingen met gediagnosticeerde leerproblemen werden geïncludeerd. In totaal bestond de review uit 14 studies ( $N = 491$  leerlingen met leerproblemen): negen experimentele, drie quasi-experimentele en twee single-case studies. In die studies werden in totaal 18 interventies aangeboden, waarvan de meeste (13) betrekking hadden op hoofdgedachtestrategieën. Solis et al. rapporteerden over het algemeen middelgrote tot grote effecten op niet-gestandaardiseerde toetsen en kleinere effecten op gestandaardiseerde toetsen. De auteurs maakten gebruik van een gestandaardiseerde procedure voor quality appraisal (o.a. gebaseerd op de WWC-richtlijnen). Studies werden alleen uitgesloten op basis van publicatiestatus; er is dan ook variatie in studiekwaliteit (bijv. in de rapportage van implementatiekwaliteit). De auteurs gingen niet in op mogelijke publicatiebias.

## Kenmerken van de leeromgeving

**Graphic organizers.** Kim et al. (2004) testten het effect van graphic organizers (GO's) op het leesbegrip van leerlingen met leerproblemen in basis- en voortgezet onderwijs (kleuters-Grade 12).<sup>20</sup> GO's werden gedefinieerd als visuele of ruimtelijke weergaven ontworpen om de verwerking van teksten te ondersteunen. Ze maken gebruik van lijnen, pijlen en/of een ruimtelijke indeling, die tekstinhoud, tekststructuur en belangrijke conceptuele verbanden verbeelden. De auteurs onderscheidden interventies met:

1. *Semantic organizers*, waarbij leerlingen GO's krijgen te zien die verbanden tussen concepten en kenmerken van concepten (bijv. verwante woordenschat) weergeven.
2. *Cognitive maps* met een *mnemonic* of "geheugensteuntje", waarbij leerlingen met behulp van een vast stappenplan worden geholpen om op een gestructureerde manier een GO te maken. Kim et al. verwezen bijvoorbeeld naar de RELATE-aanpak (Boyle, 2000): lees een tekst snel door om verschillende onderwerpen te identificeren (Read), teken een cirkel per onderwerp (Etch), zoek in de tekst naar ideeën die bij één en bij meerdere onderwerpen passen (Look), verbind elk onderwerp met drie unieke ideeën (Ancor), verbind ideeën die bij meerdere onderwerpen staan (Tie) en markeer de ideeën die bij alle onderwerpen staan (Enclose).
3. Cognitive maps zonder mnemonic, waarbij leerlingen GO's zoals onder (2) krijgen aangereikt zonder dat ze een stappenplan doorwerken.
4. *Framed outlines*: grafisch weergegeven lesplannen of andere overzichten die leraren en leerlingen helpen bij het identificeren van hoofdgedachten en belangrijke feiten uit een tekst.

Leesbegrip werd meestal met door onderzoekers ontwikkelde toetsen gemeten, waarbij in verschillende gevallen werd gekeken naar de vaardigheid van leerlingen om concepten zoals verwerkt in de GO's toe te passen in concrete situaties. Leerlingen met leerproblemen werden gedefinieerd als leerlingen met een gemiddeld IQ, maar met "poor performance in at least one academic or related behavioral domain" (Kim et al., 2004, p. 106). De review bestond uit 21 studies met 848 leerlingen met leerproblemen: 15 van de studies hadden alleen betrekking op middelbare scholieren, één studie op leerlingen in basis- en voortgezet onderwijs. Kim et al. vonden positieve effecten voor vo-leerlingen: voor leerlingen in junior high school en high

---

<sup>20</sup> Graphic organizers kwamen ook in verschillende andere literatuurstudies terug, maar alleen Kim et al. (2004) keken uitsluitend naar het effect van GO's.

school waren de gemiddelde effectgrootten respectievelijk  $d = 1.18$  en  $1.08$ . Over het geheel bleken alle typen GO's effectief, maar er werd geen uitsplitsing gemaakt naar onderwijsfasen; of die conclusie ook gold voor vo-leerlingen is dus niet duidelijk. De auteurs volgden geen standaardprocedure om studiekwaliteit te bepalen, maar codeerden alleen het design: in 15 studies werden een experimentele en controleconditie vergeleken, in 5 studies kreeg één groep meerdere interventies. Design was ook een exclusie criterium: single-case studies werden uitgesloten. Kim et al. maakten geen analyse van mogelijke publicatiebias.

**Peer mediation/coöperatief leren.** Alzahrani en Leko (2018) testten de effecten van *peer tutoring* op het leesbegrip van middelbare scholieren met een beperking. Peer tutoring werd gedefinieerd als een set van praktijken en strategieën waarbij leerlingen tijdens de uitvoering van leestaken worden ingezet om medeleerlingen van geïndividualiseerde instructie, oefening, herhaling en verduidelijking te voorzien. Hoe leesbegrip werd gemeten, werd niet gespecificeerd. Bij beperkingen ging het om emotionele of gedragsproblematiek, leerproblemen en/of cognitieve beperkingen. De review bestond uit tien (quasi-)experimentele en single-case studies. Van de in totaal 443 leerlingen waren er 86 afkomstig van middle schools en 357 van high schools. In alle studies ging de inzet van peer tutoring samen met verbeterd leesbegrip. De auteurs gebruikten een standaardprocedure voor quality appraisal: de CEC-criteria (zie eerder). Studiekwaliteit was geen exclusie criterium: studies werden alleen uitgesloten als ze niet waren gepubliceerd in een peer-reviewed tijdschrift. Desalniettemin voldeden alle studies voor ten minste 80 procent aan de CEC-criteria. Publicatiebias werd niet onderzocht.

Ook Okilwa en Shelby (2010) keken naar de effectiviteit van peer tutoring op (onder meer) het leesbegrip van leerlingen met een beperking in het voortgezet onderwijs (Grade 6-12). De auteurs onderscheidden vier typen peer tutoring:

1. *Classwide peer tutoring*: alle leerlingen in een klas werken samen in paren aan taken die door de leraar zijn gestructureerd en waarbij ze gebruikmaken van een standaard procedure voor de correctie van fouten. De paren worden vervolgens toegewezen aan groepen die met elkaar in competitie gaan.
2. *Reciprocal relationship*. Deze aanpak lijkt op het eerste type met het verschil dat tutor en *tutee* op bepaalde momenten van rol wisselen.
3. *Cross-age matching*. Bij deze benadering bestaan paren uit leerlingen van verschillende leeftijden/uit verschillende leerjaren.
4. Overige vormen, waarbij het bijvoorbeeld juist de leerlingen met een beperking zijn die de tutor-rol aannemen.

Leesbegrip werd uitsluitend gemeten met door onderzoekers ontwikkelde tests. De auteurs onderscheidden verschillende typen beperkingen: leerstoornissen, emotionele en gedragsstoornissen en mentale beperkingen, leerlingen die remediërend onderwijs kregen en leerlingen met “overige” beperkingen. De auteurs analyseerden 12 studies ( $N = 571$ ): (quasi-)experimentele studies, single-case studies en multiple-baseline studies waarin interventies altemnerend werden aangeboden. Drie van de studies (alle drie gericht op rolwisselend lezen) maten leesbegrip; in alle gevallen was er verbeterd leesbegrip na inzet van peer tutoring. Okilwa en Shelby voerden geen gestandaardiseerde quality appraisal uit, maar rapporteerden wel design en implementatiekwaliteit. Studies werden niet uitgesloten op basis van kwaliteit, alleen op basis van publicatiestatus. Hoewel designs varieerden (zie hiervoor), werd in alle studies gerapporteerd over implementatiekwaliteit, die vaak hoog was. De auteurs maakten geen analyse van mogelijke publicatiebias.

Wexler et al. (2015) onderzochten het effect van peer mediation op de schoolprestaties van middelbare scholieren (Grade 6-12). Ze keken onder meer naar resultaten op leesbegrip, gemeten met gestandaardiseerde en door onderzoekers ontwikkelde toetsen. Peer mediation had betrekking op situaties waarin leerlingen van dezelfde leeftijd elkaar tutoeren of samenwerken in duo's/kleine groepjes, met als doel een bepaalde taak af te ronden. Een vereiste voor inclusie was dat leerlingen instructie kregen over tutoring of samenwerking. De review richtte zich op leerlingen “identified as at risk, low achievement, struggling, below grade level, or diagnosed with a learning disability (LD)” (p. 454) op het gebied van lezen, geletterdheid en/of rekenen. De review bestond uit 13 studies ( $N = 653$ ), waarin ofwel een vergelijking tussen een experimentele en controlegroep werd gemaakt ( $k = 10$ ) ofwel een single-case design werd gebruikt ( $k = 3$ ). Wexler et al. concludeerden dat peer mediation bijdroeg aan leesbegrip, vooral wanneer er feedback aan was gekoppeld. De auteurs volgden een gestandaardiseerde procedure voor het vaststellen van studiekwaliteit: de WWC-richtlijnen. Alleen design en publicatiestatus werden gebruikt om studies uit te sluiten. Er was variatie in studiekwaliteit, maar die verschilde wel per indicator: zo werd in alle *treatment-control*-studies random toewijzing aan condities toegepast, maar werd in sommige studies niet of nauwelijks aandacht besteed aan implementatiekwaliteit. De auteurs maakten geen analyse van mogelijke publicatiebias.

## Combinatie van variabelen

Cisco en Padrón (2012) onderzochten de rol van woordenschat en leesstrategieën bij de ontwikkeling van het leesbegrip van meertalige leerlingen in het voortgezet onderwijs (Grade

5-8) alsmede de effecten van woordenschat- en leesstrategie-instructie op die ontwikkeling. De review bestond uit 11 studies (totale aantal leerlingen onbekend), die sterk verschilden van aard: de dataset bevatte quasi-experimentele studies, case studies, interviewstudies en hardop-denkstudies. In alle gevallen ging het overigens om leerlingen in de VS die het Engels als tweede taal leerden. De auteurs vatten de uitkomsten samen met een thematische analyse. Ze trokken drie conclusies:

1. Woordenschat in de tweede taal (T2) speelt een cruciale rol bij het leesbegrip van deze leerlingen. Cisco en Padrón constateerden dat zwakke tweetalige lezers doorgaans een beperkte T2-woordenschat hebben en dat sterke tweetalige lezers meer strategieën inzetten om onbekende woorden te achterhalen dan zwakke.
2. Vaardigheid in de eerste taal (T1) draagt bij aan leesbegrip in de T2: tijdens het lezen van een tekst in de tweede taal is er sprake van een transfer van inhoudelijke kennis, processen en strategieën van T1 naar T2. Sterke tweetalige lezers zetten tijdens het lezen van een tekst in de T2 bovendien strategieën in waarbij ze gebruikmaken van hun T1.
3. Woordenschat- en strategie-instructie en instructie waarbij de T1 wordt ingezet als hulpmiddel om T2-teksten te begrijpen, kunnen bijdragen aan het leesbegrip in een tweede taal. Wel maakte het beperkte aantal studies en de verschillen in focus het moeilijk om te bepalen wat precies de effectieve componenten van zulke instructie zijn.

De auteurs gebruikten een gestandaardiseerde procedure om de kwaliteit van de studies te bepalen: de richtlijnen van de *National Research Council's Scientific Research in Education*. Bij de selectie van studies letten ze met name op de vraag of er sprake was van empirisch onderzoek, hoe sterk de theoretische/conceptuele onderbouwing van de studie was, of het design aansloot bij de onderzoeksvragen, hoe helder en precies de onderzoeksmethode was beschreven en hoe plausibel de conclusies waren. Vier studies werden op basis van studiekwaliteit geëxcludeerd, maar variatie in kwaliteit van de overgebleven studies werd niet expliciet gerapporteerd. Cisco en Padrón maakten geen analyse van mogelijke publicatiebias.

Garwood et al. (2014) onderzochten de effecten van verschillende typen leesinterventies op het leesbegrip van adolescenten (11-16 jaar oud) met gedrags- en emotionele stoornissen. De interventies richtten zich zowel op vloeiend als op begrijpend lezen en vielen uiteen in zeven typen, waarbij in sommige programma's meerdere interventies werden gecombineerd: (1) herhaald (hardop) lezen; (2) *cognitive/text mapping*:

het maken van een visuele representatie van een tekst, bijvoorbeeld door middel van een graphic organizer; (3) een combinatie van keuze (leerlingen mochten zelf voor een instructiewijze kiezen) en beloningen; (4) *listening-while-reading*, waarbij leerlingen een tekst lezen terwijl die ook wordt voorgelezen (door een persoon of computer), (5) *corrective reading*: directe instructie gericht op accuraat, vlot en begrijpend lezen; (6) *peer-assisted learning* (zie eerder); en (7) *self-graphing*, waarbij leerlingen zelf bijhouden hoe accuraat en vlot ze lezen. Leesbegrip werd gemeten met gestandaardiseerde en door onderzoekers ontwikkelde toetsen en met verschillende toetswijzen (meerkeuzevragen over teksten, cloze-taken, mondelinge [open] vragen en essays). Leerlingen met stoornissen werden geïdentificeerd op basis van de richtlijnen van hun school of met behulp van landelijke of wettelijke richtlijnen. De review bestond uit negen single-subject studies met in totaal 38 leerlingen. Uitkomsten varieerden: van de vijf studies met leesbegrip als uitkomst resulteerden er twee in zichtbare vooruitgang bij leerlingen, drie niet. In de eerste twee studies werd gebruikgemaakt van story mapping. De auteurs maakten geen gebruik van een standaardprocedure om studiekwaliteit te bepalen, maar ze codeerden wel enkele kwaliteitsaspecten (interbeoordelaarsbetrouwbaarheid bij de scoring van uitkomstmaten, implementatiekwaliteit en *social validity*, d.w.z. hoe leerlingen de interventies waardeerden). Er werden geen studies uitgesloten op basis van kwaliteit, hoewel die kwaliteit wel varieerde. Garwood et al. maakten geen analyse van mogelijke publicatiebias.

Huddle et al. (2017) onderzochten de effecten van leesinterventies voor tweedetaalleerders met leesproblemen of -stoornissen (niet nader gespecificeerd) in het basis- en voortgezet onderwijs (Grade 4-12). In de meeste gevallen ging het om meervoudige interventies, dat wil zeggen, om interventies waarin aandacht werd besteed aan verschillende aspecten van leesvaardigheid (bijv. fonologisch bewustzijn, woordenschat, fluency, leesbegrip). Leesbegrip werd gemeten met “any measure that required students to read the text and answer questions or complete sentences within the text” (p. 7). De review bestond uit acht (quasi-)experimentele studies, waarvan er zeven zijn uitgevoerd in het voortgezet onderwijs (middle school:  $k = 6$ , high school:  $k = 1$ ); maar één studie had betrekking op basisschoolleerlingen. De effecten op leesbegrip varieerden van negatief en klein tot positief en middelgroot ( $g = -0.20-0.71$ ). Huddle et al. gebruikten een gestandaardiseerde procedure voor quality appraisal (de WWC-richtlijnen). Ze sloten alleen studies uit op basis van design: uitsluitend (quasi-)experimenten werden geïnccludeerd (zie hiervoor). Vier van de acht studies voldeden aan alle WWC-criteria, drie studies deden dat deels en één studie voldeed niet. De auteurs maakten geen analyse van mogelijke publicatiebias.

Kim et al. (2012) onderzochten het effect van twee verschillende instructiemethoden op het leesbegrip van leerlingen met leerstoornissen in het voortgezet onderwijs (middle school):

1. *Instructional modifications*: interventies waarbij de inhoud van leermaterialen werd bewerkt om die toegankelijker te maken. Het ging hierbij met name om de inzet van graphic organizers.
2. Strategie-instructie: interventies waarbij leerlingen hulpmiddelen kregen aangereikt om teksten zelfstandig te verwerken. Hieronder vielen strategieën als de hoofdgedachte van een tekst achterhalen, jezelf vragen stellen tijdens het lezen en samenvatten.

Leesbegrip werd gemeten met gestandaardiseerde en door onderzoekers ontwikkelde toetsen en bij de identificatie van leerstoornissen (niet nader toegelicht) werd uitgegaan van “federal and state criteria” (p. 68). De review bestond uit 14 (quasi-)experimentele studies afkomstig uit 11 publicaties. Voor instructional modifications ( $k = 6$ ) waren de effecten niet eenduidig, voor strategie-instructie ( $k = 8$ ) waren de effecten over het algemeen positief; met name studies waarin de effecten van samenvatten en het achterhalen van de hoofdgedachte werden getest ( $k = 4$ ), lieten grote effecten zien, zowel over alle effectmaten heen ( $g = 1.41$ ) als op gestandaardiseerde tests ( $g = 0.84$ ). In drie studies die het effect van tekststructuurinterventies testten, werden ook positieve effecten gevonden. Kim et al. voerden geen gestandaardiseerde quality appraisal uit, maar codeerden wel bepaalde aspecten van studiekwaliteit (toewijzing aan condities, implementatiekwaliteit) en sloten op basis daarvan studies uit: ze namen alleen (quasi-)experimentele studies met een implementatiecheck op. De auteurs maakten geen analyse van mogelijke publicatiebias.

Payne en Yoon (2024) onderzochten de effecten van verschillende interventies op de schoolprestaties (waaronder het leesbegrip) van leerlingen in het voortgezet onderwijs (high school). Studies werden geselecteerd als ten minste de helft van de steekproef bestond uit leerlingen met een combinatie van leer- en gedragsproblemen. De auteurs analyseerden negen single-case studies met in totaal 27 leerlingen. Van de drie studies met leesbegrip als uitkomstmaat (niet verder gespecificeerd) hadden er twee betrekking op interventies met graphic organizers en één op een strategie-interventie. In alle drie de gevallen was er sprake van grote effecten. Payne en Yoon maakten gebruik van een gestandaardiseerde procedure voor het bepalen van de kwaliteit van single-case studies (Horner et al., 2005), waarbij ze keken naar design, toewijzing, uitval, implementatiekwaliteit en sociale validiteit (zie eerder).

Design was een inclusie criterium, maar, hoewel de auteurs ook zochten naar quasi-experimenten, vonden ze die niet. Er was variatie in studiekwaliteit, met name op het gebied van de toewijzing van leerlingen aan interventies. De auteurs maakten geen analyse van mogelijke publicatiebias.

## Gecombineerde meta-analyses en systematische reviews: leerlingen met speciale onderwijsbehoeften

### Kennis en vaardigheden

**Zelfregulatie, metacognitie en leesstrategieën.** Stevens et al. (2019) onderzochten het effect van strategie-interventies op het leesbegrip van zwakke lezers uit basis- en voortgezet onderwijs (Grade 3-12). Die interventies waren uitsluitend gericht op het leren samenvatten of achterhalen van de hoofdgedachte van teksten. De auteurs maakten een onderscheid tussen verschillende typen interventies:

1. Tekststructuurinstructie, waarbij leerlingen verschillende typen (informatieve) tekststructuren leren herkennen, signaalwoorden leren identificeren of alert worden gemaakt op bepaalde tekstelementen (bijv. afsluitende zinnen).
2. Leren parafraseren via een stapsgewijze procedure (bijv. lees een alinea, bepaal de hoofdgedachte, kies twee relevante details en beschrijf hoofdgedachte en details in je eigen woorden).
3. Het gebruik van graphic organizers.
4. Het aanleren van *microprocessing*-strategieën. Hierbij gaat het om interventies gericht op het actief verwerken van de microstructuur van een tekst (bijv. op hoe verbindingswoorden zoals “want” of “omdat” relaties tussen tekstdelen uitdrukken).
5. Interventies waarin het leren samenvatten of achterhalen van hoofdgedachten wordt gecombineerd met self-monitoring: het volgen en aansturen van het eigen leesproces door doelen te stellen, jezelf vragen te stellen tijdens het lezen en genomen stappen te evalueren.

Leesbegrip werd gemeten met gestandaardiseerde en door onderzoekers ontwikkelde toetsen. Zwakke lezers werden gedefinieerd als leerlingen met een gediagnosticeerde lees- of leerstoornis of leerlingen van wie (het risico op) leesachterstanden werd bepaald door lage leesscores, plaatsing in een remediëringstraject of richtlijnen van de school of het onderwijsdistrict. Stevens et al. analyseerden 30 studies met in totaal 983 leerlingen

(basisschool:  $k = 10$ ; middle school:  $k = 9$ ; high school:  $k = 8$ ; middle en high school:  $k = 3$ ). Van de 30 studies volgden er 24 een (quasi-)experimentele opzet, in zes studies werd een single-case design gebruikt. Met de eerstgenoemde studies werd een meta-analyse uitgevoerd: die liet over alle interventies heen een significant, positief en groot effect op leesbegrip zien ( $g = 0.97$ ,  $p < .001$ ). Onderwijsfase bleek geen moderator: de gunstige effecten golden dus ook voor middelbare scholieren. Ook de single-case studies lieten positieve uitkomsten zien van samenvatten/hoofdgedachte achterhalen, met name op het (mondeling/schriftelijk) navertellen van teksten. Alleen voor de single-case studies maakten de auteurs gebruik van een standaardprocedure voor quality appraisal: de WWC-richtlijnen. Vijf van die studies voldeden (met reserves) aan de richtlijnen, één studie niet. Voor de meta-analyse was design een inclusie criterium: in verreweg de meeste gevallen ( $k = 20$ ) gebruikten studies een experimentele opzet. De auteurs keken niet naar moderatoreffecten van studiekwaliteit. Wel analyseerden ze publicatiebias: ze vonden geen aanwijzingen dat die de resultaten beïnvloedde.

## Kenmerken van de leeromgeving

**Vakintegratie.** Swanson et al. (2014) onderzochten het effect van interventies gericht op leesbegrip binnen de mens- en maatschappijvakken (social studies) bij leerlingen met (taal)leerproblemen en leerlingen met milde cognitieve beperkingen in basis- en voortgezet onderwijs (Grade 3-12). Ze keken naar effecten op leesbegrip (gemeten met gestandaardiseerde en door onderzoekers ontwikkelde toetsen) en de verwerving van vakkennis. De review bestond uit 27 studies: 12 (quasi-)experimentele studies waarin één experimentele groep en één controlegroep werden vergeleken, drie studies waarin meerdere experimentele condities werden vergeleken (al dan niet met een controlegroep), negen studies waarin alleen een experimentele groep werd onderzocht en drie single-case studies. Op 16 studies werd een meta-analyse uitgevoerd, op 11 studies een systematische review. Selectie voor de meta-analyse was afhankelijk van beschikbare kwantitatieve informatie, niet noodzakelijk van design: de meta-analyse bevatte bijvoorbeeld ook een single-case studie. In de studies werden verschillende typen interventies toegepast: graphic organizers ( $k = 7$ ), mnemonics (bijv. ondersteuning van begrip met plaatjes;  $k = 6$ ), begripsvragen over teksten ( $k = 4$ ), *guided notes* (door de leraar voorbereide aantekeningen;  $k = 3$ ), meervoudige interventies ( $k = 3$ ) en “overige” interventies ( $k = 7$ ). In de 25 studies die hierover rapporteerden, zijn er 13 uitgevoerd bij leerlingen uit Grade 3 tot en met 8 en acht bij leerlingen uit Grade 9-12; vier studies hadden betrekking op leerlingen uit middle en high

school. Over alle interventies en leeftijdsgroepen heen liet de meta-analyse een significant, positief en groot effect op leesbegrip zien ( $d = 1.02, p < .001$ ). Onderwijsfase was geen significante moderator: het effect gold dus ook voor leerlingen in het voortgezet onderwijs. Ook in de systematische review werden positieve uitkomsten gerapporteerd. Voor voerleerlingen werd overigens niet onderzocht of er verschil in effecten was tussen de eerdergenoemde interventietypen. De auteurs voerden een kwaliteitscheck uit en maakten daarvoor gebruik van de WWC-richtlijnen. Studiekwaliteit was geen exclusie criterium. Er was dan ook aanzienlijke variatie in kwaliteit: niet alleen in design (zie eerder), maar bijvoorbeeld ook in de vraag of er werd gerapporteerd over implementatiekwaliteit. In de meta-analyse werden geen aspecten van studiekwaliteit opgenomen als moderator. Noch in de meta-analyse noch in de systematische review werd gekeken naar mogelijke publicatiebias.

### Overzichtstabel

In Tabel 5 geven we een overzicht van onze conclusies over de bewijskracht en de effecten per variabele. We hebben de bevindingen voor zich normaal ontwikkelende leerlingen en leerlingen met speciale onderwijsbehoeften gecombineerd in dezelfde tabel. Als er uitsluitend studies waren naar zich normaal ontwikkelende leerlingen staat er in elke cel maar één cijfer. Als er studies waren naar beide doelgroepen staat het cijfer voor de zich normaal ontwikkelende leerlingen voor de / en staat het cijfer voor de leerlingen met speciale behoeften achter de /. Als er uitsluitend studies waren naar leerlingen met speciale behoeften (alleen het geval bij graphic organizers) staat er ook een / voor het cijfer. De conclusies over de bewijskracht en effecten per individuele studie zijn te vinden in Bijlage 3.

**Tabel 5**

*Bewijskracht en effecten per (sub)categorie (indien relevant: zich normaal ontwikkelende leerlingen staan voor de / en leerlingen met speciale onderwijsbehoeften staan achter de /)*

(Sub)categorie	<i>k</i>	Bewijskracht			Effecten		
		+	~	-	+	~	0
<i>Kennis en vaardigheden</i>							
Inferenties maken/achtergrondkennis	4	1	1	2	4		
Integreren van meerdere teksten	1			1	1		
Kennis van tekststructuur	2/1		2/1		2/1		
(Mondelinge) taalvaardigheid	4/2		4/2		0/2	1/0	3/0
Schrijfvaardigheid*	3		3		2		1
Technisch lezen (decoderen, vloeiend lezen)	1/3		1/2	0/1			1/3
Zelfregulatie, metacognitie en leesstrategieën	8/13	0/2	6/10	2/1	6/12	1/1	1/0
<i>Leesgedrag en -houding</i>							
	6		3	3	3		3
<i>Leeromgeving</i>							
Digitale technologie	4/1		3/0	1/1	0/1	3/0	1/0
Graphic organizers	/3		/3		/2	/1	
Peer mediation/coöperatief leren	1/3		1/3		0/3		1/0
Personalisatie	1		1				1
Tutoring	1		1		1		
Vakintegratie	3/2		3/1	0/1	1/2	1/0	1/0
<i>Schoolorganisatie</i>							
	1		1				1
<i>Combinatie van variabelen</i>							
	1/3		1/2	0/1	0/1	0/2	1/0

\* Hebert et al. (2013) vergeleken verschillende soorten schrijfinstructie, maar maakten geen vergelijking met een controlegroep. Omdat ze dus geen inschatting van de effecten van schrijfinstructie op zich, is deze studie in de tabel buiten beschouwing gelaten.

## Conceptuele overzichtsstudies

De vier gevonden conceptuele overzichtsstudies hadden betrekking op twee begrippen: leesmotivatie (Conradi et al., 2014; Schiefele et al., 2012) en reading engagement of leesbetrokkenheid (Lee et al., 2021; Unrau & Quirk, 2014).<sup>21</sup> Conradi et al. (2014) maakten een inventarisatie van termen die in de wetenschappelijke literatuur worden gebruikt om te verwijzen naar leesmotivatie. De auteurs analyseerden daarvoor 92 empirische studies: de meeste waren kwantitatief en met name correlatief (47%) van aard, hoewel er ook enkele kwalitatieve studies werden geïncludeerd. De overzichtsstudie was niet gericht op een bepaalde doelgroep; wel had het merendeel van de onderzochte studies (65%) betrekking op leerlingen uit Grade 4-12. Conradi et al. (2014) identificeerden negen relevante termen: *agency, attitude, expectancy, goal, interest, reading motivation, self-concept, self-efficacy* en *value*. Voor elke term formuleerden ze een definitie, gebaseerd op informatie waarover consensus bestond, dat wil zeggen: informatie die vrij systematisch terugkwam in omschrijvingen van de termen in de literatuur. De (vertalingen van de) definities zijn te vinden in Tabel 6. De termen brachten ze vervolgens onder in een hiërarchie (zie Figuur 2), die ze baseerden op de drie centrale concepten uit hun definitie van leesmotivatie: doelen, overtuigingen en (pre)disposities/attitudes; deze definitie vormde overigens de basis voor de eerdergenoemde meta-analyses van Toste et al. (2020) en Van der Sande et al. (2023). De termen komen voort uit verschillende theorieën. Deze werden door Conradi et al. (2014) wel benoemd (bijv. *self-determination theory, reading engagement model, expectancy-value theory*), maar niet nader beschreven. Bovendien legden de auteurs geen verbinding tussen leesmotivatie (of een verwant concept) en leesbegrip.

---

<sup>21</sup> We selecteerden aanvankelijk vijf studies, maar de studie van Jang et al. (2015) bleek een bewerking van de studie van Conradi et al. (2014) voor *The Reading Teacher*, een tijdschrift dat op de onderwijspraktijk is georiënteerd. Daarom hebben we de studie van Jang et al. (2015) alsnog uitgesloten. In die studie werden overigens wel suggesties voor instrumenten en vragenlijstitems gedaan.

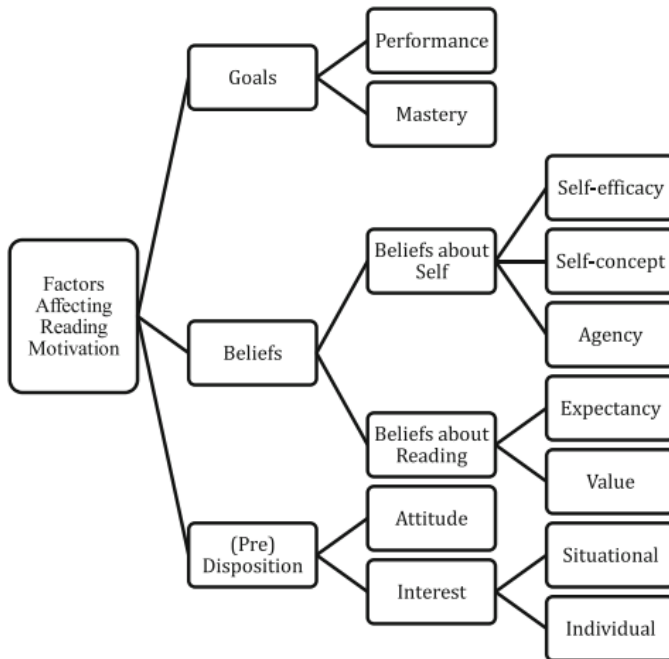
**Tabel 6***Vertaalde definities uit Conradi et al. (2014)*

<b>Concept</b>	<b>Definitie</b>
Agency	Iemands (gepercipieerde) vermogen om de eigen betrokkenheid bij het leesproces en leesactiviteiten te sturen.
Attitude	Een set van verworven gevoelens over lezen die ervoor zorgt dat iemand gaat lezen of lezen juist vermijdt.
Expectancy	De mate waarin iemand verwacht dat een leeservaring resulteert in succes of mislukking.
Goal	Iemands oriëntatie op en intenties ten aanzien van lezen (d.w.z. wat iemand beoogt met zijn of haar leesactiviteiten), waarbij verder onderscheid kan worden gemaakt tussen prestatie- en beheersingsdoelen (zie Toste et al., 2020).
Interest	Een positieve oriëntatie op lezen over een bepaald onderwerp. Interesse kan individueel of situationeel zijn. Bij individuele interesse gaat het om een relatief stabiele en blijvende oriëntatie op lezen over een bepaald onderwerp. Bij situationele interesse gaat het om een context-specifieke oriëntatie op lezen over een bepaald onderwerp, die zich doorgaans beperkt tot bepaalde momenten.
Reading motivation	De neiging om te lezen die voortkomt uit een uitgebreide set van overtuigingen, attitudes en doelen met betrekking tot lezen. Intrinsieke motivatie is de neiging om te lezen voor interne doeleinden, zoals plezier hebben, persoonlijke doelen behalen of het bevredigen van nieuwsgierigheid. Extrinsieke motivatie is de neiging om te lezen voor externe doeleinden, zoals beloningen of erkenning.
Self-concept	De algemene perceptie die iemand van zichzelf heeft als lezer, inclusief het gevoel van competentie en de rol die lezen speelt in zijn of haar identiteit.
Self-efficacy	De inschatting van het eigen vermogen om een specifieke leestaak tot een goed einde te brengen.

Value

Iemands opvattingen over de mate waarin lezen over het algemeen nuttig, vermakelijk of op een andere manier belangrijk is.

Figuur 2: Hiërarchie van motivatie-gerelateerde constructen (Conradi et al., 2014)



Schiefele et al. (2012) probeerden op basis van een synthese van 34 kwantitatieve en kwalitatieve studies tot een overzicht te komen van de dimensies van leesmotivatie en hoe die dimensies zich tot elkaar verhouden. Ook onderzochten ze de relatie tussen dimensies van leesmotivatie en leesvaardigheid (waaronder leesbegrip). De auteurs maakten allereerst onderscheid tussen *actuele* en *habituële* leesmotivatie. Bij het eerste gaat het om de intentie die iemand heeft om een bepaalde tekst in een bepaalde situatie te lezen, bij het tweede om de relatief stabiele bereidheid die iemand heeft om te lezen. Daarnaast maakten Schiefele et al. onderscheid tussen wat de kernelementen van leesmotivatie kunnen worden genoemd<sup>22</sup> en “antecedenten”, dat wil zeggen: variabelen die leesmotivatie voorspellen, maar geen deel uitmaken van het construct. Bij de kernelementen gaat het om het onderscheid tussen intrinsieke en extrinsieke motivatie. Intrinsieke motivatie kan verder worden onderverdeeld in object-specifieke motivatie (een lezer is dan gemotiveerd om te lezen omdat hij of zij geïnteresseerd is in het onderwerp van de tekst) en activiteit-specifieke motivatie, wat

<sup>22</sup> Schiefele et al. spreken van “genuine reading motivation constructs” (p. 431).

inhoudt dat een lezer gemotiveerd is om te lezen omdat de leesactiviteit zelf samengaat met positieve ervaringen (zoals opgaan in een verhaal). Bij extrinsieke motivatie wordt lezen gestuurd door externe redenen, die te maken hebben met de uitkomsten van de leesactiviteit, niet met de leesactiviteit zelf. Schiefele et al. stelden vast dat extrinsieke motivatie doorgaans niet gerelateerd is aan leesvaardigheid en dat intrinsieke motivatie een positieve relatie heeft met leesvaardigheid. Onder de antecedenten van leesmotivatie vallen allereerst de waarden die worden gehecht aan lezen: leest iemand bijvoorbeeld omdat hij of zij lezen nuttig vindt (*utility value*) of omdat hij of zij het belangrijk vindt om beheersing te bereiken (*attainment value*)? Andere antecedenten zijn self-concept en self-efficacy (zie ook Conradi et al., 2014).

Lee et al. (2021) inventariseerden hoe onderzoekers leesbetrokkenheid definiëren, welke dimensies van dat concept ze doorgaans onderzoeken en hoe ze die dimensies meten. De auteurs analyseerden in totaal 60 empirische studies. Ongeveer de helft ( $k = 33$ ) bestond uit kwantitatieve studies waarin leesbetrokkenheid in kaart werd gebracht met surveys (meestal zelfrapportages); bij de andere 27 studies ging het om kwalitatief onderzoek waarin betrokkenheid werd gemeten met behulp van interviews of observaties ( $k = 27$ ). Van de 60 studies hadden er 45 betrekking op leerlingen uit basis- en voortgezet onderwijs (kleuterbouw-Grade 12); studies gericht op leerlingen uit Grade 6 tot en met 8 vormden daarbinnen de grootste groep ( $k = 19$ ). Bij de definitie van de dimensies van leesbetrokkenheid maakten de auteurs twee onderscheidingen. Allereerst zagen ze dat onderzoekers betrokkenheid ofwel definieerden als een *aptitude* (d.w.z. als een relatief stabiel persoonskenmerk, dat overigens wel beïnvloedbaar is) ofwel als een *event* (d.w.z. als een handeling of activiteit). Daarnaast maakten ze onderscheid tussen gedragsmatige, cognitieve, affectieve en sociale betrokkenheid. Bij gedragsmatige betrokkenheid gaat het om gedragingen direct gerelateerd aan het leerproces of om de deelname aan activiteiten binnen en buiten de klas. Gedragsmatige betrokkenheid bij lezen wordt vaak gemeten door te kijken naar de tijd die leerlingen besteden aan lezen, via rapportages van hun inzet en volharding bij de uitvoering van leestaken of door middel van observaties van het leesgedrag. Cognitieve betrokkenheid bij lezen definieerden de auteurs op basis van Guthrie et al. (2012) als de bereidheid om de mentale inspanning te leveren die nodig is om teksten te begrijpen en uitdagende leestaken te volbrengen. Metingen hebben betrekking op de gerapporteerde voorkeur voor bepaalde doeloriëntaties (prestatiedoelen, beheersingsdoelen) of het gebruik van (metacognitieve) strategieën. Affectieve betrokkenheid bij lezen verwijst naar de emoties die ervoor zorgen dat leerlingen betrokken raken bij een leesactiviteit, -taak of -ervaring; deze dimensie vertoont volgens de auteurs sterke overlap met leesmotivatie. Sociale betrokkenheid

bij lezen bestaat uit een aantal elementen: het gaat om de uitwisselingen die leerlingen met elkaar en met de leraar hebben over teksten, om de mate waarin leerlingen (expliciete of impliciete) gedragsregels volgen tijdens leeslessen, maar ook om de rol die leerlingen zelf spelen in de verhouding met anderen en met het boek. Lee et al. stellen “Readers are agentic beings who have reciprocal and transactional relationships [...] not only with the characters inside books, but also with themselves and with others...” (p. 546).

De overzichtsstudie van Unrau en Quirk (2014) had betrekking op de verhouding tussen leesmotivatie en leesbetrokkenheid. Anders dan in de vorige drie conceptuele studies, gingen de auteurs niet in op hun onderzoeksmethode, waardoor bijvoorbeeld een overzicht van het aantal onderliggende studies, de aard van die studies en de participanten ontbreekt. Ze definieerden leesmotivatie als de gedachten, overtuigingen en zelfpercepties die iemand ertoe aanzetten te lezen of leesactiviteiten vol te houden. Leesbetrokkenheid omschreven ze als de observeerbare en niet-observeerbare handelingen die gepaard gaan met iemands leesactiviteiten. Net zoals Lee et al. onderscheidden Unrau en Quirk gedragsmatige, cognitieve en affectieve betrokkenheid; sociale betrokkenheid maakte geen onderdeel uit van hun definitie. Gedragsmatige betrokkenheid definieerden ze als alle aan lezen gerelateerd gedrag, zoals het lezen zelf, maar ook het kiezen van een boek. Affectieve betrokkenheid omschreven ze als iemands emotionele reacties tijdens het lezen van een tekst, zoals interesse, verveling, blijdschap, verdriet of stress. Onder cognitieve betrokkenheid verstonden ze de inzet van strategieën of metacognitieve handelingen tijdens het lezen; het gaat daarbij vooral om acties die betrekking hebben op kennisconstructie, zoals het formuleren van hypothesen over de inhoud van de tekst en het evalueren van die hypothesen, of het herlezen van stukken tekst bij begripsproblemen.

## Discussie

In deze literatuurstudie stonden twee vragen centraal:

1. Welke kenmerken van leesonderwijs (d.w.z. domeinspecifieke kenmerken en algemene onderwijskundige principes en professionele kwaliteiten van leraren) en welke aspecten van het leesgedrag en de leeshouding van leerlingen en leraren dragen bij aan het leesbegrip van adolescenten/leerlingen in het voortgezet onderwijs?
2. Hoe worden centrale concepten op het gebied van leesgedrag en -houding voor deze doelgroep gedefinieerd en geoperationaliseerd?

Het doel was om met de beantwoording van deze vragen te bepalen welke factoren relevant zijn om mee te nemen in de landelijke peiling naar leesvaardigheid aan het einde van leerjaar 2 van het voortgezet onderwijs. Meer specifiek moesten (a) aanbevelingen worden gedaan voor het opnemen van factoren in vragenlijsten, (b) een eventuele prioritering worden voorgesteld van deze factoren, en (c) suggesties worden gedaan voor factoren die het meest relevant zijn om op te nemen in een trendanalyse voor de periode 2023–2028.

Om de onderzoeksvragen te beantwoorden, voerden we een umbrella review uit, die bestond uit twee onderdelen:

- Een review van overzichtsstudies (d.w.z. meta-analyses en systematische reviews) naar de relatie tussen leesbegrip en kenmerken van het onderwijsleerproces, het leesgedrag en de leeshouding van leerlingen en leraren.
- Een review van (conceptuele) overzichtsstudies naar definities en operationalisaties van centrale begrippen rond leesgedrag en leeshouding.

In de volgende paragraaf beantwoorden we de eerste onderzoeksvraag en doen we aanbevelingen voor op te nemen factoren, voor een prioritering en voor een selectie van factoren voor de trendanalyse. In de daaropvolgende paragraaf beantwoorden we de tweede onderzoeksvraag: we geven daarin een overzicht van de uitkomsten van de conceptuele studies die we vonden.

## Onderzoeksvraag 1: kenmerken van leesonderwijs, leeshouding en leesgedrag

We vonden 56 literatuurstudies (29 meta-analyses, 25 systematische reviews en 2 combinaties). Op basis van een analyse van de bevindingen uit die literatuurstudies maakten we een categorisering van onderzochte variabelen. We brachten de uitkomsten van de studies onder in vier hoofdcategorieën: (1) kennis en vaardigheden, (2) leesgedrag en -houding, (3) kenmerken van de leeromgeving en (4) kenmerken van de schoolorganisatie; in vier literatuurstudies werd gekeken naar meerdere variabelen. Categorie (1) heeft betrekking op de in Onderzoeksvraag 1 genoemde domeinspecifieke onderwijskenmerken, terwijl categorie (3) betrekking heeft op de meer algemene onderwijskundige principes. In Tabel 4 presenteerden we een overzicht van de hoofd- en subcategorieën en gaven we omschrijvingen van elk van die (sub)categorieën. Om tot aanbevelingen te komen voor op te nemen factoren en een prioritering voor het peilingsonderzoek, bepaalden we vervolgens welke conclusies in elke literatuurstudie werden getrokken over relaties met leesbegrip (in geval van correlatieve onderzoek) en effecten op leesbegrip (in geval van interventie-onderzoek) alsmede wat de bewijskracht van de studie was. Die bevindingen presenteerden we in Tabel 5.

Als we kijken naar de aantallen studies dan vallen twee zaken op. Weliswaar gingen de meeste studies over zich normaal ontwikkelende leerlingen ( $k = 33$ ; 23 meta-analyses en 10 systematische reviews), maar toch waren er ook relatief veel studies over leerlingen met speciale onderwijsbehoeften ( $k = 23$ ; 6 meta-analyses, 15 systematische reviews en 2 combinaties). Daarnaast zijn er aanzienlijke verschillen in aantallen studies per (sub)categorie: 38 studies hadden betrekking op kennis en vaardigheden, 18 studies hadden betrekking op de leeromgeving en zeven hadden betrekking op leeshouding en -gedrag; in maar één studie werd gekeken naar kenmerken van de schoolorganisatie.<sup>23</sup> Binnen deze hoofdcategorieën was verreweg de meeste aandacht voor onderzoek naar zelfregulatie, metacognitie en leesstrategieën: in maar liefst 21 van de studies is gekeken naar relaties tussen deze variabelen en leesbegrip of naar de effecten van interventies die zich daarop richtten. Andere variabelen die in meerdere literatuurstudies werden onderzocht, zijn, op volgorde van frequentie: (mondelinge) taalvaardigheid ( $k = 6$ ), digitale technologie ( $k = 5$ ), vakintegratie ( $k = 5$ ), inferenties maken/achtergrondkennis ( $k = 4$ ), peer mediation/coöperatief

---

<sup>23</sup> De som van deze aantallen is groter dan 56, omdat in sommige studies naar meerdere variabelen werd gekeken.

leren ( $k = 4$ ), schrijfvaardigheid ( $k = 4$ ), technisch lezen ( $k = 4$ ), graphic organizers ( $k = 3$ ) en kennis van tekststructuur ( $k = 3$ ). Drie variabelen kwamen ieder in maar één studie terug: integreren van meerdere teksten, personalisatie en tutoring; die laatste twee in één en dezelfde meta-analyse, namelijk die van Baye et al. (2019).

Op basis van Tabel 5 kunnen we een aantal conclusies trekken over de onderzochte variabelen. We bespreken die in de hierna volgende onderdelen. We doen dat op volgorde van de mate waarin we ondersteuning vonden voor positieve relaties met en effecten op leesbegrip. Die volgorde suggereert een prioritering voor de selectie van factoren voor het peilingsonderzoek.

### Variabelen met systematische ondersteuning voor relaties/effecten

In twee gevallen vonden we vrij systematische ondersteuning voor een positieve relatie met of effecten op leesbegrip. Voor zelfregulatie, metacognitie en leesstrategieën lieten 18 van de 21 studies gunstige resultaten zien; voor 15 van die studies was de bewijskracht ten minste redelijk. We vonden maar drie literatuurstudies naar kennis van tekststructuur, maar in alle drie de gevallen was er sprake van positieve effecten en een redelijke bewijskracht. De resultaten die we vonden, golden zowel voor zich normaal ontwikkelende leerlingen als voor leerlingen met speciale onderwijsbehoeften. De uitkomsten betekenen dat instructie die erop is gericht leerlingen te helpen hun leesproces bewust aan te sturen door de inzet van leesstrategieën en instructie waarmee leerlingen inzicht krijgen in hoe teksten zijn georganiseerd, bijdragen aan leesbegrip. Die twee vormen van instructie staan overigens niet noodzakelijk los van elkaar: tekststructuurinstructie kan leerlingen helpen hun leesstrategieën toe te snijden op de aard van een tekst (Bogaerds-Hazenbergh, 2023).

### Variabelen met ondersteuning voor relaties/effecten voor leerlingen met speciale behoeften, maar niet voor zich normaal ontwikkelende leerlingen

Voor (mondelijke) taalvaardigheid zagen we in drie van de vier meta-analyses uitgevoerd voor zich normaal ontwikkelende leerlingen *geen* effecten op leesbegrip; in de vierde meta-analyse waren de effecten wisselend. In twee van deze studies werden de effecten van woordenschatinterventies onderzocht (Baye et al. 2019; Elleman et al., 2009). Die studies lieten geen positieve effecten op gestandaardiseerde toetsen zien, maar, in geval van Elleman et al., wel op door onderzoekers ontwikkelde tests. In die laatste tests bevatten de gebruikte teksten woorden die tijdens de instructie werden aangeleerd. Deze observatie sluit aan bij die uit eerder onderzoek waaruit bleek dat (expliciete) woordenschatinstructie vooral een positief effect heeft op het begrip van teksten waarin nieuw geleerde woorden voorkomen (Swart &

Verhoeven, 2022). Goodwin en Ahn (2013) onderzochten de effecten van morfologische instructie en concludeerden dat effecten afhankelijk zijn van leeftijd: ze vonden wel effecten voor jongere, maar niet voor oudere leerlingen. Rogde et al. (2019) vonden geen effecten van een brede variëteit aan taalbegripsinterventies.

De literatuurstudies gericht op leerlingen met speciale behoeften—een meta-analyse over leerlingen met gediagnosticeerde leerproblemen (Kaldenberg et al., 2015) en een systematische review over meertalige leerlingen (Cisco & Padrón, 2012)—lieten allebei wel positieve effecten zien op leesbegrip. In beide gevallen ging het om woordenschatinterventies. Mogelijk hebben zulke interventies juist voor leerlingen met speciale behoeften—vanwege hun vaak achterblijvende (mondelijke) taalvaardigheid—wel een grotere toegevoegde waarde. Interessant is dat Cisco en Padrón concludeerden dat woordenschatinterventies waarbij de eerste taal van leerlingen werd ingezet gunstige effecten sorteerden.

Ook voor peer mediation/coöperatief leren bleken er alleen effecten in studies naar leerlingen met speciale onderwijsbehoeften; die studies (Alzahrani & Leko, 2018; Okilwa & Shelby, 2010; Wexler et al., 2015) hadden alle drie redelijke bewijskracht. Deze conclusie sluit aan bij ander onderzoek waaruit bleek dat met name leerlingen met leerproblemen profijt hebben van instructievormen waarbij ze samenwerken met meer vaardige leerlingen (Van Ryzin et al., 2025). Peer mediation/coöperatief leren lijkt voor zich normaal ontwikkelende leerlingen niet effectief voor leesbegrip; wel tekenen we aan dat die conclusie is gebaseerd op maar één meta-analyse waarin uitsluitend naar uitkomsten op gestandaardiseerde toetsen is gekeken (Baye et al., 2019).

## Variabelen met ondersteuning voor relaties/effecten die maar voor een van beide doelgroepen zijn onderzocht

Sommige variabelen zijn uitsluitend voor een van de twee doelgroepen onderzocht. Zo vonden we alleen voor zich normaal ontwikkelende leerlingen literatuurstudies die zich richtten op inferenties maken/achtergrondkennis en op schrijfvaardigheid. In het eerste geval waren de uitkomsten van alle vier de studies—twee meta-analyses (Elleman, 2017; Rice & Wijekumar, 2024) en twee systematische reviews (Hattan et al., 2024; Sulaiman et al., 2023)—positief. Van de systematische reviews was de bewijskracht echter zwak: zo sloten Hattan et al. (2024) geen studies uit op grond van methodologische kwaliteit, maar voerden ze ook geen enkele kwaliteitscheck of controle op mogelijke publicatiebias uit. We hebben daardoor geen informatie over de kwaliteit van en balans tussen de primaire studies.

Conclusies over effecten kunnen dus het beste worden getrokken op basis van de meta-analyses, die wel een redelijke tot sterke bewijskracht hadden. Beide meta-analyses suggereren dat instructie gericht op het leren maken van inferenties en het ontwikkelen, activeren en inzetten van achtergrondkennis bijdraagt aan leesbegrip; de grootste effecten werden (logischerwijs) gevonden op inferentieel begrip ( $d = 0.68, p < .001$ ). Welke instructiestrategie daarbij het beste werkt, blijft onduidelijk: met name Elleman (2017) onderscheidde verschillende strategieën, maar onderzocht geen moderatoreffecten exclusief voor vo-leerlingen.

De vier studies die zich richtten op schrijfvaardigheid waren alle vier meta-analyses met een redelijke bewijskracht. Twee meta-analyses (Graham & Hebert, 2011; Graham et al., 2023) lieten positieve effecten zien en hoewel in beide studies verschillende instructiebenaderingen werden onderzocht, bleek alleen expliciete training in het uitvoeren van schrijfactiviteiten een significante moderator (Graham & Hebert, 2011). Op basis van de meta-analyse van Hebert et al. (2013) kunnen alleen conclusies worden getrokken over verschillen in effecten in type schrijfinstructie, niet van schrijfinstructie *an sich*: er bleken echter geen verschillen tussen de vier onderzochte instructietypen (samenvattingen schrijven, vragen beantwoorden, notities maken en extended writing). De meta-analyse van Baye et al. (2019) liet geen effecten van schrijfinstructie zien. Die uitkomst heeft niet noodzakelijk te maken met het gegeven dat de auteurs alleen resultaten op gestandaardiseerde toetsen analyseerden (zie eerder): Graham en Hebert (2011) maakten een vergelijking tussen de effecten op gestandaardiseerde en niet-gestandaardiseerde toetsen en vonden in beide gevallen positieve effecten.

We kwamen de inzet van graphic organizers tegen in verschillende literatuurstudies, maar vaak als onderdeel van een andere interventie (bijv. Elleman, 2017; Pyle et al., 2017; Okkinga et al., 2018). Drie systematische reviews gingen exclusief over het gebruik van graphic organizers en die hadden alle drie betrekking op leerlingen met speciale behoeften. Twee van die studies lieten positieve effecten zien (Kim et al., 2004; Payne & Yoon, 2024), één gemengde (Kim et al., 2012); in alle drie de gevallen was er sprake van redelijke bewijskracht. Kim et al. suggereren dat de wisselende uitkomsten in hun studie te maken hebben met het feit dat leerlingen niet in alle studies werden betrokken bij het maken van de graphic organizers, terwijl actieve verwerking nodig is om effecten te kunnen sorteren.

## Variabelen zonder ondersteuning voor relaties/effecten of met wisselende ondersteuning

Voor de overige variabelen waarnaar meerdere studies zijn gedaan, waren relaties/effecten afwezig of waren de uitkomsten gemengd, zowel in termen van effecten als in termen van bewijskracht. Dat eerste gold voor technisch lezen (decoderen, vloeiend lezen) en dan zowel voor zich normaal ontwikkelende leerlingen als voor leerlingen met speciale onderwijsbehoeften: de beide meta-analyses (Baye et al., 2019; Edmonds et al., 2009) en de beide systematische reviews (Steinle et al., 2022; Wexler et al., 2008) lieten geen van alle interventie-effecten op leesbegrip zien; in de laatste twee studies was er overigens wel sprake van effecten op fluency (Steinle et al., 2022; Wexler et al., 2008). In alle gevallen op één na (Edmonds et al., 2009) was de bewijskracht redelijk. Deze uitkomsten sluiten aan bij de veronderstelling dat het gewicht van technisch lezen bij de verklaring van verschillen in leesbegrip afneemt naarmate leerlingen ouder worden en een bepaald drempelniveau is bereikt (Denton et al., 2011; Van Steensel, Oostdam et al., 2016). Kennelijk geldt dat ook voor zwakke lezers en leerlingen met lees- en leerproblemen.

Uit de studies naar leesgedrag en -houding lijkt het beeld naar voren te komen dat, wanneer de bewijskracht redelijk is, er geen relaties of effecten waarneembaar zijn (Altamura et al., 2025; Bus et al., 2024; Van der Sande et al., 2023), terwijl er wel relaties/effecten zichtbaar zijn als de bewijskracht zwak is (Anderson et al., 2023; Mol & Bus, 2011; Toste et al., 2020). De focus van de studies varieerde. De meta-analyse van Altamura et al. (2025) ging over de samenhang tussen leesbegrip en digitale leesgewoonten: die correlatie was marginaal. De studies van Bus et al. (2024) en Mol en Bus (2011) richtten zich op de frequentie van het lezen van papieren boeken (of in elk geval op lezen in het algemeen). Uit de eerste studie lijkt naar voren te komen dat investeren in meer lezen op school (d.m.v. vrij lezen/stillezen) niet noodzakelijk leidt tot beter lezen, hoewel een recente meta-analyse wel liet zien dat toevoegingen aan vrij lezen dat gebrek aan effect kunnen compenseren, met name voor zwakke lezers (Merke et al., 2024).<sup>24</sup> De overige studies hadden betrekking op leesbetrokkenheid (Anderson et al., 2023) en leesmotivatie (Toste et al., 2020; Van der Sande et al., 2023). We vonden op basis van die studies onvoldoende ondersteuning voor de gedachte dat het veranderen van de houding van middelbare scholieren ten aanzien van lezen leidt tot beter begrip.

---

<sup>24</sup> We hebben de meta-analyse van Merke et al. (2024) overigens niet opgenomen in ons bestand, omdat de afhankelijke variabele een combinatie van begrijpend en technisch lezen was.

Voor digitale technologie waren zowel de uitkomsten als de bewijskracht wisselend. Maar één literatuurstudie—de systematische review van Wood et al. (2008)—liet positieve effecten zien van technologie (meer specifiek, van tekst-naar-spraaktechnologie) op het leesbegrip van leerlingen met speciale behoeften, maar we beoordeelden de bewijskracht van die studie als zwak. In de andere gevallen waren relaties/effecten wisselend of afwezig. Het onderwerp van die studies varieerde; wel gingen ze allemaal over zich normaal ontwikkelende leerlingen. Zowel Delgado et al. (2018) als Hare et al. (2024) onderzochten of tekstbegrip wordt bepaald door de vraag of een tekst wordt gelezen van papier of van een scherm: in geen van beide studies was sprake van een duidelijk voordeel van het ene of het andere medium. Zowel Berkeley et al. (2015) als Ter Beek et al. (2018) onderzochten de effecten van digitale tekstinterventies. We beoordeelden de bewijskracht van de eerste studie als redelijk en die van de tweede als zwak. De resultaten van de studie van Berkeley et al. (2015) varieerden: ze vonden wel een effect van digitale aanpassingen die het begrip van een tekst ondersteunen (denk aan beeld, geluid, animatie, concept maps of prompts), maar geen effect van aanpassingen die de toegankelijkheid van teksten vergroten (font aanpassen, vertaalfunctie).

Naar vakintegratie zijn vier literatuurstudies gedaan: één meta-analyse (Baye et al., 2019), twee systematische reviews (McCulley & Osman, 2015; Wade, 1983) en één gecombineerde meta-analyse en systematische review (Swanson et al., 2014). Twee studies waarvan we de bewijskracht als redelijk beoordeelden, lieten positieve effecten zien: die van McCulley en Osman (2015) en die van Swanson et al. (2014). In het laatste geval gold de beoordeling van de bewijskracht alleen voor de systematische review; voor de meta-analyse was die beoordeling zwak. In beide studies werden verschillende benaderingen gebruikt om vakintegratie vorm te geven, maar werden geen uitspraken gedaan over de effecten van specifieke aanpakken. De studies van Baye et al. (2019) en van Wade (1983)—beide met een redelijke bewijskracht—lieten respectievelijk geen en gemengde effecten zien van vakintegratie op leesbegrip. Bij Baye et al. (2019) ging het specifiek om instructie in metacognitieve strategieën gericht op kennisverwerving binnen bepaalde vakken; in die studie werden alleen resultaten van gestandaardiseerde toetsen meegenomen (zie eerder). Wade (1983) vond met name effecten wanneer vakintegratie onderdeel was van een breder pakket van interventies en wanneer leraren actief betrokken waren bij ontwerp en uitvoering.

In de laatste vier studies ging het om de effecten van meervoudige interventies, waarin verschillende dimensies van lees- en/of taalvaardigheid werden gecombineerd (fonologisch bewustzijn, decoderen, vloeiend lezen, woordenschat, kennis van tekststructuur,

achtergrondkennis). In alle gevallen op één na werd hun bewijskracht als redelijk beoordeeld; de bewijskracht van de studie van Garwood et al. (2014) was zwak. Voor zich normaal ontwikkelende leerlingen leek zo'n meervoudige aanpak niet tot resultaten te leiden: Paul en Clarke (2016) vonden geen effecten op leesbegrip. Voor leerlingen met speciale behoeften waren de effecten wisselend (Huddle et al., 2017; Garwood et al., 2014) of positief (Sohn et al., 2023).

### Variabelen waarnaar nog maar weinig literatuurstudies zijn gedaan

Naar interventies waarin leerlingen leren meerdere teksten te integreren, tutoring, personalisatie en kenmerken van de schoolorganisatie is maar weinig literatuuronderzoek gedaan. De laatste drie thema's zijn zelfs onderzocht in één en dezelfde meta-analyse (Baye et al., 2019). Alleen voor tutoring bleken positieve effecten waarvoor ook voldoende bewijskracht was.

### Consequenties voor het peilingsonderzoek

Voor het peilingsonderzoek laten de hierboven beschreven conclusies zien dat er met name ondersteuning is voor selectie van de volgende variabelen: (1) zelfregulatie, metacognitie en leesstrategieën, (2) kennis van tekststructuur, (3) mondelinge taalvaardigheid, (4) peer mediation/coöperatief leren, (5) inferenties maken/achtergrondkennis, (6) schrijfvaardigheid en (7) graphic organizers. Variabelen (3), (4) en (7) lijken vooral relevant voor leerlingen met speciale onderwijsbehoeften; de selectie van variabelen (5) en (6) is alleen gebaseerd op onderzoek naar zich normaal ontwikkelende leerlingen. Deze selectie wordt uiteraard beïnvloed door het beschikbare onderzoek (zie ook Lacunes in de review); onze conclusies nemen niet weg dat ook aan andere mogelijk relevante thema's aandacht kan worden besteed die in ons bestand onderbelicht zijn gebleven.

Een van de doelen van de umbrella review was na te gaan in hoeverre de uitkomsten ervan overeenstemmen met de conclusies uit eerdere overzichtsstudies naar kenmerken van effectief leesonderwijs (samengevat in Tabel 1). Onze bevindingen sluiten in belangrijke mate aan bij de kernelementen uit die eerdere studies: net als in die studies wijst onze review op het belang van achtergrondkennis, de blootstelling aan en instructie over verschillende teksttypen en genres, het toepassen van leesstrategieën, interactie en discussiëren (via peer mediation/coöperatief leren) en de integratie van lezen en schrijven (kernelement 2 tot en met 6). Voor de rol van leesmotivatie (kernelement 7) vonden we geen overtuigende ondersteuning. We vonden geen literatuurstudies waarin de twee andere kernelementen—(1)

monitoren en differentiëren en (8) lezen in authentieke situaties—werden onderzocht. Ook verschillende van de kenmerken van het onderwijsleerproces die in het vorige peilingsonderzoek in het voortgezet onderwijs (Inspectie van het Onderwijs, 2024) in kaart zijn gebracht, vonden we in de meta-analyses en systematische reviews niet terug. Het gaat dan bijvoorbeeld om het op schoolniveau stellen van doelen op het gebied van lezen, het aanbod van schoolbrede lees(bevorderings)activiteiten (denk aan de samenwerking met de bibliotheek), de tijd die wordt ingeroosterd voor leesonderwijs, het aanbod van extra (bij)lessen of de frequentie waarmee in verschillende genres wordt gelezen.

We adviseren om in de trendanalyse in elk geval de vragenlijstonderdelen op te nemen die te maken hebben met de variabelen die we hierboven noemden. We geven hieronder een (niet-uitputtende) lijst met voorbeelden:

1. Uit *Mate waarin leerlingen verschillende leesstrategieën nuttig vinden voor het begrijpen en onthouden van teksten*: “Ik vat de tekst in mijn eigen woorden samen” (variabele: zelfregulatie, metacognitie en leesstrategieën).
2. Uit *Leesactiviteiten tijdens de les Nederlands voor de verwerking van non-fictie*: “Evalueren en bekritisieren van de stijl of structuur van de tekst” (variabele: kennis van tekststructuur).
3. Uit *Leesactiviteiten tijdens de les Nederlands voor de verwerking van fictie*: “De betekenis opzoeken van onbekende woorden in de tekst” (variabele: [mondelijke] taalvaardigheid).
4. Uit *Leesactiviteiten tijdens de les Nederlands voor de verwerking van fictie*: “In kleine groepen discussiëren met andere leerlingen die dezelfde tekst hebben gelezen” (variabele: peer mediation/coöperatief leren).
5. Uit *Leesactiviteiten tijdens de les Nederlands voor de verwerking van fictie*: “De inhoud van het verhaal met hun eigen ervaringen vergelijken” (variabele: inferenties maken/achtergrondkennis).
6. Uit *Leesactiviteiten tijdens de les Nederlands voor de verwerking van fictie*: “Een tekst schrijven over iets dat verband houdt met wat ze gelezen hebben” (variabele: schrijfvaardigheid).

Voor de variabele graphic organizers vonden we geen voorbeelden in het voorgaande peilingsonderzoek.

## Onderzoeksvraag 2: definitie van concepten op het gebied van leesgedrag en -houding

We vonden vier conceptuele overzichtsstudies op het gebied van leesgedrag en -houding. De studies hadden betrekking op twee concepten: leesmotivatie en leesbetrokkenheid. De belangrijkste conclusie uit de studies is dat het in beide gevallen om complexe begrippen gaat die uit meerdere dimensies bestaan. Wat de twee studies naar leesmotivatie (Conradi et al., 2014; Schiefele et al., 2012) gemeen hebben, is dat ze de antecedenten of voorspellers van leesmotivatie (bijv. doelen, zelfpercepties, waarde, attitude) onderscheiden van “echte” leesmotivatie; zowel Conradi et al. (2014) als Schiefele et al. (2012) verwijzen bij dat laatste naar het onderscheid tussen intrinsieke en extrinsieke motivatie. Bij de definitie van leesbetrokkenheid wordt door zowel Lee et al. (2021) als door Unrau en Quirk (2014) onderscheid gemaakt tussen gedragsmatige, affectieve en cognitieve betrokkenheid; Lee et al. (2021) onderscheidde ook nog sociale betrokkenheid.

### Lacunae in de review

Zoals we eerder schetsten, was er aanzienlijke variatie in de beschikbaarheid van literatuurstudies. Daardoor zijn met name studies naar zelfregulatie, metacognitie en leesstrategieën oververtegenwoordigd. Andere thema's, die wellicht relevant zijn voor het peilingsonderzoek of waarom verzocht is bij het verstrekken van de opdracht, zijn juist onderbelicht gebleven; deels komt dat omdat het nog om relatief nieuwe onderzoeksdomeinen gaat. Het gaat dan bijvoorbeeld om het integreren van meerdere teksten, tutoring, personalisatie en kenmerken van de schoolorganisatie, waarover telkens maar één literatuurstudie is gevonden, maar ook om thema's waarover we überhaupt geen geschikte literatuurstudie vonden. Een voorbeeld van dat laatste is het leesgedrag en de leeshouding van leraren. Naar dat onderwerp is maar beperkt onderzoek gedaan, zo bleek uit een recente studie waarbij de derde auteur betrokken was (Kieft et al., 2022). Er zijn wel enkele studies waarin is gekeken naar de relatie tussen leesgedrag en -houding van leraren en de vormgeving van hun lesonderwijs (bijv. McKool & Gespass, 2009) of naar de effecten van leraarprofessionalisering op het gebied van leesbevordering op de leesmotivatie van leerlingen (bijv. De Naeghel et al., 2016), maar het type studies waarnaar wij op zoek waren, is vermoedelijk schaars. Een ander voorbeeld is de omgang met literaire teksten. Hoewel we termen als *literature* en *fiction* opnamen in onze zoekopdracht en tijdens onze expertraadpleging ook met literatuurwetenschappers hebben gesproken, vonden we geen

meta-analyses of systematische reviews waarin de effecten van literatuuronderwijs op leesbegrip zijn nagegaan. Ook gezien de aandacht die daarvoor is in de nieuwe (concept)kerndoelen is dat een gemis.

## Conclusie

Ter voorbereiding op de landelijke peiling naar leesvaardigheid aan het einde van leerjaar 2 van het vo, voerden we in opdracht van de Inspectie van het Onderwijs en NRO een umbrella review uit waarin we probeerden te achterhalen welke kenmerken van leesonderwijs en welke aspecten van het leesgedrag en de leeshouding van leerlingen en leraren bijdragen aan het leesbegrip van leerlingen in het voortgezet onderwijs, en hoe centrale concepten op het gebied van leesgedrag en -houding voor deze doelgroep gedefinieerd en geoperationaliseerd worden. Op basis van een inventarisatie van 56 meta-analyses en systematische reviews concludeerden we dat voor de volgende variabelen de meeste ondersteuning is voor een relatie met leesbegrip: (1) zelfregulatie, metacognitie en leesstrategieën, (2) kennis van tekststructuur, (3) (mondelinge) taalvaardigheid, (4) peer mediation/coöperatief leren, (5) inferenties maken/achtergrondkennis, (6) schrijfvaardigheid en (7) graphic organizers. We adviseren dan ook om in het peilingsonderzoek in elk geval aan deze variabelen aandacht te besteden. Verder bieden de vier conceptuele overzichtsstudies die we tegenkwamen een raamwerk om twee centrale concepten op het gebied van leeshouding en -gedrag te operationaliseren: leesmotivatie en leesbetrokkenheid. We zagen ook belangrijke lacunes: over het leren integreren van meerdere teksten, personalisatie, tutoring, schoolorganisatiekenmerken, het leesgedrag en de leeshouding van leraren, en de relatie tussen literatuuronderwijs en leesbegrip zijn nog niet of nauwelijks overzichtsstudies verschenen. Daar ligt een belangrijke opdracht voor vervolgonderzoek.

## Referenties

\*Referenties met een asterisk zijn geïnccludeerd in de umbrella review.

\*Altamura, L., Vargas, C., & Salmerón, L. (2025). Do new forms of reading pay off? A meta-analysis on the relationship between leisure digital reading habits and text comprehension. *Review of Educational Research*, 95(1), 53–88.

<https://doi.org/10.3102/00346543231216463>

\*Alzahrani, T., & Leko, M. (2018). The effects of peer tutoring on the reading comprehension performance of secondary students with disabilities: A systematic review. *Reading & Writing Quarterly*, 34(1), 1–17. <https://doi.org/10.1080/10573569.2017.1302372>

\*Anderson, L. L., Meline, M., & Harn, B. (2023). Student engagement within adolescent reading comprehension interventions: A systematic literature review. *Journal of Education*, 203(2), 258–268. <https://doi.org/10.1177/00220574211032327>

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, 13(3), 132–140. <https://doi.org/10.1097/XEB.0000000000000055>

\*Barzilai, S., Zohar, A. R. & Mor-Hagani, S. (2018). Promoting integration of multiple texts: a review of instructional approaches and practices. *Educational Psychology Review* 30, 973–999. <https://doi.org/10.1007/s10648-018-9436-8>

\*Baye, A., Inns, A., Lake, C., & Slavin, R. E. (2019). A synthesis of quantitative research on reading programs for secondary students. *Reading Research Quarterly*, 54(2), 133–166. <https://doi.org/10.1002/rrq.229>

\*Berkeley, S., Kurz, L. A., Boykin, A., & Evmenova, A. S. (2015). Improving reading comprehension using digital text: A meta-analysis of interventions. *International*

*Journal for Research in Learning Disabilities*, 2(2), 18–43. <https://journals.ub.uni-koeln.de/index.php/IJRLD/article/view/1853>

Bogaerds-Hazenberg, S.T.M. (2023). Text structure instruction in Dutch primary education: Building bridges between research and practice [Proefschrift, Universiteit Utrecht]. <https://doi.org/10.33540/1775>

Bogaerds-Hazenberg, S., Van den Bergh, H., & Evers-Vermeul, J. (2022). Onderwijs in tekststructuur. In Th. Houtveen & R. van Steensel (Eds.), *De zeven pijlers van onderwijs in begrijpend lezen* (pp. 85-109). Eburon. <https://www.lezen.nl/publicatie/kennisbundel-de-zeven-pijlers-van-onderwijs-in-begrijpend-lezen-gepubliceerd/>

Boute, D. (2024). *Begrijpen is meer dan lezen: Onderzoek naar principes van effectief begrijpend lezen in Blink Lezen* [Bachelorscriptie, Erasmus Universiteit Rotterdam].

\*Bus, A. G., Shang, Y., & Roskos, K. (2024). Building a stronger case for independent reading at school. *AERA Open*, 10. <https://doi.org/10.1177/23328584241267843>

Caird, J., Sutcliffe, K., Kwan, I., Dickson, K., & Thomas, J. (2015). Mediating policy-relevant evidence at speed: Are systematic reviews of systematic reviews a useful approach? *Evidence & Policy*, 11(1), 81-97. <https://doi.org/10.1332/174426514X13988609036850>

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest : A Journal of the American Psychological Society*, 19(1), 5–51. <https://doi.org/10.1177/1529100618772271>

\*Cisco, B. K., & Padrón, Y. (2012). Investigating vocabulary and reading strategies with middle grades English language learners: A research synthesis. *RMLE Online*, 36(4), 1–23. <https://doi.org/10.1080/19404476.2012.11462097>

- \*Conradi, K., Jang, B. G., & McKenna, M. C. (2014). Motivation terminology in reading research: A conceptual review. *Educational Psychology Review*, 26(1), 127–164.  
<https://doi.org/10.1007/s10648-013-9245-z>
- Cook, B., Buysse, V., Klingner, J., Landrum, T., McWilliam, R., Tankersley, M., & Test, D. (2014). *Council for exceptional children standards for evidence-based practices in special education*. Council for Exceptional Children.
- \*Cure, G., Batu, E. S., & Gulboy, E. (2021). Effectiveness of the story-mapping strategy in students with disabilities: Meta-analysis of single-case experimental design studies. *Reading & Writing Quarterly*, 37(6), 513–534. <https://doi-org.eur.idm.oclc.org/10.1080/10573569.2020.1858211>
- \*Daniel, J., Capin, P., & Steinle, P. (2021). A synthesis of the sustainability of remedial reading intervention effects for struggling adolescent readers. *Journal of Learning Disabilities*, 54(3), 170–186. <https://doi.org/10.1177/0022219420972184>
- \*Daniel, J., & Williams, K. J. (2021). Self-questioning strategy for struggling readers: A synthesis. *Remedial and Special Education*, 42(4), 248–261.  
<https://doi.org/10.1177/0741932519880338>
- Darch, C., & Eaves, R. C. (1986). Visual displays to increase comprehension of high school learning-disabled students. *The Journal of Special Education*, 20(3), 309–318.  
<https://doi.org/10.1177/002246698602000305>
- \*Davis, D. S. (2010). *A meta-analysis of comprehension strategy instruction for upper elementary and middle school students* (Publicatienr. 3430730) [Proefschrift, Vanderbilt University]. ProQuest Dissertations & Theses Global.  
<https://www.proquest.com/dissertations-theses/meta-analysis-comprehension-strategy-instruction/docview/769909285/se-2>

De Naeghel, J., Van Keer, H., Vansteenkiste, M., Haerens, L., & Aelterman, N. (2016).

Promoting elementary school students' autonomous reading motivation: Effects of a teacher professional development workshop. *The Journal of Educational Research, 109*(3), 232–252. <https://doi.org/10.1080/00220671.2014.942032>

\*Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review, 25*, 23–38. <https://doi.org/10.1016/j.edurev.2018.09.003>

Denton, C. A., Barth, A. E., Fletcher, J. M., Wexler, J., Vaughn, S., Cirino, P. T., Romain, M., & Francis, D. J. (2011). The relations among oral and silent reading fluency and comprehension in middle school: Implications for identification and instruction of students with reading difficulties. *Scientific Studies of Reading, 15*(2), 109–135. <https://doi-org.eur.idm.oclc.org/10.1080/10888431003623546>

Dixon, M., & Oakhill, J. (2024). Exploring teachers teaching reading comprehension: Knowledge, behaviours and attitudes. *Education, 52*(7), 963–978. <https://doi-org.eur.idm.oclc.org/10.1080/03004279.2024.2357884>

Dymock, S. (2007). Comprehension strategy instruction: Teaching narrative text structure awareness. *The Reading Teacher, 61*(2), 161–167. <https://doi.org/10.1598/RT.61.2.6>

\*Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., & Schnakenberg, J. W. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*(1), 262–300. <https://doi-org.eur.idm.oclc.org/10.3102/0034654308325998>

\*Elleman, A. M. (2017). Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic review.

*Journal of Educational Psychology*, 109(6), 761–781.

<http://dx.doi.org.eur.idm.oclc.org/10.1037/edu0000180>

\*Elleman, A., Lindo, E., Morphy, P., & Compton, D. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, 2(1), 1–44.

<https://doi.org/10.1080/19345740802539200>

\*Fan, W. (1993). *Metacognition and comprehension: A quantitative synthesis of metacognitive strategy instruction* (Publicatienr. 9407806) [Proefschrift, University of Cincinnati]. ProQuest Dissertations & Theses Global.

<https://www.proquest.com/dissertations-theses/metacognition-comprehension-quantitative/docview/304022051/se-2>

Filderman, M. J., Austin, C. R., Boucher, A. N., O'Donnell, K., & Swanson, E. A. (2022). A meta-analysis of the effects of reading comprehension interventions on the reading comprehension outcomes of struggling readers in third through 12th grades.

*Exceptional Children*, 88(2), 163-184. <https://doi.org/10.1177/00144029211050860>

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College*

*Composition and Communication*, 32(4), 365–387. <https://doi.org/10.2307/356600>

\*Garwood, J. D., Brunsting, N. C., & Fox, L. C. (2014). Improving reading comprehension and fluency outcomes for adolescents with emotional-behavioral disorders: Recent research synthesized. *Remedial and Special Education*, 35(3), 181-194. [https://doi-](https://doi.org.eur.idm.oclc.org/10.1177/0741932513514856)

[org.eur.idm.oclc.org/10.1177/0741932513514856](https://doi.org.eur.idm.oclc.org/10.1177/0741932513514856)

Geudens, A., Schraeyen, K., Bellens, K., Taelman, H., Trioen, M., Casteleyn, J., Simons, M., & Smits, T.F.H. (2022). *Les in lezen. Umbrella review van effectief leesonderwijs in het kleuter-, lager en secundair onderwijs*. Universiteit Antwerpen.

<https://hdl.handle.net/10067/1965490151162165141>

- Gobyn, S., Merchie, E., De Bruyne, E., De Smedt, F., Schiepers, M., Vanbuel, M., Verstedden, P., Van den Branden, K., Ghesquière, P., & Van Keer, H. (2019). *Sleutels voor effectief begrijpend lezen: Inspiratie voor een eigentijdse didactiek in het basisonderwijs*. Vlaamse Onderwijsraad. <https://biblio.ugent.be/publication/8632072>
- \*Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading, 17*(4), 257–285. <https://doi.org/10.1080/10888438.2012.689791>
- \*Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744. <https://doi.org/10.17763/haer.81.4.t2k0m13756113566>
- \*Graham, S., Kim, Y.-S., Cao, Y., Lee, W., Tate, T., Collins, P., Cho, M., Moon, Y., Chung, H. Q., & Olson, C. B. (2023). A meta-analysis of writing treatments for students in grades 6-12. *Journal of Educational Psychology, 115*(7), 1004–1027. <http://dx.doi.org.eur.idm.oclc.org/10.1037/edu0000819>
- Hajer, M. (2005). Taalgericht vakonderwijs: Tijd voor een nieuw vijfjarenplan. *Levende Talen Tijdschrift 6*(1). 3-11. <https://lt-tijdschriften.nl/ojs/index.php/lt/article/view/475>
- \*Hall-Mills, S. S., & Marante, L. M. (2022). Explicit text structure instruction supports expository text comprehension for adolescents with learning disabilities: A systematic review. *Learning Disability Quarterly, 45*(1), 55–68. <https://doi.org/10.1177/0731948720906490>
- \*Haller, E. P., Child, D. A., & Walberg, H. J. (1988). Can comprehension be taught? A quantitative synthesis of “metacognitive” studies. *Educational Researcher, 17*(9), 5–8. <https://doi-org.eur.idm.oclc.org/10.3102/0013189X017009005>
- \*Hare, C., Johnson, B., Vlahiotis, M., Panda, E. J., Tekok-Kilic, A., & Curtin, S. (2024). Children’s reading outcomes in digital and print mediums: A systematic review.

*Journal of Research in Reading*, 47(3), 309–329. <https://doi.org/10.1111/1467-9817.12461>

\*Hattan, C., Alexander, P. A., & Lupo, S. M. (2024). Leveraging what students know to make sense of texts: What the research says about prior knowledge activation. *Review of Educational Research*, 94(1), 73–111. <https://doi.org/10.3102/00346543221148478>

\*Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. (2016). The effects of text structure instruction on expository reading comprehension: A meta-analysis. *Journal of Educational Psychology*, 108(5), 609–629.  
<http://dx.doi.org.eur.idm.oclc.org/10.1037/edu0000082>

\*Hebert, M., Simpson, A., & Graham, S. (2013). Comparing effects of different writing activities on reading comprehension: A meta-analysis. *Reading and Writing: An Interdisciplinary Journal*, 26(1), 111–138. <https://doi.org/10.1007/s11145-012-9386-3>

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2(2), 127–160. <https://doi.org/10.1007/BF00401799>

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165–179. <https://doi.org/10.1177/001440290507100203>

Houtveen, Th., & Van Steensel, R. (2022). *De zeven pijlers van onderwijs in begrijpend lezen*. Eburon. <https://www.lezen.nl/publicatie/kennisbundel-de-zeven-pijlers-van-onderwijs-in-begrijpend-lezen-gepubliceerd/>

Houtveen, Th., Van Steensel, R., & De la Rie, S. (2019). *De vele kanten van leesbegrip. Literatuurstudie naar onderwijs in begrijpend lezen in opdracht van het Nationaal Regieorgaan Onderwijsonderzoek en de Inspectie van het Onderwijs*. Erasmus Universiteit Rotterdam. <http://hdl.handle.net/1765/119125>

- \*Huddle, S., Hosp, J., & Watt, S. (2017). Reading interventions for adolescent English language learners with reading difficulties: A synthesis of the evidence. *Multiple Voices for Ethnically Diverse Exceptional Learners*, 17(2), 3–18.  
<https://doi.org/10.56829/2158-396X.17.2.3>
- Inspectie van het Onderwijs (2024). *Peil.Leesvaardigheid: Einde leerjaar 2 voortgezet (speciaal) onderwijs schooljaar 2022-2023*. Inspectie van het Onderwijs.
- Institute of Education Sciences (2003). *What Works Clearinghouse study review standards*.
- Jang, B. G., Conradi, K., McKenna, M. C., & Jones, J. S. (2015). Motivation: Approaching an elusive concept through the factors that shape it. *Reading Teacher*, 69(2), 239–247.  
<https://doi.org/10.1002/trtr.1365>
- \*Kaldenberg, E. R., Watt, S. J., & Therrien, W. J. (2015). Reading instruction in science for students with learning disabilities: A meta-analysis. *Learning Disability Quarterly*, 38(3), 160–173. <https://doi.org/10.1177/0731948714550204>
- Kendeou, P., Van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning Disabilities Research & Practice*, 29(1), 10–16. <https://doi-org.eur.idm.oclc.org/10.1111/ldrp.12025>
- Kieft, M., Van Steensel, R., Jongstra, W., & Van den Eijnden, J. (2022). *Literatuurstudie leraren als leesbevorderaars*. Hogeschool KPZ. [https://hbo-kennisbank.nl/details/sharekit\\_kpz:oai:surfsharekit.nl:6cfcf58a-5fae-47c4-a878-745cb3790ebb](https://hbo-kennisbank.nl/details/sharekit_kpz:oai:surfsharekit.nl:6cfcf58a-5fae-47c4-a878-745cb3790ebb)
- \*Kim, A.-H., Vaughn, S., Wanzek, J., & Wei, S. (2004). Graphic organizers and their effects on the reading comprehension of students with LD: A synthesis of research. *Journal of Learning Disabilities*, 37(2), 105–118.  
<https://doi.org/10.1177/00222194040370020201>

- \*Kim, W., Linan-Thompson, S., & Misquitta, R. (2012). Critical factors in reading comprehension instruction for students with learning disabilities: A research synthesis. *Learning Disabilities Research & Practice, 27*(2), 66–78.  
<https://doi.org/10.1111/j.1540-5826.2012.00352.x>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26–38.  
<https://doi.org/10.1177/0741932512452794>
- Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly, 45*(2), 230–251. <https://doi.org/10.1598/RRQ.45.2.4>
- \*Lan, Y.-C., Lo, Y.-L., & Hsu, Y.-S. (2014). The effects of meta-cognitive instruction on students' reading comprehension in computerized reading contexts: A quantitative meta-analysis. *Educational Technology & Society, 17*(4), 186–202.  
<https://www.jstor.org/stable/jeductechsoci.17.4.186>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. <https://doi.org/10.2307/2529310>
- \*Lee, Y., Jang, B. G., & Conradi Smith, K. (2021). A systematic review of reading engagement research: What do we mean, what do we know, and where do we need to go? *Reading Psychology, 42*(5), 540–576.  
<https://doi.org/10.1080/02702711.2021.1888359>
- Maheady, L., Harper, G. F., & Sacca, M. K. (1988). Peer-mediated instruction: A promising approach to meeting the diverse needs of LD adolescents. *Learning Disability Quarterly, 11*(2), 108–113. <https://doi.org/10.2307/1510988>

- McCrudden, M. T., Kulikowich, J. M., Lyu, B., & Huynh, L. (2022). Promoting integration and learning from multiple complementary texts. *Journal of Educational Psychology*, *114*(8), 1832–1843. <https://doi.org/10.1037/edu0000746>
- \*McCulley, L. V., & Osman, D. J. (2015). Effects of reading instruction on learning outcomes in social studies: A synthesis of quantitative research. *The Journal of Social Studies Research*, *39*(4), 183–195. <https://doi.org/10.1016/j.jssr.2015.06.002>
- McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children’s attitudes toward reading: A national survey. *Reading Research Quarterly*, *30*(4), 934-956. <https://doi.org/10.2307/748205>
- McKool, S. S., & Gespass, S. (2009). Does Johnny’s reading teacher love to read? How teachers’ personal reading habits affect instructional practices. *Literacy Research and Instruction*, *48*(3), 264–276. <https://doi.org/10.1080/19388070802443700>
- Merke, S., Ganushchak, L., & Van Steensel, R. (2024). Effects of additions to independent silent reading on students’ reading proficiency, motivation, and behavior: Results of a meta-analysis. *Educational Research Review*, *42*.  
<https://doi.org/10.1016/j.edurev.2023.100572>
- \*Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267–296.  
<https://doi.org/10.1037/a0021890>
- O'Donnell, A. M. (2006). The role of peers and group learning. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 781–802). Lawrence Erlbaum Associates Publishers.
- \*Okilwa, N. S. A., & Shelby, L. (2010). The effects of peer tutoring on academic performance of students with disabilities in grades 6 through 12: A synthesis of the

literature. *Remedial and Special Education*, 31(6), 450–463.

<https://doi.org/10.1177/0741932509355991>

\*Okkinga, M., Van Steensel, R., Van Gelderen, A. J. S., Van Schooten, E., Slegers, P. J. C., & Arends, L. R. (2018). Effectiveness of reading-strategy interventions in whole classrooms: A meta-analysis. *Educational Psychology Review*, 30(4), 1215–1239.

<https://doi.org/10.1007/s10648-018-9445-7>

\*Paul, S.-A. S., & Clarke, P. J. (2016). A systematic review of reading interventions for secondary school students. *International Journal of Educational Research*, 79, 116–127. <https://doi.org/10.1016/j.ijer.2016.05.011>

\*Payne, S. B., & Yoon, N. Y. (2024). A synthesis of academic interventions for high school students with behavioral and learning difficulties. *Behavioral Disorders* 50(4).

<https://doi.org/10.1177/01987429241261969>

\*Peng, P., Wang, W., Filderman, M. J., Zhang, W., & Lin, L. (2023). The active ingredient in reading comprehension strategy intervention for struggling readers: A Bayesian network meta-analysis. *Review of Educational Research*, 94(2), 228-267.

<https://doi.org/10.3102/00346543231171345>

Petscher, Y. (2010). A meta-analysis of the relationship between student attitudes towards reading and achievement in reading. *Journal of Research in Reading*, 33(4), 335–355. <https://doi.org/10.1111/j.1467-9817.2009.01418.x>

Polanin, J. R., Maynard, B. R., & Dell, N. A. (2017). Overviews in education research: A systematic review and analysis. *Review of Educational Research*, 87(1), 172-

203. <https://doi.org/10.3102/0034654316631117>

Pollock, M., Fernandes, R. M., Becker, L. A., Pieper, D., Hartling, L. (2024). Chapter v: Overviews of reviews. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T.

- Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions version 6.5*. Cochrane. [cochrane.org/handbook](https://www.cochrane.org/handbook)
- \*Pyle, N., Vasquez, A. C., Lignugaris/Kraft, B., Gillam, S. L., Reutzel, D. R., Olszewski, A., Segura, H., Hartzheim, D., Laing, W., & Pyle, D. (2017). Effects of expository text structure interventions on comprehension: A meta-analysis. *Reading Research Quarterly*, 52(4), 469–501. <https://doi-org.eur.idm.oclc.org/10.1002/rrq.179>
- \*Rice, M., & Wijekumar, K. (2024). Inference skills for reading: A meta-analysis of instructional practices. *Journal of Educational Psychology*, 116(4), 569–589. <https://doi.org/10.1037/edu0000855>
- \*Rogde, K., Hagen, Åste M., Melby-Lervåg, M., & Lervåg, A. (2019). The effect of linguistic comprehension instruction on generalized language and reading comprehension skills: A systematic review. *Campbell Systematic Reviews*, 15(4). <https://doi.org/10.1002/cl2.1059>
- Rogiers, A., Bogaert, R., Van Ammel, K., Merchie, E., Van Keer, H. (2022). Leer leerlingen strategische lezers te worden. In Th. Houtveen & R. van Steensel (Eds.), *De zeven pijlers van onderwijs in begrijpend lezen* (pp. 85-109). Eburon. <https://www.lezen.nl/publicatie/kennisbundel-de-zeven-pijlers-van-onderwijs-in-begrijpend-lezen-gepubliceerd/>
- Salmerón, L., Strømsø, H. I., Kammerer, Y. A. K., Stadtler, M., & Van den Broek, P. (2018). Comprehension processes in digital reading. In M. Barzillai, J. Thomson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 91-120). John Benjamins Publishing Company. <https://doi.org/10.1075/swll.17>
- Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., & Torgesen, J. K. (2007). Interventions for adolescent struggling readers: A meta-

analysis with implications for practice. *Center on Instruction*.

<https://files.eric.ed.gov/fulltext/ED521837.pdf>

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4-12: 1980-2011. *Journal of Learning Disabilities*, 48(4), 369–390. <https://doi->

[ur.idm.oclc.org/10.1177/0022219413504995](https://doi-ur.idm.oclc.org/10.1177/0022219413504995)

\*Schiefele, U., Schaffner, E., Möller, J., Wigfield, A., Nolen, S., & Baker, L. (2012).

Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly*, 47(4), 427–463.

<https://doi.org/10.1002/RRQ.030>

Slavin, R. E. (2011). Instruction based on cooperative learning. In R.E. Mayer & P.A.

Alexander (Eds.), *Handbook of research on learning and instruction* (pp.344-360).

Routledge.

Snel, M. (2024). *Kwaliteitswaaier effectief onderwijs in begrijpend lezen*. Platform samen

onderzoeken. <https://www.platformsamenerzoeken.nl/wp->

[content/uploads/2023/01/Kwaliteitswaaier-Effectief-Onderwijs-in-Begrijpend-Lezen-versie-2024.pdf](https://www.platformsamenerzoeken.nl/wp-content/uploads/2023/01/Kwaliteitswaaier-Effectief-Onderwijs-in-Begrijpend-Lezen-versie-2024.pdf)

\*Sohn, H., Acosta, K., Brownell, M. T., Gage, N. A., Tompson, E., & Pudvah, C. (2023). A

meta-analysis of interventions to improve reading comprehension outcomes for adolescents with reading difficulties. *Learning Disabilities Research & Practice*,

38(2), 85–103. <https://doi.org/10.1111/ldrp.12307>

\*Solis, M., Ciullo, S., Vaughn, S., Pyle, N., Hassaram, B., & Leroux, A. (2012). Reading

comprehension interventions for middle school students with learning disabilities: A

synthesis of 30 years of research. *Journal of Learning Disabilities*, 45(4), 327–340.

<https://doi.org/10.1177/0022219411402691>

\*Steinle, P. K., Stevens, E. A., & Vaughn, S. (2022). Fluency interventions for struggling readers in grades 6 to 12: A research synthesis. *Journal of Learning Disabilities*, 55(1), 3–21. <https://doi.org/10.1177/0022219421991249>

\*Stevens, E. A., Park, S., & Vaughn, S. (2019). A review of summarizing and main idea interventions for struggling readers in grades 3 through 12: 1978–2016. *Remedial and Special Education*, 40(3), 131–149. <https://doi.org/10.1177/0741932517749940>

\*Sulaiman, Puspidalia, Y. S., Susanty, L., Lubis, I. A., Yunidar, & Talib, J. (2023). A systematic review of LRD (Listen-Read-Discuss) strategy to teaching reading. *Journal of Higher Education Theory and Practice*, 23(5).

<https://doi.org/10.33423/jhetp.v23i5.5927>

\*Sun, Y., Wang, J., Dong, Y., Zheng, H., Yang, J., Zhao, Y., & Dong, W. (2021). The relationship between reading strategy and reading comprehension: A meta-analysis. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.635289>

\*Swanson, E., Hairrell, A., Kent, S., Ciullo, S., Wanzek, J. A., & Vaughn, S. (2014). A synthesis and meta-analysis of reading interventions using social studies content for students with learning disabilities. *Journal of Learning Disabilities*, 47(2), 178–195.

<https://doi.org/10.1177/0022219412451131>

Swart, N. M., & Verhoeven, L. (2022). Zonder woordkennis geen begrip. In Th. Houtveen & R. van Steensel (Eds.), *De zeven pijlers van onderwijs in begrijpend lezen* (pp. 37-58).

Eburon. <https://www.lezen.nl/publicatie/kennisbundel-de-zeven-pijlers-van-onderwijs-in-begrijpend-lezen-gepubliceerd/>

- \*Ter Beek, M., Brummer, L., Donker, A. S., & Opdenakker, M.-C. J. L. (2018). Supporting secondary school students' reading comprehension in computer environments: A systematic review. *Journal of Computer Assisted Learning*, 34(5), 557–566.  
<https://doi.org/10.1111/jcal.12260>
- Timmers, M. (2024). *Gefundeerd leesonderwijs in de praktijk: Hoe geven actuele leesmethoden vorm aan wetenschappelijke inzichten over effectief onderwijs in begrijpend lezen?* [Bachelorscriptie, Erasmus Universiteit Rotterdam].
- \*Toste, J. R., Didion, L., Peng, P., Filderman, M. J., & McClelland, A. M. (2020). A meta-analytic review of the relations between motivation and reading achievement for K–12 Students. *Review of Educational Research*, 90(3), 420–456.  
<https://doi.org/10.3102/0034654320919352>
- \*Unrau, N. J., & Quirk, M. (2014). Reading motivation and reading engagement: Clarifying commingled conceptions. *Reading Psychology*, 35(3), 260–284.  
<https://doi.org/10.1080/02702711.2012.684426>
- Van den Broek, P., Helder, A., Espin, C., & Van der Liende, M. (2021). *Sturen op begrip: Effectief leesonderwijs in Nederland. Rapportage aan de Vaste 2e Kamer Commissie voor OCW*. Universiteit Leiden. <https://hdl.handle.net/1887/3567187>
- Van der Hoeven, J. Kooiker-den Boer, H., van den Bergh, H., & Evers-Vermeul, J. (2022). Combineren van lees- en schrijfonderwijs loont. In Th. Houtveen & R. van Steensel (Eds.), *De zeven pijlers van onderwijs in begrijpend lezen* (pp. 129-146). Eburon.  
<https://www.lezen.nl/publicatie/kennisbundel-de-zeven-pijlers-van-onderwijs-in-begrijpend-lezen-gepubliceerd/>
- \*Van Ryzin, M. J., Murray, C., & Roseth, C. J. (2024). The effects of cooperative learning on self-reported peer relations, peer support, and classroom engagement among students

with disabilities. *The Journal of Educational Research*, 117(6), 355-364.

<https://doi.org/10.1080/00220671.2024.2410494>

\*Van der Sande, L., Van Steensel, R., Fikrat-Wevers, S., & Arends, L. (2023). Effectiveness of interventions that foster reading motivation: A meta-analysis. *Educational Psychology Review*, 35(1). <https://doi.org/10.1007/s10648-023-09719-3>

Van Steensel, R. (2022). Zorg voor een motiverende leesomgeving. In Th. Houtveen & R. van Steensel (Eds.), *De zeven pijlers van onderwijs in begrijpend lezen* (pp. 147-163). Eburon. <https://www.lezen.nl/publicatie/kennisbundel-de-zeven-pijlers-van-onderwijs-in-begrijpend-lezen-gepubliceerd/>

Van Steensel, R., Oostdam, R., van Gelderen, A., & van Schooten, E. (2016). The role of word decoding, vocabulary knowledge and meta-cognitive knowledge in monolingual and bilingual low-achieving adolescents' reading comprehension. *Journal of Research in Reading*, 39(3), 312-329. <https://doi.org/10.1111/1467-9817.12042>

Van Steensel, R., Van der Sande, N.E., Bramer, W.M., & Arends, L. (2016). *Effecten van leesmotivatie-interventies: Uitkomsten van een meta-analyse. Reviewstudie in opdracht van NRO*. Erasmus Universiteit Rotterdam. <http://hdl.handle.net/1765/110821>

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, 1, 3-14. <https://doi.org/10.1007/s11409-006-6893-0>

\*Wade, S. E. (1983). A synthesis of the research for improving reading in the social studies. *Review of Educational Research*, 53(4), 461-497. <https://doi-org.eur.idm.oclc.org/10.3102/00346543053004461>

- \*Wexler, J., Reed, D. K., Pyle, N., Mitchell, M., & Barton, E. E. (2015). A synthesis of peer-mediated academic interventions for secondary struggling learners. *Journal of Learning Disabilities, 48*(5), 451–470. <https://doi.org/10.1177/0022219413504997>
- \*Wexler, J., Vaughn, S., Edmonds, M., & Reutebuch, C. K. (2008). A synthesis of fluency interventions for secondary struggling readers. *Reading and Writing: An Interdisciplinary Journal, 21*(4), 317–347. <https://doi.org/10.1007/s11145-007-9085-7>
- \*Wood, S. G., Moxley, J. H., Tighe, E. L., & Wagner, R. K. (2018). Does use of text-to-speech and related read-aloud tools improve reading comprehension for students with reading disabilities? A meta-analysis. *Journal of Learning Disabilities, 51*(1), 73–84. <https://doi.org/10.1177/0022219416688170>
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50031-7>

## Bijlage 1: Zoektermen per database

### ERIC

(\* Reading/ OR exp "Reading Instruction"/ OR "Literacy Education"/ OR "Reading Motivation"/ OR exp "Reading Skills"/ OR "Reading Teachers"/ OR "Reading Strategies"/ OR "Silent Reading"/ OR "Reading Improvement"/ OR "Reading Achievement"/ OR "Reading Research"/ OR "Reading Ability"/ OR "Reading Attitudes"/ OR "Reading Habits"/ OR "Reading Programs"/ OR "Recreational Reading"/ OR (((Reading OR reader\* OR Literac\*) ADJ3 (Instruct\* OR educat\* OR Content-Area OR silent\* OR individuali\* OR Motivat\* OR skill\* OR teach\* OR strateg\* OR comprehens\* OR struggl\* OR proficien\* OR improve\* OR Achieve\* OR abilit\* OR Research OR promotion\* OR attitude\* OR habit\* OR Program\* OR Recreation\* OR Performan\* OR intervention\* OR didactic\* OR literature\* OR fiction\* OR leisure\*)) OR ((multiple-text\* OR multitext\*) ADJ3 (comprehen\* OR integrat\*))).ab,ti OR (Reading OR reader\* OR Literac\*).ti.) AND ("Secondary Education"/ OR "Elementary Secondary Education"/ OR "Grade 7"/ OR "Grade 8"/ OR "Grade 9"/ OR "Grade 10"/ OR "Grade 11"/ OR "Grade 12"/ OR "High Schools"/ OR "High School Students"/ OR "Junior High Schools"/ OR "Junior High School Students"/ OR "Secondary School Curriculum"/ OR "Secondary School Students"/ OR "Secondary School Teachers"/ OR "Secondary Schools"/ OR Adolescents/ OR "Early Adolescents"/ OR "Late Adolescents"/ OR (Secondary-Education OR highschool\* OR high-school\* OR Junior-High OR JuniorHigh OR Adolescen\* OR ((Grade\* OR year\*) ADJ2 ("7" OR "8" OR "9" OR "10" OR "11" OR "12"))).ab,ti OR (k-12 OR "3-12" OR "4-12").ti.) AND ("Meta Analysis"/ OR ((systematic ADJ3 review\*) OR Meta-Analy\* OR (research ADJ3 synthesis\*)).ab,ti OR ((systematic AND review) OR synthesis\*).ti.) NOT ((health-litera\*).ti.)

### PsycINFO

(\* Reading/ OR "Reading Ability"/ OR "Reading Achievement"/ OR "Reading Comprehension"/ OR "Reading Development"/ OR "Reading Education"/ OR "Reading Materials"/ OR "Reading Skills"/ OR "Reading Speed"/ OR "Remedial Reading"/ OR "Silent Reading"/ OR (((Reading OR reader\* OR Literac\*) ADJ3 (Instruct\* OR educat\* OR Content-Area OR silent\* OR individuali\* OR Motivat\* OR skill\* OR teach\* OR strateg\* OR comprehens\* OR struggl\* OR proficien\* OR improve\* OR Achieve\* OR abilit\* OR Research OR promotion\* OR attitude\* OR habit\* OR Program\* OR Recreation\* OR

Performan\* OR intervention\* OR didactic\* OR literature\* OR fiction\* OR leisure\*)) OR ((multiple-text\* OR multitext\*) ADJ3 (comprehen\* OR integrat\*))).ab,ti OR (Reading OR reader\* OR Literac\*).ti.) AND ("Secondary Education"/ OR "High School Education"/ OR "Junior High Schools"/ OR "High Schools"/ OR 200.ag. OR (Secondary-Education OR highschool\* OR high-school\* OR Junior-High OR JuniorHigh OR Adolescen\* OR ((Grade\* OR year\*) ADJ2 ("7" OR "8" OR "9" OR "10" OR "11" OR "12"))).ab,ti OR (k-12 OR "3-12" OR "4-12").ti.) AND ("Meta Analysis"/ OR "Systematic Review"/ OR ((systematic ADJ3 review\*) OR Meta-Analy\* OR (research ADJ3 synthesis\*))).ab,ti OR ((systematic AND review) OR synthesis\*).ti.) NOT ((health-litera\*).ti.)

### **Web of science**

(TS=(((Reading OR reader\* OR Literac\*) NEAR/2 (Instruct\* OR educat\* OR Content-Area OR silent\* OR individuali\* OR Motivat\* OR skill\* OR teach\* OR strateg\* OR comprehens\* OR struggl\* OR proficien\* OR improve\* OR Achieve\* OR abilit\* OR Research OR promotion\* OR attitude\* OR habit\* OR Program\* OR Recreation\* OR Performan\* OR intervention\* OR didactic\* OR literature\* OR fiction\* OR leisure\*)) OR ((multiple-text\* OR multitext\*) NEAR/2 (comprehen\* OR integrat\*))) OR TI=(Reading)) AND (TS=(Secondary-Education OR highschool\* OR high-school\* OR Junior-High OR JuniorHigh OR Adolescen\* OR ((Grade\* OR year\*) NEAR/2 ("7" OR "8" OR "9" OR "10" OR "11" OR "12"))) OR TI=(k-12 OR "3-12" OR "4-12")) AND (TS=((systematic NEAR/2 review\*) OR Meta-Analy\* OR (research NEAR/2 synthesis\*)) OR TI=((systematic AND review) OR synthesis\*)) NOT TI=((health-litera\*)) NOT DT=(Meeting Abstract OR Meeting Summary) AND LA=(English)

### **Scopus**

(TITLE-ABS-KEY(((Reading OR reader\* OR Literac\*) W/2 (Instruct\* OR educat\* OR Content-Area OR silent\* OR individuali\* OR Motivat\* OR skill\* OR teach\* OR strateg\* OR comprehens\* OR struggl\* OR proficien\* OR improve\* OR Achieve\* OR abilit\* OR Research OR promotion\* OR attitude\* OR habit\* OR Program\* OR Recreation\* OR Performan\* OR intervention\* OR didactic\* OR literature\* OR fiction\* OR leisure\*)) OR ((multiple-text\* OR multitext\*) W/2 (comprehen\* OR integrat\*))) OR TITLE(Reading)) AND (TITLE-ABS-KEY(Secondary-Education OR highschool\* OR high-school\* OR Junior-High OR JuniorHigh OR Adolescen\* OR ((Grade\* OR year\*) W/2 ("7" OR "8" OR "9" OR "10" OR "11" OR "12"))) OR TITLE(k-12 OR "3-12" OR "4-12")) AND (TITLE-

ABS-KEY((systematic W/2 review\*) OR Meta-Analy\* OR (research NEAR/2 synthesis\*))  
OR TITLE((systematic AND review) OR synthesis\*) AND NOT TITLE((health-litera\*))  
AND NOT DOCTYPE(ab) AND LANGUAGE(English)

### **Google scholar**

'Reading

Instruction|education|Motivation|skills|teaching|strategies|comprehension|intervention|proficiency|promotion' 'Secondary Education'|highschool'|high school'|Adolescents 'systematic review'|'Meta-Analysis' -'health-literacy'

### **Google Nederlandstalig**

"lees

onderwijs|motivatie|vaardigheid|bevordering"|leesonderwijs|leesmotivatie|leesvaardigheid|leesbevordering "voortgezet|secundair onderwijs"|tieners|vo filetype:pdf

allintitle:"lees

onderwijs|motivatie|vaardigheid|bevordering"|leesonderwijs|leesmotivatie|leesvaardigheid|leesbevordering onderwijs filetype:pdf

### **Web of science (voor conceptuele studies)**

TI=((motivation\* OR attitude\* OR engagement\* OR self-efficacy\* OR self-concept\* OR confidence\* OR behav\* OR habits\* OR engag\* OR frequency\* OR time\* OR activity\* OR voluntary\* OR leisure\* OR comprehend\*) NEAR/5 reading) AND TI=(conceptu\* OR conception\* OR construct\* OR dimension\* OR classific\* OR defin\* OR clarif\* OR consens\* OR validit\* OR validat\* OR distinct\* OR correlate\*) NOT DT=(Meeting Abstract OR Meeting Summary) AND LA=(English)

## Bijlage 2: Codeerschema voor reviews van interventies

1	<p><b>Citation details</b></p> <p>According to APA 7 guidelines</p>	
2	<p><b>Type of study</b></p> <p>(Meta-analysis/systematic review/conceptual review/other review, namely ...)</p>	
3	<p><b>Goal of the study and/or research questions and/or hypotheses</b></p> <p>Only report the research questions and hypotheses that are relevant to this umbrella review.</p>	
4	<p><b>Number of studies included</b></p>	
5	<p><b>Research design of included studies</b></p> <p>(quasi-)experimental studies, correlational studies, etc. Also report the number of studies per research design.</p>	
6	<p><b>Population + participants</b></p> <p>Total number of participants and a general description of participants and distribution of participants in terms of:</p> <ul style="list-style-type: none"> <li>• Age/grade level,</li> <li>• Gender</li> </ul>	

	<ul style="list-style-type: none"> <li>• Special demographic (reading problems, reading disorders, bilingualism, etc.)</li> <li>• Background characteristics (SES, education level, culture, etc.)</li> </ul> <p>(If the number of participants per demographic was not provided, report the number of studies per demographic)</p>	
7	<p><b>Quality assessment</b></p> <ul style="list-style-type: none"> <li>• Was there an assessment of study quality of the empirical studies (yes/no)?</li> <li>• What instrument or method was used to assess the quality of the empirical studies?</li> <li>• What aspects were included in the quality assessment? (e.g. research design, implementation fidelity, control group, randomization, counterbalancing, etc.)</li> <li>• Was the presence of (publication) bias examined? If yes, how?</li> <li>• Were studies excluded based on (low) quality? (Report whether only peer-reviewed articles were included as well)</li> </ul>	
8	<p><b>Context</b></p> <ul style="list-style-type: none"> <li>• School, lab, or leisure?</li> <li>• School subject (language, social studies, geography, etc.)</li> <li>• Type of lesson (group size, online/offline education, etc.)</li> </ul>	

	<ul style="list-style-type: none"> <li>• Geographical location(s) of included studies (i.e., countries where the primary studies were conducted, or urban/rural contexts)</li> <li>• <i>If this is not coded as an independent variable:</i> were the interventions implemented by researchers or teachers?</li> </ul>	
<b>9</b>	<p><b>Analysis</b></p> <p>In what way were the studies analyzed? How was the coding procedure done (particularly, what measures were taken to increase reliability, e.g., multiple coders)?</p> <p>For meta-analyses: What type of analysis was used (fixed/random-effects model)?</p> <p>For systematic reviews: How were the studies synthesized?</p>	
<b>10</b>	<b>Timespan of included studies</b>	
<b>11</b>	<b>Language of the review and the included primary studies</b>	
<b>12</b>	<p><b>Independent variable(s)</b></p> <ul style="list-style-type: none"> <li>• Which independent variables were examined?</li> <li>• How were these (theoretically) defined?</li> <li>• How were these measured?</li> <li>• <i>For interventions:</i> what did these look like? (e.g. how was a certain strategy operationalized?)</li> </ul>	

	<p>Include participant variables, study design features, etc., not just intervention characteristics.</p>	
<p><b>13</b></p>	<p><b>Dependent variable(s)</b></p> <ul style="list-style-type: none"> <li>• What is/are the central or dependent variable(s)?</li> <li>• How is/are these (theoretically) defined? (Use the authors' definition and indicate whether such a definition was absent.)</li> <li>• How is/are these measured?</li> </ul>	
<p><b>14</b></p>	<p><b>Relevant results</b></p> <p>Provide the results of the relevant research questions/goals from point 3.</p> <p>Report the exact effect sizes, coefficients, and correlations.</p> <ul style="list-style-type: none"> <li>• For systematic reviews, provide a short description of the reported findings by the authors.</li> </ul>	

15	<p><b>Relevant moderators</b></p> <p><i>Only for meta-analyses</i></p> <p>Significant moderators on the level of, for example:</p> <ul style="list-style-type: none"> <li>• Age/ grade level (report this moderator when the population of the review is broader than secondary education/adolescents, even if it is not significant)</li> <li>• Features of (reading) education</li> <li>• Type of measurement instrument (or test)</li> </ul> <p>Report the effect size of each significant moderator.</p>	
16	<p><b>Limitations</b></p> <p>What limitations do the authors provide for the relevant research questions?</p>	
17	<p><b>Conclusions</b></p> <p>What answers do the authors provide for the relevant research questions (point 3)?</p>	
18	<p><b>Comments</b></p> <p>This is the space to name certain aspects of the review that didn't fit into other parts of the coding scheme.</p>	

## Bijlage 3: Codeerschema voor conceptuele studies

1	<p><b>Citation details</b></p> <p>According to APA 7 guidelines</p>	
2	<p><b>Concept</b></p> <p>Which concepts were addressed? (Reading motivation, reading engagement, reading behavior, etc.)</p>	
3	<p><b>Purpose/research questions</b></p> <p>What was the purpose/aim of the conceptual review? What were the research questions?</p>	
4	<p><b>Demographic</b></p> <p>Was the conceptual review aimed at a specific population? If yes, which population? If not, what do the authors report about the age of the participants of included studies?</p>	
5	<p><b>Method</b></p> <p>What did the authors report about the</p>	

	<p>methodology of the review?</p> <p>If possible, provide information on:</p> <ul style="list-style-type: none"> <li>• Inclusion criteria</li> <li>• Coding procedures (how many coders?)</li> </ul>	
6	<b>Number of included studies</b>	
7	<b>Timespan of included studies</b>	
8	<p><b>Theories</b></p> <p>Which theories were described by the authors? What do these theories pose?</p>	
9	<p><b>Definition</b></p> <p>How was the concept defined? What operationalizations were provided? Here, a ‘consensus definition’ can be described, as well as the definitions found in the included studies.</p>	

<b>10</b>	<b>Reading comprehension</b> Did the authors draw their own conclusions about the relationship between the addressed construct and reading comprehension?	
<b>11</b>	<b>Comments</b> This is the space to describe certain aspects of the review that didn't fit into other parts of the coding scheme.	

## Bijlage 3: Bewijskracht en effecten per studie

Studie	Bewijskracht (+, ~, -)	Effect (+, ~, 0)
Elleman (2017)	+	+
Rice & Wijekumar (2024)	~	+
Hebert et al. (2016)	~	+
Pyle et al. (2019)	~	+
Elleman et al. (2009)	~	~
Goodwin & Ahn (2013)	~	0
Rogde et al. (2019)	~	0
Graham & Hebert (2011)	~	+
Graham et al. (2023)	~	+
Hebert et al. (2013)	~	n.v.t.*
Sun et al. (2021)	-	+
Davis (2010)	~	+
Fan (1993)	~	+
Haller et al. (1993)	~	+
Okkinga et al. (2018)	~	~
Altamura et al. (2025)	~	0
Mol & Bus (2011)	-	+
Bus et al. (2024)	~	0
Toste et al. (2020)	-	+
van der Sande et al. (2023)	~	0
Berkeley et al. (2015)	~	~
Delgado et al. (2018)	~	0
Baye et al. (2019)	~	~
Cure et al. (2021)	~	+
Peng et al. (2023)	+	+
Wood et al. (2018)	-	+
Edmonds et al. (2009)	-	+
Kaldenberg et al. (2015)	~	+
Sohn et al. (2023)	~	+
Hattan et al. (2024)	-	+
Sulaiman et al. (2023)	-	+
Barzilai et al. (2018)	-	+
Lan et al. (2014)	-	+
Anderson et al. (2023)	-	+
Hare et al. (2024)	~	~
ter Beek et al. (2018)	-	~
McCulley & Osman (2015)	~	+
Wade (1983)	~	~
Paul & Clarke (2016)	~	0
Hall-Mills et al. (2022)	~	+
Steinle et al. (2022)	~	0
Wexler et al. (2008)	~	0
Daniel et al. (2021)	~	+
Daniel & Williams (2021)	~	~

<b>Studie</b>	<b>Bewijskracht (+, ~, -)</b>	<b>Effect (+, ~, 0)</b>
Solis et al. (2012)	~	+
Kim et al. (2004)	~	+
Alzahrani & Leko (2018)	~	+
Okilwa & Shelby (2010)	~	+
Wexler et al. (2015)	~	+
Cisco & Padrón (2012)	~	+
Garwood et al. (2014)	-	~
Huddle et al. (2017)	~	~
Kim et al. (2012)	~	~
Payne & Yoon (2024)	~	+
Stevens et al. (2019) (meta-analyse)	~	+
Stevens et al. (2019) (systematische review)	+	+
Swanson et al. (2014) (meta-analyse)	-	+
Swanson et al. (2014) systematische review	~	+

**Noot:** De studies zijn gepresenteerd in de volgorde waarin ze in het resultatenhoofdstuk voorkomen.

\*De studie van Hebert et al. (2013) vergeleek verschillende soorten schrijfinstructie, en gaf dus geen inschatting van de effecten van schrijfinstructie op zich.