



university of
 groningen

Autonomiebehoefte en ervaren autonomieondersteuning

Een systematische review en psychometrische kwaliteitsanalyse van vragenlijsten gericht op autonomiebehoefte en ervaren autonomieondersteuning van leerlingen.

Student: Léon Postma (S3185613)

Eerste beoordelaar: Dr. R. Hoekstra

Tweede beoordelaar: Dr. M. Hingstman

Woord aantal: 8894

Rijksuniversiteit Groningen

Faculteit der Gedrags- en Maatschappijwetenschappen

Masterthesis Onderwijswetenschappen (Track Onderwijsinnovatie)

In samenwerking met: D. Džananović, S.E. van de Laar, E.C. van der Kooi

Januari 2026

Abstract

Motivation plays an important role in students' learning processes and academic performance. Within Self-Determination Theory (SDT), autonomy, competence, and relatedness form the three basic psychological needs for motivation. This study focuses on students' autonomy need and their perceived autonomy support in primary and secondary education. To assess these constructs adequately, reliable and valid measurement instruments are needed. To create an overview of the available instruments and their psychometric quality, a systematic review was conducted following the PRISMA guidelines. The aim was to provide an up-to-date overview of questionnaires used to measure autonomy and autonomy support in educational settings and to evaluate their psychometric properties. Four databases were searched (ERIC, Web of Science, Scopus, and PubMed). This resulted in twelve included studies. The instruments were assessed using the COSMIN guidelines and with Dima's (2018) six-step framework for scale development. The results showed variation in the conceptualisation of autonomy across instruments, limiting direct comparability. Internal consistency and structural validity were generally well reported. Test-retest reliability was examined only minimally. None of the studies reported measurement error according to COSMIN definitions (e.g., SEM or SDC), although modern psychometric models did incorporate aspects of measurement error implicitly. Responsiveness was not reported in any of the studies. The Dima analysis indicated that most instruments had been validated primarily at the scale level, while item-level analyses (such as Mokken, IRT, or Rasch) were not used. Instruments such as the MD-PASS-PE, IBQ-PE, SIS/SIS-PE, and BPNSFS showed relatively strong psychometric properties. It is important to note that this review provides an initial overview of more recent instruments and that the conclusions should be interpreted with caution given the limitations of this study. Future research may benefit from combining classical COSMIN indicators with modern model-based psychometric approaches to obtain a more complete and reliable understanding of autonomy instruments and their practical applicability.

Keywords: autonomy, Self-Determination Theory, autonomy support, students, measurement instruments, psychometric properties, systematic review, PRISMA, COSMIN.

Samenvatting

Motivatie speelt een belangrijke rol in het leerproces en de schoolprestaties van leerlingen. Binnen de zelfdeterminatietheorie (SDT) vormen autonomie, competentie en verbondenheid de drie basisbehoeften die motivatie versterken. In dit onderzoek staat de behoefte aan autonomie en de ervaren autonomieondersteuning binnen het primair en voortgezet onderwijs centraal. Om deze elementen goed in kaart te brengen zijn betrouwbare en valide instrumenten nodig. Om hiervan een overzicht te verkrijgen is een systematische review uitgevoerd gebaseerd op de PRISMA-richtlijnen. Het doel was om een actueel overzicht te bieden van beschikbare vragenlijsten en hun psychometrische kwaliteit. Er zijn vier databases doorzocht (ERIC, Web of Science, Scopus en PubMed), wat uiteindelijk resulteerde in twaalf geïnccludeerde studies. De instrumenten werden beoordeeld volgens de COSMIN-richtlijnen en daarna aangevuld met het zesstappenmodel van Dima (2018). De resultaten suggereren dat er verschillen waren in de conceptualisering van autonomie, wat de vergelijkbaarheid beperkt. De interne consistentie en factorstructuur waren over het algemeen goed onderbouwd. Test-hertestbetrouwbaarheid werd slechts beperkt onderzocht. Geen van de artikelen rapporteerde meetfout volgens COSMIN-richtlijnen (zoals SEM of SDC), terwijl meetfout in moderne psychometrische modellen wel impliciet werd meegenomen. Responsiviteit werd in geen van de geïnccludeerde studies gerapporteerd. Uit de Dima-analyse bleek dat de meeste instrumenten vooral op schaalniveau zijn gevalideerd, terwijl itemniveau-analyses (zoals Mokken, IRT of Rasch) bijna niet worden gebruikt. Instrumenten zoals de MD-PASS-PE, IBQ-PE, SIS/SIS-PE en BPNSFS lieten sterke psychometrische waarden zien. Het is belangrijk te benoemen dat deze review een eerste overzicht biedt van recentere beschikbare vragenlijsten, maar dat de conclusies en resultaten voorzichtigheid moeten worden geïnterpreteerd gezien de beperkingen die dit onderzoek met zich meebrengt. In toekomstig onderzoek zouden klassieke COSMIN-indicatoren wellicht kunnen worden gecombineerd met moderne modelgebaseerde analyses om zo tot een vollediger en betrouwbaarder psychometrisch beeld van autonomie-instrumenten en hun bruikbaarheid te komen.

Trefwoorden: Autonomie, zelfdeterminatietheorie, autonomie-ondersteuning, leerlingen, meetinstrumenten, psychometrische eigenschappen, systematische review, PRISMA, COSMIN.

Autonomiebehoefte en -ondersteuning vanuit de zelfdeterminatietheorie

Context

In deze scriptie staan de autonomiebehoefte van leerlingen en de mate waarin zij autonomieondersteuning ervaren centraal. Om de relevantie hiervan te onderbouwen wordt gebruikgemaakt van de zelfdeterminatietheorie (SDT) van Deci en Ryan (2000). Binnen SDT wordt autonomie beschreven als een psychologische basisbehoefte die samenhangt met motivatie en daarmee met het leerproces van leerlingen (Deci & Ryan, 2000; Ryan & Deci, 2020; Vansteenkiste et al., 2012).

Zelfdeterminatietheorie (SDT)

Een belangrijke factor binnen het leerproces van leerlingen en hun schoolprestaties betreft motivatie (Deci & Ryan, 2000). Motivatie valt in te delen in subthema's die invloed uitoefenen op de mate waarin leerlingen gemotiveerd zijn (Ryan & Deci, 2020; Skinner & Pitzer, 2012). Intrinsieke motivatie is volgens Deci en Ryan (2000) onder andere afhankelijk van de drie psychologische basisbehoeften: autonomie, competentie en verbondenheid. Deze basisbehoeften vormen de kern van de zelfdeterminatietheorie (SDT). Wanneer in deze behoeften wordt voorzien, kan dit samenhangen met hogere motivatie, wat een positief effect kan hebben op het leerproces en de bijbehorende prestaties (Deci & Ryan, 2000). Zowel de behoefte aan autonomie als de mate waarin leerlingen autonomieondersteuning ervaren vanuit hun docent spelen hierin een centrale rol. Autonomie gaat volgens Deci en Ryan (2000) om de mate waarin een leerling keuzes kan maken binnen het eigen leerproces, zeggenschap heeft over de aangeboden leeractiviteiten en in hoeverre de leerling zich eigenaar voelt van het proces. Wanneer wordt voorzien in de autonomiebehoefte, ontstaat er doorgaans een intrinsieke motivatie die de leerprestaties positief beïnvloedt. Bij competentie gaat het om het gevoel dat een leerling zichzelf voldoende in staat acht om een taak te voltooien. Hierbij gaat het zowel om het efficiënt afronden van de taak als om het hebben van de overtuiging dat mogelijke obstakels die horen bij het voltooien van een taak kunnen worden overkomen. Bij verbondenheid gaat het om de relatie die een leerling heeft met de omgeving. Het gaat hierbij vooral om vertrouwen in elkaar, een veilige klassenomgeving en het gevoel hebbende dat fouten maken mag, en erbij hoort (Ryan & Deci, 2000; Vansteenkiste et al., 2012; Verbeek, 2010). Het is in de context van deze drie behoeften belangrijk om stil te staan bij het gegeven dat de weg naar het vervullen van deze behoeften niet voor elke leerling hetzelfde is. Wat voor de ene leerling sterk motiverend werkt kan voor een andere leerling veel minder passend zijn (Vansteenkiste & Soenens, 2015).

Ook recentelijk onderzoek binnen de onderwijscontext laat zien dat autonomie belangrijk is voor de intrinsieke motivatie (Mammadov & Schroeder, 2023; Yang et al., 2022). In dit onderzoek zal de focus daarom liggen op autonomie, hierbij wordt gekeken naar de autonomiebehoefte en ervaren autonomieondersteuning van leerlingen. Autonomie wordt gezien als een belangrijke voorwaarde voor het bereiken van motivatie. Verder onderzoek naar autonomie kan helpen om nog beter in de behoeften van leerling te voorzien, zeker gezien uit eerdere studies blijkt dat leraren en leerlingen soms een ander beeld hebben van de ondersteuning die daadwerkelijk wordt gegeven (Admiraal et al., 2019). Daarom is verder onderzoek naar dit thema relevant aangezien uit diverse studies blijkt dat autonomieondersteuning samenhang vertoont met positieve uitkomsten, zoals een verhoogde betrokkenheid, zelfregulatie, positief sociaal gedrag en verhoogde leerresultaten (Reeve & Cheon, 2021; Cheon et al., 2023; Miao & Ma, 2023). Door autonomieondersteuning op een goede manier te kunnen meten kan het meer zeggen over de effectiviteit van de onderwijspraktijk. Bij het kijken naar het proces van autonomie speelt structuur ook een rol. Uit de literatuur komt namelijk naar voren dat autonomie-ondersteuning en het hebben van een duidelijke structuur complementair zijn aan elkaar (Jang et al., 2010; Vansteenkiste et al., 2012). Hierdoor is het niet alleen relevant om te kijken naar de wens van leerlingen op het gebied, maar voornamelijk in hoeverre zij ondersteuning ervaren en of deze voorzien is van voldoende structuur.

Autonomiebehoefte en -ondersteuning

Bij autonomiebehoefte gaat het om de mate waarin leerlingen keuze en eigenaarschap hebben in hun leerproces. Dit kan zich uiten door middel van het maken van keuzes, controle hebben over het eigen leerproces of het bezitten van verantwoordelijkheden. Wanneer er wordt voldaan aan deze voorwaarden voelt een leerling zich mogelijk meer betrokken bij het leerproces en kan de intrinsieke motivatie groeien (Vansteenkiste et al., 2012). Op het gebied van de ervaren autonomieondersteuning gaat het om hoe leerlingen de autonomie-ondersteuning vanuit de betrokken docenten ervaren. Wanneer de autonomiebehoefte onvoldoende wordt vervuld, zou dit kunnen leiden tot gevoelens van controleverlies, frustratie en verminderde motivatie (Vansteenkiste & Soenens, 2015). Minnaert en Odenthal (2018) benadrukken dat er niet alleen naar de leerling moet worden gekeken, maar ook naar de rol die de docent speelt. De interactie tussen leerling en docent blijkt namelijk belangrijk voor motivatie, die wordt gestimuleerd door autonomie-ondersteunend gedrag van docenten. Autonomie-ondersteuning wordt hierbij omschreven als de gedragingen en interacties van een

docent die erop zijn gericht om de autonomiebehoefte van leerlingen te vervullen. Onderzoek suggereert dat autonomieondersteuning vanuit docenten een positieve invloed kan hebben op leerprestaties en de betrokkenheid van leerlingen (Reeve et al., 2004). Anderzijds zou te weinig autonomieondersteuning kunnen leiden tot een negatieve invloed op schoolprestaties (Jang, Reeve & Deci, 2010).

Wanneer wordt gekeken naar autonomieondersteuning vanuit de docent kunnen er vier onderdelen worden onderscheiden (Reeve, 2006; Patall et al., 2008; Hornstra et al., 2016). Het eerste onderdeel bestaat uit het geven van een betekenisvolle uitleg. Dit ontstaat door zo veel mogelijk aan te sluiten bij de belevingswereld van de leerlingen en de actualiteit. Daarnaast is aandacht voor het ‘waarom’ van belang. Leerlingen krijgen hierdoor inzicht in de relevantie van de taak. Het tweede onderdeel is het gebruikmaken van informatieve taal. Er wordt hierbij onderscheid gemaakt tussen informatieve- en dwingende taal. Informatieve taal kenmerkt zich door het benadrukken van het belang en het waarom in een bepaalde taak. Dwingende taal kenmerkt zich door het ‘moeten’, hetgeen weerstand kan opleveren. Het derde onderdeel betreft het bieden van betekenisvolle opties die voorzien zijn van de juiste structuur. Het geven van keuzes in het met wie zij willen werken, wanneer zij dit willen doen, de aanpak hiervan en de volgorde waarop het werk voltooid dient te worden is daarbij van belang. Keuzemogelijkheden zouden idealiter wel aangeboden worden binnen een duidelijke structuur. Het laatste onderdeel gaat over het inlevingsvermogen. Van docenten wordt verwacht dat zij luisteren en rekening houden met hoe een leerling zich voelt (Hornstra et al., 2016). Het doel van de autonomie-ondersteunende benadering is het creëren van intrinsieke motivatie en activiteit binnen het klaslokaal, hierbij zijn zowel keuzevrijheid als het hebben van een duidelijke structuur van grote invloed (Vansteenkiste et al., 2012).

Binnen de literatuur wordt het belang van een zekere mate van autonomie-ondersteuning erkend, toch ziet men in de praktijk vaak dat autonomie-ondersteuning te weinig wordt toegepast. Docenten overschatten namelijk de mate waarin zij een leerling autonomie-ondersteuning bieden, terwijl leerlingen minder autonomie ervaren dan dat de docent ze denkt te bieden (Admiraal et al., 2019). Door deze tegenstrijdigheid is het van belang dat er meetinstrumenten zijn waarin de opvattingen van de leerling duidelijk naar voren komen.

Autonomie meten

Om de dynamiek tussen autonomieondersteuning en behoefte van autonomie beter te begrijpen, is het belangrijk om te beschikken over betrouwbare en valide meetinstrumenten

die zowel de autonomiebehoefte als de ervaren autonomieondersteuning meten. Er zijn instrumenten ontwikkeld om deze psychologische constructen te meten, maar er is nog niet voldoende overzicht van de psychometrische kwaliteit van deze instrumenten. Ook richten sommige instrumenten zich uitsluitend op autonomieondersteuning of juist alleen op behoefte, terwijl andere beide combineren. In beide gevallen ontbreekt vaak een duidelijke onderbouwing van de kwaliteit (Black & Deci, 2000).

Er bestaan veel instrumenten die autonomiebehoefte en autonomieondersteuning meten, maar de kwaliteit is niet altijd duidelijk vanwege de complexiteit van het krijgen en geven van de juiste ondersteuning. Er zijn wel meetinstrumenten ontwikkeld om deze concepten in kaart te brengen, maar er ontbreekt een systematisch overzicht van de psychometrische kwaliteit van deze instrumenten, met name in de context van leerlingen in het primair- en voortgezet onderwijs. Daarbij blijkt uit recente systematische reviews dat het aantal studies dat specifiek de ervaren autonomie-ondersteuning van leerlingen meet in het primair en voortgezet onderwijs relatief schaars is (Yang et al., 2022; Mammadov & Schroeder, 2023). In Nederland is wel onderzoek gedaan naar perceptieverschillen tussen docenten en leerlingen (Admiraal et al., 2019), maar ook hier ontbreekt een systematisch overzicht van de kwaliteit van de gebruikte instrumenten.

Hierdoor is er behoefte aan een systematisch overzicht van bestaande meetinstrumenten, met een kritische evaluatie van hun psychometrische eigenschappen. Het is waardevol om deze eigenschappen goed in beeld te krijgen. De psychometrische kwaliteit van een instrument kan worden beoordeeld aan de hand van de internationaal erkende COSMIN-richtlijnen (Mokkink et al., 2024). Deze maken onderscheid in betrouwbaarheid, validiteit en responsiviteit. Bij betrouwbaarheid gaat het hier om de mate waarin een instrument consistente resultaten geeft over tijd of tussen items. In het geval van validiteit om de mate waarin een instrument meet wat het beoogt te meten, en bij de responsiviteit gaat het om het vermogen van een instrument om relevante veranderingen in gemeten uitkomsten over de tijd heen te herkennen.

Wanneer een vragenlijst van lage of onbekende kwaliteit is op psychometrische gebied, kan dit leiden tot onbetrouwbare gegevens en daardoor kunnen er vervolgens onjuiste conclusies worden getrokken over de autonomiebehoefte of ervaren ondersteuning van leerlingen. Dit kan zowel voor de onderzoekspraktijk, als voor de onderwijspraktijk ineffectief zijn. Het kan bijvoorbeeld voorkomen dat leraren vervolgens interventies inzetten die niet aansluiten bij de werkelijke behoeften van leerlingen (Ntumi & Antwi-Agyakwa, 2022). Om hier een kanttekening bij te plaatsen neemt bovenstaande niet weg dat het trekken

van onjuiste conclusies altijd kan voorkomen, zowel bij de aanwezigheid van een hoge als een lage psychometrische kwaliteit.

Wat betreft de psychometrische eigenschappen zijn er geen vaste, algemene kenmerken van een “goede” vragenlijst. De context is van grote invloed. De betrouwbaarheid en validiteit van een instrument kan dan ook verschillen per doelgroep, het doel van de meting en de situatie waarin het wordt gebruikt. Een vragenlijst die betrouwbaar blijkt bij adolescenten in regulier onderwijs, is dus niet automatisch even geschikt voor jongere kinderen of voor gebruik in het speciaal onderwijs (COTAN, z.d.).

In dit onderzoek wordt een systematische review uitgevoerd, met specifieke aandacht voor de psychometrische kwaliteit van instrumenten die gericht zijn op het meten van autonomiebehoefte en/of ervaren autonomie-ondersteuning van leerlingen. De psychometrische kwaliteit van de instrumenten zal worden bekeken aan de hand van de COSMIN en het zesstappenmodel van Dima (2018). Bij de totstandkoming van de review zullen de PRISMA richtlijnen worden gevolgd. Op deze manier kan dit onderzoek naast een wetenschappelijke bijdrage mogelijk ook praktische handvatten bieden voor docenten, schoolleiders en onderzoekers. Bij het onderzoek zijn op basis van het theoretisch kader de volgende onderzoeksvragen ontwikkeld.

Centrale onderzoeksvraag (hoofdvraag)

Welke meetinstrumenten zijn er beschikbaar voor het meten van autonomiebehoefte en ervaren autonomieondersteuning bij leerlingen in het primair en voortgezet onderwijs, en wat is de psychometrische kwaliteit van deze instrumenten?

Deelvragen

1. Welke meetinstrumenten zijn er ontwikkeld voor het meten van autonomiebehoefte en/of ervaren autonomieondersteuning bij leerlingen en hoe worden deze concepten gedefinieerd en gemeten?
2. Welke psychometrische eigenschappen worden gerapporteerd over deze meetinstrumenten en wat zijn de uitkomsten ervan?
3. Binnen welke onderwijscontexten zijn deze meetinstrumenten gebruikt en wat is de bijbehorende methodologische kwaliteit van deze studies waarin de psychometrische eigenschappen van deze instrumenten worden gerapporteerd?

Methode

Onderzoeksopzet

Dit onderzoek zal een systematische review zijn naar meetinstrumenten die gericht zijn op autonomiebehoefte en de ervaren autonomie-ondersteuning van leerlingen in het primair en voortgezet onderwijs. De focus op primair en voortgezet onderwijs is gekozen om enerzijds zo breed mogelijk te kunnen inventariseren welke instrumenten worden gebruikt, en anderzijds een duidelijke afbakening te hanteren in leeftijd en ontwikkelingsfase. Doordat deze groepen relatief vergelijkbaar zijn is de kans groter dat de gevonden resultaten onderling beter vergelijkbaar zijn.

De opzet volgt het PRISMA-protocol (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*; PRISMA, 2024) voor het transparant en systematisch opzetten, uitvoeren en rapporteren van systematische reviews.

Voor het beoordelen van de psychometrische eigenschappen van de instrumenten wordt het COSMIN (*COnsensus-based Standards for the selection of health Measurement INstruments*) gebruikt (Mokkink et al., 2024). COSMIN vindt zijn oorsprong in de medische wetenschap, maar het is een middel dat ook in de context van het onderwijs kan worden gebruikt. Binnen zowel psychologisch als onderwijskundig onderzoek kan COSMIN gebruikt worden voor het selectieproces en ook voor het evaluatieproces van meetinstrumenten. Bovendien is de checklist te gebruiken bij het vellen van een oordeel over de methodologische kwaliteit van psychometrische studies (Mokkink et al., 2010; COSMIN, 2024).

Naast bovengenoemde middelen zal er ook gebruik worden gemaakt van het zesstappenplan van Dima (2018). Het plan is ontwikkeld voor de beoordeling van schaalvalidatie. Dima (2018) is waardevol binnen dit onderzoek aangezien het de onderbouwing versterkt, het kan worden gezien als een aanvulling op COSMIN. Binnen de COSMIN-criteria kijken we naar de kwaliteit van de studie en met behulp van Dima kunnen we ook kijken naar de mate waarin het validatieproces in orde is.

Wat betreft het zoeken zal er een ondergrens worden gebruikt vanaf 2016. Verouderde instrumenten sluiten vaak niet meer aan bij de actuele context, daarom is gekozen om recentere instrumenten te zoeken die beter aansluiten bij de recente theorie en actualiteit.

Selectiecriteria

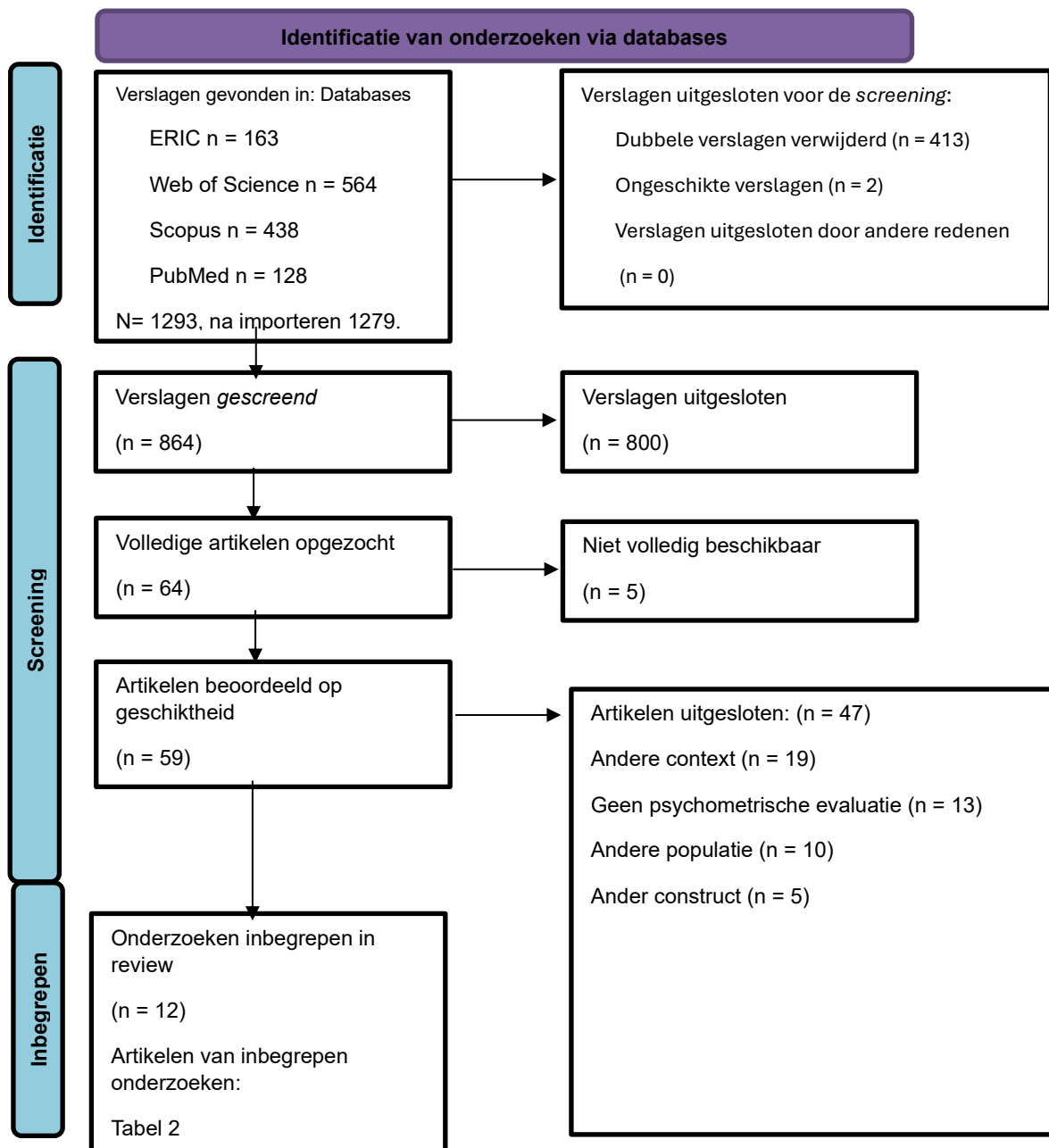
Voor het onderzoek zijn de volgende inclusiecriteria omschreven, deze zijn te lezen in Tabel 1.

Tabel 1
Inclusiecriteria

Categorie	Inclusiecriteria
Populatie	Onderzoek gericht op leerlingen in het primair of voortgezet onderwijs.
Context	Studies uitgevoerd binnen een onderwijscontext.
Concept	Autonomie of autonomie-ondersteuning, en theoretisch onderbouwd vanuit de SDT.
Psychometrische kwaliteit	Rapportage van psychometrische gegevens zoals betrouwbaarheid, validiteit en/of responsiviteit van het instrument.
Publicatietype	Peer-reviewed
Toegankelijkheid	Full-tekst is beschikbaar.
Focus van het instrument	Instrument meet autonomie of autonomie-ondersteuning of een nauw verwant construct.

Figuur 1

Zoekstrategie volgens de PRISMA-richtlijnen weergegeven in een flowdiagram.



Noot: Aangepast naar Page et al. (2021).

Selectieprocedure

In dit onderzoek is gezocht in vier wetenschappelijke databases: ERIC, Web of Science, Scopus en PubMed. De zoekstrategie was gericht op instrumenten die autonomiebehoefte of autonomieondersteuning meten bij leerlingen in het primair en voortgezet onderwijs, en psychometrische kenmerken rapporteren. In totaal werden 1293 studies geïdentificeerd (ERIC n = 163, Web of Science n = 564, Scopus n = 438, PubMed n = 128).

Eerst zijn alle gegevens opgenomen in EndNote, vervolgens zijn alle duplicaten verwijderd. Na het importeren in EndNote en het verwijderen van duplicaten werden er 413 artikelen verwijderd, waardoor 864 hits overbleven voor titel- en abstract screening. Daarna zijn handmatig alle titels en samenvattingen bekeken en zijn nog enkele duplicaten verwijderd, omdat kleine verschillen in de titel er al voor zorgden dat EndNote deze artikelen als een aparte versie zag. Tijdens de screening op titel en abstract werden uiteindelijk 800 studies uitgesloten.

Na de screeningsfase werden de artikelen volledig gelezen op geschiktheid. Voor 59 artikelen werd de volledige tekst opgevraagd. Uiteindelijk werden 47 artikelen uitgesloten. De reden van uitsluiting was de aanwezigheid van een andere context ($n = 19$), geen psychometrische evaluatie ($n = 13$), andere populatie ($n = 10$), ander construct ($n = 5$). In totaal zijn 12 onderzoeken opgenomen in de systematische review. De gehele documentatie van dit proces is te zien in Figuur 1 en de exacte zoekstrategie is terug te vinden in Bijlage A. De screening van de uiteindelijk in het onderzoek opgenomen artikelen werd ook gedaan door een medestudent.

De studies en de bijbehorende instrumenten zijn weergegeven in Tabel 2, de kenmerken van de instrumenten in Tabel 3. In Tabel 4 staan de gerapporteerde psychometrische eigenschappen, in Tabel 5 de COSMIN-beoordelingen en in Tabel 6 het overzicht volgens de Dima-stappenanalyse (2018).

Data-extractie

Voor elk meetinstrument dat in het onderzoek is vermeld werden naam, doel, doelgroep, context, aantal items, aantal schalen/dimensies, de gebruikte analyses (betrouwbaarheid, validiteit), de gerapporteerde psychometrische eigenschappen en de onderliggende theorie of conceptualisering van autonomie gerapporteerd.

Beoordelingskader

De beoordeling van de psychometrische kwaliteit is gebeurd op basis van het COSMIN (COSMIN, z.d.). Hierbij is gekeken naar aspecten als betrouwbaarheid, validiteit en responsiviteit waarbij deze zijn geëvalueerd aan de hand van de bijbehorende checklists (Mokkink et al., 2024; Prinsen et al., 2018). Daarnaast is er aanvullend gebruik gemaakt van het stappenplan van Dima (2018) voor de beoordeling van schaalvalidatie. Dima (2018) bestaat uit de volgende zes stappen:

1. Beschrijvende itemanalyse (distributie van items, floor- en ceilingeffecten, missings),
2. Non-parametrische schaalanalyse, mokkenanalyse.

3. Parametrische IRT-/Rasch-analyse,
4. Factoranalyse (EFA/CFA/ESEM en toetsing van meetinvariantie),
5. Betrouwbaarheid (Cronbach's α , McDonald's ω),
6. Evaluatie van totaalscores en externe validiteit (correlaties met andere constructen).

Voor elk instrument is gekeken welke van deze zes stappen in het artikel zijn uitgevoerd en gerapporteerd. Deze informatie staat in Tabel 6 waarbij per instrument per stap is aangegeven of de betreffende psychometrische analyse is uitgevoerd (\checkmark), deels is uitgevoerd (\pm) of volledig ontbreekt ($-$). De Dima-analyse wordt gebruikt als aanvulling op COSMIN. COSMIN kijkt of een vragenlijst in zijn geheel werkzaam is, Dima geeft meer inzicht in de diepgang van de schalen en items. Dit maakt het mogelijk om niet alleen te beoordelen of een instrument betrouwbare en valide uitkomsten rapporteert, maar maakt het ook mogelijk om nuances aan te brengen op een dieper niveau.

Analyse en synthese

De verkregen data zijn geëvalueerd en samengevat in tabellen. De nadruk ligt op het vergelijken van de eigenschappen van verschillende instrumenten. Er is per meetinstrument gerapporteerd welke psychometrische eigenschappen zijn geëvalueerd, met aandacht voor kwaliteitsscores en contextuele factoren (zoals doelgroep en onderwijsniveau). Het doel is om een overzicht te bieden welke meetinstrumenten het meest geschikt zijn en inzicht te krijgen in welke mate deze instrumenten psychometrisch onderbouwd zijn.

Ethische overwegingen

Er is gewerkt volgens de principes van wetenschappelijke integriteit. Transparantie, herhaalbaarheid, objectiviteit en correcte bronvermelding werden continu nageleefd. Aangezien er geen persoonsgegevens werden verzameld, was er geen aanvullende ethische goedkeuring vereist.

Resultaten

Beschrijving van de opgenomen artikelen

De twaalf geïnccludeerde studies zijn uitgevoerd tussen 2016 en 2025 en vonden plaats in België, Spanje, Zuid-Korea, Portugal, Turkije, China, Nederland, en Duitsland. De onderzoeken werden uitgevoerd binnen het primair- en voortgezet onderwijs. De steekproefgroottes varieerden afhankelijk van het instrument en het onderwijsniveau tussen enkele honderden tot meer dan duizend.

Alle studies hadden als doel het ontwikkelen, vertalen of valideren van een vragenlijst gericht op het meten van autonomiebehoefte of autonomieondersteuning binnen de context van de zelfdeterminatietheorie. De onderzochte instrumenten richten zich op verschillende aspecten van motivatie, zoals ervaren autonomieondersteuning, docentgedrag, (de)motiverende stijlen en behoeftebevrediging of -frustratie. Het overzicht van deze artikelen is opgenomen in onderstaande Tabel 2.

Tabel 2*Overzicht van de opgenomen artikelen en bijbehorende kenmerken*

Artikel	Land	Onderzoeksdoel	Instrument	Doelgroep	Steekproef (N)	Meetdoel
Aelterman et al. (2019)	België	Ontwikkelen en valideren docentstijlen binnen SDT-circumplex.	(SIS) Situations-in-School Questionnaire	VO	1332 leraren, 1735 leerlingen (14,6 jaar)	Autonomieondersteuning, structuur, controle, chaos
Ahn et al. (2019)	Zuid-Korea	Factorstructuur TASCQ bevestigen.	(TASCQ) Teacher as Social Context Questionnaire	PO	697 leerlingen/35 klassen (10-12 jaar).	Ervaren autonomieondersteuning, structuur en betrokkenheid.
Burgueño et al. (2024)	Spanje	Adaptatie/validatie SIS-PE.	(SIS-PE) Situations-in-School PE	VO (LO)	1441 leerlingen (15,94 jaar), 473 docenten, 654 studenten	Autonomieondersteuning, structuur, controle, chaos

Burgueño et al. (2020)	Spanje	Adaptatie en psychometrische evaluatie MD-PASS-PE.	(MD-PASS-PE) Multidimensional Perceived Autonomy Support Scale (PE)	VO (LO)	560 leerlingen (14,49 jaar).	Cognitieve, procedurele en organisatorische autonomieondersteuning.
Burgueño & Medina-Casabón (2021)	Spanje	Valideren IBQ-PE.	(IBQ-PE) Interpersonal Behaviors Questionnaire (PE)	VO (LO)	478 leerlingen (14,56 jaar).	Autonomie, competentie en relatieondersteuning, frustratie
Liu & Chung (2017)	Hong Kong	Factorstructuur + meetinvariantie NSTSS-PE.	(NSTSS-PE) Need-Supportive Teaching Style Scale (PE)	VO	605 leerlingen (13,24 jaar).	Autonomieondersteuning, structuur, betrokkenheid.
Conesa & Duñabeitia (2021)	Spanje	Ontwikkelen/valideren schaal behoeftebevrediging klascontext.	(BPN-CS) Basic Psychological Needs Satisfaction in the Classroom	PO	1344 leerlingen (10,3 jaar).	Autonomie, competentie, verbondenheid, nieuwigheidsbehoefte.

Koçak et al. (2025)	Turkije	Validatie Turkse BPNSFS + longitudinale validiteit.	(BPNSFS) Basic Psychological Needs Satisfaction and Frustration Scale	VO	1033 (T1) + 849 (T2) (15,91 jaar).	Autonomie, competentie en relatiebehoefte
Li et al. (2025)	China	Ontwikkelen schaal behoeftefrustratie op school.	(SBPNFSS) Students' Basic Psychological Needs Frustration at School Scale	PO, VO	1186 + 1614 + 1441 leerlingen (7-19 jaar).	Frustratie van autonomie, competentie en relatiebehoefte.
Van Aart et al. (2017)	Nederland	Valideren autonomie-behoefteschaal voor 9–12-jarigen.	(CARR) Competence, Autonomy, Relatedness Relations Scale	PO	172 leerlingen (9-12 jaar).	Autonomiebehoefte, competentiebehoefte, relatie met docent/klasgenoten.
Inman et al. (2023)	Portugal	Valideren Portugese ASBPNSS	(ASBPNSS) Basic Psychological Needs at School Scale	VO	1648 leerlingen (14,1 jaar).	Autonomie, competentie, relatiebehoefte.

Sturm et al. (2020)	Duitsland	Adaptatie Duitse BPNS-PE + multilevel validatie.	(GBPNS-PE) German BPNS in PE Scale	VO (LO)	481 leerlingen (9-10 jaar).	Autonomie, competentie, relatiebehoefte tijdens LO
---------------------	-----------	--	--	---------	--------------------------------	---

Noot. Weergave van de beschrijvende kenmerken.

Een deel van de studies richtte zich op de ontwikkeling van nieuwe meetinstrumenten, zoals de CARR en SBPNFSS. Andere onderzoeken hadden als doel de factorstructuur en validiteit van bestaande schalen te bevestigen of te verbeteren, zoals bij de TASCQ, NSTSS-PE en ASBPNSS. Daarnaast richtten diverse studies zich op culturele adaptatie en validatie van instrumenten in een nieuwe taal of onderwijssysteem, bijvoorbeeld de Turkse, Portugese en Duitse versies van de BPNSFS. Over het algemeen lag de focus op het waarborgen van betrouwbare en valide metingen van autonomie of autonomie gerelateerde constructen.

Binnen de twaalf studies werden verschillende vragenlijsten gebruikt die allemaal zijn ontwikkeld vanuit de zelfdeterminatietheorie. De instrumenten verschilden in structuur en focus, maar waren gericht op het meten van autonomie, autonomieondersteuning of basisbehoeften. Sommige schalen richtten zich op docentstijlen en de manier waarop leerlingen het gedrag van hun docent ervaren, zoals de SIS, TASCQ en MD-PASS-PE. Andere instrumenten waren gericht op de ervaren vervulling of frustratie van basisbehoeften van leerlingen zelf, waaronder de BPNSFS, SBPNFSS en ASBPNSS. Daarnaast waren er instrumenten die een combinatie van autonomie, competentie en verbondenheid meten, zoals de CARR en GBPNS-PE.

De doelgroep van de studies bestond voornamelijk uit leerlingen in het primair en voortgezet onderwijs. De leeftijd van de deelnemers lag tussen de 7 en 19 jaar. In sommige gevallen werden ook docenten of studenten in opleiding betrokken, vooral bij de ontwikkeling of validatie van vragenlijsten die docentgedrag betreffen. Wel is het zo dat het perspectief van de leerling in alle studies centraal staat.

Wat betreft het meetdoel richtten de instrumenten zich op relevante aspecten van autonomie binnen de onderwijscontext. Een deel van de vragenlijsten meet de ervaren autonomieondersteuning door de docent, inclusief gerelateerde dimensies zoals structuur, controle, betrokkenheid en chaotisch gedrag. Andere instrumenten zijn gericht op de ervaren vervulling van de psychologische basisbehoeften autonomie, competentie en verbondenheid. Daarnaast waren er vragenlijsten die expliciet de frustratie van deze basisbehoeften meten. Aanvullend is er nog gekeken naar de specifieke kenmerken van de vragenlijsten die in het onderzoek zijn opgenomen. Deze kenmerken zijn weergegeven in Tabel 3.

Tabel 3*Specifieke kenmerken van de instrumenten*

Instrument	Aantal items/ schalen	Type schaal	Herkomst (origineel/ aangepast)
1. SIS – Situations-in-School Questionnaire	60 items/ 4 stijlen	Likert (1-7)	Origineel ontwikkeld
2. TASCQ – Teacher as Social Context Questionnaire (short form)	24 items/ 3 schalen	Likert (1-5)	Verkorte TASCQ-versie
3. SIS-PE – Situations-in-School (Physical Education)	48 items/ 4 stijlen	Likert (1-7)	Aangepast en vertaald van SIS-PE-vragenlijst
4. MD-PASS-PE – Multidimensional Perceived Autonomy Support Scale (PE)	15 items/ 3 factoren	Likert (1–7)	Aangepast en vertaald van PASS
5. IBQ-PE – Interpersonal Behaviors Questionnaire (PE)	24 items/ 6 schalen	Likert (1–7)	Aangepast en vertaald van IBQ naar onderwijs/LO
6. NSTSS-PE – Need-Supportive Teaching Style Scale	10 items /3 schalen	Likert (1–7)	Origineel ontwikkeld
7. BPN-CS – Basic Psychological Needs Satisfaction in the Classroom	17 items/ 4 schalen	Likert (1–5)	Origineel ontwikkeld
8. BPNSFS – Basic Psychological Needs Satisfaction and	24 items/ 6 schalen	Likert (1–5)	Aangepast en vertaald naar schoolcontext

Frustration Scale
(school)

9. SBPNFSS – Students’ Basic Psychological Needs Frustration at School Scale	15 items/ 3 schalen	Likert (1–7)	Origineel ontwikkeld
10. CARR – Competence, Autonomy, Relatedness Relations Scale	18 items/ 4 schalen	Likert (1–5)	Origineel ontwikkeld
11. ASBPNSS – Adolescent Students’ Basic Psychological Needs at School Scale	15 items/ 3 factoren	Likert (1–5)	Aangepast en culturele validatie
12. GBPNS-PE – German Basic Psychological Needs Satisfaction in PE	11 items/ 3 schalen	Likert (1–5)	Aangepast en vertaald van BPNS naar LO (Duits)

Alle instrumenten bestaan uit meerdere items met Likertschaal met tussen de vijf en zeven antwoordmogelijkheden. Het aantal items varieert van 10 tot 60 items, afhankelijk van het aantal factoren of aantal dimensies dat wordt gemeten. De meeste instrumenten zijn oorspronkelijk in het Engels of Spaans ontwikkeld en vervolgens vertaald en/of geadapteerd naar andere talen (waaronder Nederlands, Turks, Duits en Chinees) volgens internationale richtlijnen voor instrumentadaptatie. Hierdoor konden instrumenten aangepast worden aan de culturele context.

Psychometrische kwaliteit

De rapportage van de psychometrische eigenschappen van de instrumenten is weergegeven in Tabel 4.

Tabel 4*Rapportage van psychometrische eigenschappen van de instrumenten volgens COSMIN*

Instrument	Interne Consistentie	Betrouwbaarheid (test-hertest)	Inhoudsvaliditeit	Constructvaliditeit	Criteriumvaliditeit
1. SIS – Situations-in-School (Aelterman et al., 2019)	α tussen 0.69 en 0.90	$r = .48-.80$	Expert, pilot, SDT-onderbouwing	Circumplex structuur bevestigd (MDS + CFA)	Significante verbanden met motivatie en welzijn
2. TASCQ – Teacher as Social Context Questionnaire (short) (Ahn et al., 2018)	Betrouwbaarheid = .91 (student) & .92 (teacher) (multilevel MCFA)	-	Vertaling richting het Koreaans	Structuur bevestigd (MEFA/MCFA)	-
3. SIS-PE – Situations-in-School PE versie (Burgueño et al., 2024)	ω .71–.89 (stijlen)	-	ITC-vertaling en culturele adaptatie	Bifactor-ESEM en circumplex structuur bevestigd	Significante verbanden met basisbehoeften
4. MD-PASS-PE – Multidimensional Perceived Autonomy	α geschikt, AVE-waarden boven .50	ICC = .82–.87	Expertpanel	CFA en bifactor-ESEM bevestigd	Predictieve validiteit met basisbehoeften

Support Scale (PE)
(Burgueño et al., 2020)

5. IBQ-PE – Interpersonal Behaviors Questionnaire (PE) (Burgueño & Medina-Casabón, 2020)	$\alpha = .75-.81$, AVE $\geq .50$ voor alle 6 factoren	ICC = $\geq .70$	CVI = 0.8	CFA, invariantie voor geslacht en leeftijd	Significante relaties met behoeftebevrediging en -frustratie
6. NSTSSPE – Need-Supportive Teaching Style Scale in PE (Liu & Chung, 2017)	$\alpha = .76-.92$ (10-item model)	Follow-up na 2 maanden, stabiel over tijd	Vertaling, pilot, items geselecteerd uit TASCQ en HCCQ	ESEM (betere fit dan CFA)	Significante relaties met autonomie-, competentie- en relatiebehoefte
7. BPN-CS – Basic Psychological Needs Satisfaction in the Classroom Scale (Conesa & Duñabeitia, 2021)	$\omega .72-.79$	-	Adaptatie en toegevoegd <i>novelty</i>	CFA 4-factor model bevestigd, meetinvariantie (geslacht)	Convergente validiteit
8. BPNSFS – Basic Psychological Needs Satisfaction and	$\omega .78-.88$	Geen test-hertest, wel longitudinale invariantie	ITC-procedure en expert review	CFA en MTMM en invariantie over tijd	Predictieve validiteit (welzijn/ depressieve gevoelens)

Frustration Scale (Koçak et al., 2025)

<p>9. SBPNFSS – Students’ Basic Psychological Needs Frustration at School Scale (Li et al., 2025)</p>	<p>Totaal $\alpha=.93$, $\omega=.94$. $r = .69-.75$ Subschalen $\alpha/\omega=.87-.90$. AutF: $\alpha=.901$, $\omega=.901$</p>	<p>Itemontwikkeling, expertpanel, pilot</p>	<p>EFA, EGA, CFA, bevestigd, invariantie (geslacht, SES, tijd)</p>	<p>Predictieve validiteit (betrokkenheid)</p>
<p>10. CARR Scale & BRPEQ-adaptatie (van Aart et al., 2017)</p>	<p>Autonomie ($\rho=0.67$) Kappa .21 en .60 Kappa .08 en .69</p>	<p>Expert, pilot, adaptatie</p>	<p>Mokken: autonomie zwakker</p>	<p>Convergente validiteit (autonome motivatie)</p>
<p>11. ASBPNSS – Adolescent Students’ Basic Psychological Needs at School Scale (Portuguese version) (Inman et al., 2023)</p>	<p>$\alpha/\omega. >.70$ -</p>	<p>ITC-vertaling, terugvertaling</p>	<p>CFA bevestigd en meetinvariantie (geslacht, SES)</p>	<p>Convergente validiteit (welzijn)</p>

12. GBPNS-PE – German Basic Psychological Needs Satisfaction in Physical Education Scale (Sturm et al., 2020)	$\alpha = .78-.85$	-	Vertaling, adaptatie, pilot	MCFA bevestigd	Significante relaties met MVPA (PE-gedrag)
--	--------------------	---	--------------------------------	----------------	---

Noot. De data vanuit de artikelen vormt de input van de rapportage. Een streepje (–) geeft aan dat het onderdeel niet is gerapporteerd. Zowel het onderdeel responsiviteit als meetfout is vanwege de leesbaarheid niet in de tabel opgenomen. Over het onderdeel responsiviteit is in de artikelen geen relevante informatie gerapporteerd wat enkel zou resulteren in de opname van (–). Wat betreft het onderdeel meetfout is er sprake van een discrepantie tussen wat COSMIN verstaat onder het rapporteren van de meetfout en op welke wijze informatie over de meetfout door de geïnccludeerde onderzoeken wordt gepresenteerd. Hier wordt aandacht aan besteed in de discussie. Afkortingen: *AutS* = *autonomie satisfactie*; *AutF* = *autonomie frustratie*; *S* = *satisfactie*; *F* = *frustratie*; *CFA* = *confirmatieve factoranalyse*; *MCFA* = *Mixture of Common Factor Analyzers*; *ESEM* = *exploratory structural equation modeling*; *MEFA* = *Multilevel Exploratory Factor Analysis*; *EFA* = *Exploratory Factor Analysis*; *EGA* = *exploratory graph analysis*; *MDS* = *Multidimensional Scaling*; *ICC* = *intraclasse correlatiecoëfficiënt*; *CFI* = *comparative fit index*; *r* = *correlatiecoëfficiënt*; *MVPA* = *Matige-tot-Intensieve Fysieke Activiteit*; *ITC* = *International Test Commission*; *CVI* = *Content Validity Index*; *MMTM* = *Multitrait-Multimethod Matrix*.

De interne consistentie van de instrumenten was voldoende tot zeer goed. De laagste waarden werden gerapporteerd bij de SIS, waar Cronbach's α varieerde tussen 0.69 en 0.90 afhankelijk van de stijl, en bij de BPN-CS, met α -waarden tussen .72 en .78 en McDonald's ω tussen .72 en .79. De hoogste interne consistentie was bij de SBPNFSS, met α/ω -waarden rond .93, wat duidt op een hoge betrouwbaarheid van de schaal. Andere instrumenten lieten goede betrouwbaarheid zien, zoals de TASCQ (α .91–.92), MD-PASS-PE (AVE > .50) en BPNSFS (ω .81–.88). De CARR-schaal liet een zwakker beeld zien, met een relatief lage betrouwbaarheid voor de autonomie-subschaal ($\rho = .67$). Over het algemeen zijn adequate waarden volgens COSMIN boven de .70.

De stabiliteit over tijd werd in slechts een beperkt aantal studies onderzocht. De MD-PASS-PE toonde goede test-hertestbetrouwbaarheid met ICC-waarden tussen .82 en .87. Ook de SBPNFSS rapporteerde een redelijke tot goede stabiliteit over een periode van één maand ($r = .69$ –.75). De NSTSSPE liet een stabiel patroon zien bij een follow-up na twee maanden. Bij de overige instrumenten werd geen klassieke test-hertestbetrouwbaarheid gerapporteerd. Voor de BPNSFS werd wel longitudinale meetinvariantie onderzocht, maar dit zegt alleen iets over de stabiliteit van de meetstructuur en niet over de stabiliteit van scores. Over het algemeen is dit onderdeel onvoldoende gerapporteerd volgens COSMIN.

De meeste instrumenten rapporteerden expliciete stappen gericht op het waarborgen van inhoudsvaliditeit. Hierbij werden onder andere expertpanels, pilotstudies en vertaal- en adaptatieprocedures ingezet. De SIS, MD-PASS-PE en SBPNFSS maakten gebruik van expertbeoordeling en pilottests, terwijl bij instrumenten zoals de TASCQ en BPNSFS een gestructureerde vertaalprocedure volgens ITC-richtlijnen werd toegepast. De BPN-CS en CARR werden aangepast aan een specifieke doelgroep, waarbij bij de BPN-CS ook een aanvullende dimensie (*novelty*) werd toegevoegd. Over het algemeen zijn deze procedures in overeenstemming met de COSMIN-richtlijnen.

Voor alle instrumenten werd constructvaliditeit onderzocht aan de hand van factoranalyse. De SIS en SIS-PE bevestigden hun circumplex-structuur met behulp van MDS, CFA en bifactor-ESEM analyses. De TASCQ, NSTSSPE, BPN-CS, ASBPNS en GBPNS-PE toonden een bevestigde factorstructuur door middel van CFA of *multigroup* CFA-analyses. Complexere modellen, zoals MTMM en EGA, werden toegepast bij de BPNSFS en SBPNFSS. Binnen COSMIN wordt constructvaliditeit bij de meeste instrumenten als voldoende tot goed onderbouwd beschouwd.

Daarnaast werd in verschillende studies meetinvariantie aangetoond voor geslacht, sociaaleconomische status of tijd, wat wijst op een stabiele factorstructuur. Verschillende

instrumenten toonden significante verbanden met theoretisch verwante constructen. Zo correleerden de SIS, NSTSSPE, IBQ-PE en SBPNFSS significant met motivatie, welzijn, betrokkenheid en basisbehoeften. De BPNSFS vertoonde voorspellende relaties met welzijn en depressieve klachten, terwijl de GBPNS-PE relaties liet zien met fysiek gedrag (MVPA). Bij enkele instrumenten werd alleen convergente validiteit gerapporteerd (TASCQ, ASBPNSS), zonder verdere toetsing aan externe criteria. Over het algemeen zien we hier adequate beoordelingen terug.

Er werden geen gegevens gerapporteerd over het onderdeel responsiviteit, deze zijn dan ook niet weergegeven, maar moeten wel meegenomen worden in de beoordeling volgens COSMIN. Op dit onderdeel scoren de instrumenten niet gerapporteerd.

Naast het onderdeel responsiviteit moet er ook gekeken worden naar het onderdeel meetfout. Als we strikt kijken naar de COSMIN-richtlijnen en de input vanuit de artikelen valt er te zien dat er vanuit de artikelen geen globale gegevens worden gerapporteerd over meetfout zoals een SEM (*Standard Error of Measurement*), SDC (*Smallest Detectable Change*) of LoA (*Limits of Agreement*), dit leidt tot een beoordeling van niet gerapporteerd volgens COSMIN. Het is echter wel degelijk zo dat de onderzoeken meetfout schatten door de keuze van de gebruikte modellen en structuren. In de discussie zal hier uitgebreid op worden ingegaan. De totaalscores zijn weergegeven in onderstaande Tabel 5.

Tabel 5

Methodologische kwaliteit volgens de COSMIN Risk of Bias checklist

Instrument	Inhouds- validiteit	Structurele- validiteit	Interne consistentie (Test- hertest)	Betrouwbaarheid (Test-hertest)	Hypothesetoetsing/ Constructvaliditeit	Cross- culturele validiteit	Eindbeoordeling
1. SIS – Aelterman et al., 2019	Goed	Goed	Adequaat	Adequaat	Goed	-	Goed
2. TASCQ- short – Ahn et al., 2018	Adequaat	Goed	Goed	-	Adequaat	-	Goed

3. SIS-PE – Burgueño et al., 2024	Goed	Goed	Adequaat	-	Goed	-	Goed
4. MD- PASS-PE – Burgueño et al., 2020	Zeer goed	Goed	Zeer goed	Goed	Goed	Goed	Zeer goed
5. IBQ-PE – Burgueño & Medina- Casabón, 2020	Zeer goed	Goed	Zeer goed	Goed	Goed	Zeer goed	Zeer goed
6. NSTSSPE – Liu & Chung, 2017	Goed	Goed	Adequaat	Zeer goed	Goed	Goed	Goed
7. BPN-CS – Conesa & Duñabeitia, 2021	Adequaat	Goed	Adequaat	-	Adequaat	Goed	Adequaat
8. BPNSFS – Koçak et al., 2025	Goed	Goed	Goed	Goed	Goed	Adequaat	Goed
9. SBPNFSS – Li et al., 2025	Goed	Zeer goed	Zeer goed	Zeer goed	Goed	Goed	Zeer goed

10. CARR Scale – van Aart et al., 2017	Adequaar	Goed	Goed	Adequaar	Adequaar	-	Adequaar
11. ASBPNSS – Inman et al., 2023	Goed	Goed	Goed	-	Goed	Goed	Goed
12. GBPNS-PE – Sturm et al., 2020	Adequaar	Goed	Adequaar	Onvoldoende	Adequaar	Adequaar	Goed

Noot. De weergegeven artikelen zijn beoordeeld volgens de COSMIN-richtlijnen via de COSMIN Risk of Bias checklist (Mokkink et al., 2024). De beoordelingen zijn gebaseerd op wat er in de artikelen is gerapporteerd. Een streepje (–) geeft aan dat het onderdeel niet is gerapporteerd. Zowel het onderdeel meetfout als responsiviteit is vanwege de leesbaarheid niet in de tabel opgenomen. Over het onderdeel responsiviteit is niks gerapporteerd wat enkel zou resulteren in de opname van (–). Bij het onderdeel meetfout is er sprake van een discrepantie tussen hetgeen COSMIN verstaat onder meetfout en wat er in de artikelen is weergegeven. Hier wordt aandacht aan besteed in de discussie.

Naast de COSMIN-beoordeling is er aanvullend gebruik gemaakt van het zesstappenmodel van Dima (2018). Dit overzicht is weergegeven in Tabel 6.

Tabel 6*Stappenanalyse Dima*

Instrument	(1) Itemanalyse	2) Mokken/ NIRT	3) Parametrische IRT/Rasch	4) Factoranalyse	5) Betrouwbaarheid ($\alpha/\omega/Cr$)	6) Totalscore & externe validiteit
SIS – (Aelterman et al., 2019)	\pm , M en SD werden gerapporteerd voor de subschalen. Missing data en item-totaal correlaties ontbreken.	-	-	✓, MDS + CFA	✓, α en Test-hertest betrouwbaarheid zijn gerapporteerd.	✓
TASCQ-short (Ahn, 2018)	\pm , M en SD werden gerapporteerd. Rapportage van scheefheid en missing data-analyse ontbreken.	-	-	✓, EFA + MCFA	✓, Betrouwbaarheidscijfers zijn gerapporteerd.	-

SIS-PE (Burgueño, 2024)	✓	-	-	✓, MDS + ESEM	±, Interne consistentie (ω) is gerapporteerd, maar test-hertest is niet gerapporteerd.	✓
MD-PASS-PE (Burgueño, 2020)	✓	-	-	✓, CFA + MGCFA	✓, Interne consistentie (α , ρ) en AVE zijn gerapporteerd. Test- hertest is bevestigd.	✓
IBQ-PE (Burgueño & Medina- Casaubón, 2020)	±, Missing data en item-totaal correlaties ontbreken.	-	-	✓, CFA	✓, Interne consistentie (α) en AVE gerapporteerd. Test- hertest is bevestigd.	✓
NSTSSPE (Liu & Chung, 2017)	±, M en SD zijn gerapporteerd. Geen missing data- analyse.	-	-	✓, ESEM	✓, Interne consistentie (α) is gerapporteerd.	✓

BPN-CS (Conesa, 2021)	✓	-	-	✓, CFA	✓, Interne consistentie (α en ω) is gerapporteerd.	✓
BPNSFS (Koçak, 2025)	✓	-	-	✓, CFA	✓, Interne consistentie (ω) is gerapporteerd.	✓
SBPNFSS (Li, 2025)	±, Geen MDA	-	-	✓, EFA + EGA + CFA	✓, (CR) en AVE gerapporteerd.	✓
CARR Scale (van Aart, 2017)	±, M en SD zijn aanwezig, geen MDA	✓, Mokkenanalyse	-	✓, Mokkenanalyse	✓, Interne consistentie (ρ) en Test-hertest (Kappa's) gerapporteerd.	✓

ASBPNSS (Inman, 2023)	✓	-	-	✓, CFA	✓, Interne consistentie met zowel α als ω is gerapporteerd.	✓
GBPNS-PE (Sturm, 2020)	✓	-	-	✓, MCFA	✓, α en (CR) zijn gerapporteerd.	✓

Noot. Per stap is op basis van het stappenplan van Dima (2018) nagegaan of dit wel, deels, of niet aanwezig is in het artikel. De stappen zijn terug te vinden in de horizontale kolommen. Als de stap voldoet is er een (✓) opgenomen, als deze deels voldoet is er een (±) opgenomen, is er niet voldaan dan is er een (-) opgenomen. Afkortingen: *M* = Gemiddelde; *SD* = Standaarddeviatie; *MDA* = Missing data-analyse; *CR* = Composite Reliability; *CFA* = confirmatieve factoranalyse; *MCFA* = Mixture of Common Factor Analyzers; *MGCFA* = Multiple-Group Confirmatory Factor Analysis; *ESEM* = exploratory structural equation modeling; *MEFA* = Multilevel Exploratory Factor Analysis; *EFA* = Exploratory Factor Analysis; *EGA* = exploratory graph analysis; *MDS* = Multidimensional Scaling; *AVE* = Average Variance Extracted.

De resultaten vanuit Dima (2018) laten zien dat de meeste elementen vanuit Dima terugkomen. Op het gebied van de itemanalyse rapporteerden alle studies beschrijvende statistieken, maar diepgaandere itemanalyses kwamen nagenoeg niet voor. Van de twaalf studies was er maar één studie die mokkenanalyse gebruikte. Stap drie ontbrak bij alle studies. Op het gebied van factoranalyse maakten alle instrumenten gebruik van bevestigende factoranalyse, vaak met geavanceerdere modellen zoals ESEM, multilevel CFA of bifactor-modellen. De interne consistentie was voldoende tot goed, hierbij werden zowel Cronbach's alfa als McDonald's omega gebruikt. In enkele gevallen werd ook test-hertestbetrouwbaarheid onderzocht, met acceptabele tot hoge correlaties. Op de TASCQ-short na lieten alle instrumenten verbanden zien met relevante variabelen zoals motivatie, behoeftebevrediging en welzijn. Dit duidt op ondersteuning van de criteriumvaliditeit.

Conclusie en Discussie

Deze systematische review had als doel een overzicht te creëren van beschikbare meetinstrumenten voor het in kaart brengen van autonomiebehoefte en de ervaren autonomieondersteuning bij leerlingen in het primair en voortgezet onderwijs. Daarnaast werd gekeken naar de psychometrische kwaliteit van deze instrumenten. Autonomie wordt binnen de zelfdeterminatietheorie (SDT) doorgaans gezien als de ervaring van keuzevrijheid, zelfsturing en eigenaarschap (Deci & Ryan, 2000; Ryan & Deci, 2020; Reeve, 2006). Deze review draagt bij aan de bestaande literatuur door niet alleen te inventariseren welke instrumenten worden gebruikt, maar ook door te beschrijven hoe autonomie wordt geoperationaliseerd en hoe de psychometrische kwaliteit van deze instrumenten is onderbouwd (Mokkink et al., 2024; Prinsen et al., 2018; Dima, 2018).

Uit de resultaten blijkt duidelijk dat autonomie vaak niet als geïsoleerd construct wordt gemeten. De meeste instrumenten operationaliseren autonomie als onderdeel van bredere theoretische kaders, zoals de basisbehoeften binnen de ZDT (Ryan & Deci, 2020). De rol van autonomie blijft daardoor mogelijk onderbelicht. De studie van Koçak et al. (2025) rapporteert bijvoorbeeld samenhang tussen *need satisfaction* of *need frustration* en motivatie of welzijn, zonder specifieke analyse van de bijdrage van autonomie. Hierdoor blijft onduidelijk in hoeverre gevonden effecten toe te schrijven zijn aan autonomie zelf. De manier waarop autonomie geconceptualiseerd wordt heeft dus invloed op de interpretatie van de resultaten (Mokkink et al., 2024).

Daarnaast moet er een verschil worden aangeduid tussen instrumenten die de opvattingen van leerlingen over autonomieondersteuning meten en instrumenten die de autonomiebehoefte en de behoefte op het gebied van voldoening en frustratie meten. De typen instrumenten vallen allemaal onder de noemer van autonomie, maar er zijn wel inhoudelijk verschillen te zien. Bij autonomieondersteuning gaat het over wat leerlingen ervaren dat de docent doet, terwijl het bij autonomiebehoefte, bevrediging en frustratie gaat over de ervaring van de leerling zelf (Deci & Ryan, 2000; Reeve, 2006; Ryan & Deci, 2020). Dit wijst op een heterogeen construct, daardoor moet er opgepast worden met het maken van vergelijkingen tussen instrumenten en studies. Het is daarom belangrijk om in kaart te brengen welk autonomie-aspect wordt gemeten, aangezien een vergelijkbare gedraging in een andere context een andere betekenis kan hebben (Putnick & Bornstein, 2016).

Wat betreft de psychometrische kwaliteit laten de resultaten zien dat de meeste instrumenten voldoende tot goede interne consistentie rapporteren en dat de factorstructuur in

veel studies wordt bevestigd. Dit wijst op een redelijke betrouwbaarheid en constructvaliditeit van verschillende schalen (Mokkink et al., 2024; Prinsen et al., 2018). Met name instrumenten als de MD-PASS-PE, IBQ-PE, SIS/SIS-PE, BPNSFS en SBPNFSS laten relatief sterke psychometrische waarden zien en kunnen daarom in dit onderzoek als de sterkere meetinstrumenten worden beschouwd. Tegelijkertijd laten schalen als de CARR en subschalen van de SIS lagere betrouwbaarheid zien, hierdoor is voorzichtigheid vereist bij de interpretatie en toepassing.

Verder valt op dat bepaalde meeteigenschappen die juist belangrijk zijn wanneer je verandering over tijd wilt meten beperkt worden gerapporteerd, zoals responsiviteit en meetfout volgens de COSMIN-richtlijnen. Daardoor blijft het op basis van de huidige literatuur lastiger om te bepalen hoe gevoelig sommige instrumenten zijn voor verandering en hoe precies scoreverschillen geïnterpreteerd kunnen worden (Mokkink et al., 2024; Prinsen et al., 2018).

Een belangrijke methodologische kanttekening die bij de interpretatie van de resultaten moet worden geplaatst is de manier waarop het onderdeel meetfout wordt beoordeeld. Binnen de COSMIN-richtlijnen wordt meetfout opgevat als absolute meetonnauwkeurigheid en wordt aanbevolen dit te rapporteren met globale foutmaten zoals de *Standard Error of Measurement* (SEM; standaardmeetfout), de *Smallest Detectable Change* (SDC) of *Limits of Agreement* (LoA) (Mokkink et al., 2024; Prinsen et al., 2018). In geen van de studies werden deze maten gerapporteerd, waardoor dit onderdeel volgens COSMIN gescoord moest worden als 'niet gerapporteerd'. Dit maakt duidelijk dat er vooral sprake is van een verschil bij het rapporteren. Waarbij volgens COSMIN de meetfout ontbreekt, omdat SEM/SDC/LoA niet worden gerapporteerd, terwijl studies vanuit moderne modelrapportage vaak op een andere manier laten zien dat metingen niet perfect zijn. De conclusie is dus dat de COSMIN-uitkomst op meetfout hier voorzichtig geïnterpreteerd moet worden (Mokkink et al., 2024; Prinsen et al., 2018).

Daarnaast blijkt vanuit het stappenplan van Dima (2018) dat de meeste instrumenten pas in een later stadium van schaalontwikkeling grondig zijn geëvalueerd. Vroege stappen zoals systematische itemanalyse, Mokken-schaalanalyse en IRT/Rasch-modellen worden nauwelijks gerapporteerd, terwijl Dima (2018) juist benadrukt dat dit soort stappen helpen om het functioneren van individuele items beter te begrijpen (Kline, 2016). Alleen in de studie van Van Aart et al. (2017) is een Mokkenanalyse uitgevoerd, terwijl parametrische IRT/Rasch in geen van de studies voorkomt. Dit betekent dat de literatuur vooral sterk is in rapportage op schaalniveau (factorstructuur en interne consistentie), maar dat diepgaandere itemniveau-

evaluatie beperkt blijft.

Over het algemeen kunnen we op basis van de resultaten stellen dat er bruikbare en theoretisch sterke meetinstrumenten beschikbaar zijn voor het meten van autonomie in het onderwijs, maar hierbij is het wel belangrijk om te kijken in binnen welke context het instrument heeft gefunctioneerd en op welke manier autonomie gedefinieerd is (Dima, 2018; Mokkink et al., 2024).

Beperkingen

Hoewel in deze review vier grote databases zijn geraadpleegd (ERIC, Web of Science, Scopus en PubMed), kan niet worden uitgesloten dat relevante studies buiten de gehanteerde zoekstrategie zijn gebleven. Dit hangt mede samen met de gekozen afbakening in publicatieperiode (vanaf 2016). Deze ondergrens is bewust gekozen om een actueel overzicht te bieden van recent ontwikkelde of recent opnieuw gevalideerde meetinstrumenten. Tegelijkertijd kan deze keuze ertoe hebben geleid dat oudere, mogelijk veelgebruikte instrumenten en eerdere validatiestudies buiten beeld zijn gebleven. Hierdoor kan het beeld dat deze scriptie schetst mogelijk niet representatief zijn voor het gehele veld. De gekozen afbakening beperkt daarmee de volledigheid van het overzicht en daardoor is er voorzichtigheid nodig bij eventuele generalisatie.

Daarbij is in deze review gekozen om alleen peer-reviewed wetenschappelijke artikelen te includeren en geen grijze literatuur mee te nemen. Deze keuze hangt samen met het doel van de review om instrumenten te beoordelen op de psychometrische waarde. Voor een consistente COSMIN-beoordeling is er gedetailleerde en controleerbare rapportage nodig. Hoewel peer review geen garantie biedt voor volledige rapportage, vergroot het wel de kans dat deze informatie beschikbaar is. Dit ondersteunt op zijn beurt de reproduceerbaarheid en consistentie van de COSMIN-beoordeling (Mokkink et al., 2024; Prinsen et al., 2018). Tegelijkertijd betekent dit dat relevante instrumenten of aanvullende validatiestudies die buiten het peer-reviewproces zijn verspreid zijn gemist. In lijn met transparante rapportageprincipes in systematische reviews wordt deze beperking daarom expliciet meegenomen in de interpretatie van de resultaten (Page et al., 2021).

Een tweede beperking betreft praktische tijdsbeperkingen die samenhangen met de uitvoering van deze masterthesis. Hoewel een systematische zoekstrategie is gehanteerd en het screeningproces is afgerond op basis van de vooraf vastgestelde inclusiecriteria was het door tijdgebrek niet mogelijk om het reviewproces verder te verbreden met bijvoorbeeld snowballing en een volledig dubbel uitgevoerd screeningproces. Het uiteindelijke aantal

geïnccludeerde studies (n = 12) is het resultaat van de gehanteerde inclusiecriteria en de beschikbare literatuur binnen de gekozen afbakening. Deze beperkingen vergroten echter wel de kans dat relevante instrumenten of aanvullende validatiestudies niet zijn meegenomen.

Een andere beperking betreft de kwaliteit en volledigheid van de rapportage in de geïnccludeerde studies zelf. Veel artikelen rapporteren slechts een beperkt aantal psychometrische eigenschappen, waardoor de COSMIN-onderdelen van responsiviteit en meetfout als ‘niet gerapporteerd’ werden beoordeeld (Mokkink et al., 2024; Prinsen et al., 2018). Voor responsiviteit betekent dit daadwerkelijk dat er onvoldoende informatie beschikbaar is over de gevoeligheid van instrumenten voor verandering.

Ook moet worden vermeld dat het volledige screeningproces niet door twee onafhankelijke beoordelaars is uitgevoerd. Alleen de definitieve selectie van de geïnccludeerde studies en instrumenten is gezamenlijk besproken en dubbel gecheckt. Dit vergroot het risico op selectiebias en benadrukt het belang van voorzichtigheid bij interpretatie. Bovendien is de beoordeling van de psychometrische kwaliteit deels afhankelijk van interpretatie. Hoewel er is getracht zo consistent mogelijk te werken binnen de beoordelingskaders kan subjectiviteit niet volledig worden uitgesloten.

Ten slotte zijn de geïnccludeerde studies uitgevoerd in verschillende landen en onderwijscontexten. Dit is inhoudelijk waardevol omdat het laat zien dat instrumenten breed worden toegepast, maar het beperkt ook de culturele generaliseerbaarheid van de bevindingen. Autonomie kan in verschillende onderwijs- en cultuurcontexten een andere betekenis hebben, waardoor resultaten en psychometrische kenmerken niet automatisch te generaliseren zijn (Putnick & Bornstein 2016).

Wetenschappelijke implicaties

De bevindingen van dit onderzoek hebben implicaties voor zowel de onderwijspraktijk als wetenschappelijk onderzoek. Voor de praktijk is het aan te raden om gebruik te maken van vragenlijsten met een relatief sterk onderbouwde psychometrische basis, zoals de MD-PASS-PE, SIS/SIS-PE en SBPNFSS. Op basis van de in deze review gerapporteerde meeteigenschappen bieden deze instrumenten een relatief betrouwbare en valide manier om autonomieondersteuning en autonomiebehoeften bij leerlingen in kaart te brengen binnen primair en voortgezet onderwijs (Mokkink et al., 2024; Prinsen et al., 2018). Om hier een kanttekening bij te plaatsen is het wel van belang om te kijken naar de haalbaarheid van een instrument. Zowel aan de praktische kant waarbij het volgens Mokkink et al. (2024) gaat om bijvoorbeeld tijd en geld als aan de ecologische kant waarbij het volgens Putnick en Bornstein

(2016) gaat om het afnemen van de test onder verschillende omstandigheden.

Daarnaast vraagt het gebruik van deze instrumenten om voorzichtigheid bij interpretatie van de resultaten, vooral wanneer men uitspraken wil doen over verandering of effecten van interventies.

Voor toekomstig onderzoek is het aan te raden om de psychometrische onderbouwing van bestaande instrumenten verder te versterken. Er is met name behoefte aan systematisch onderzoek naar test-hertestbetrouwbaarheid en responsiviteit. Door dit beter in kaart te brengen kan worden vastgesteld in hoeverre instrumenten geschikt zijn om veranderingen over tijd te meten (Mokkink et al., 2024; Prinsen et al., 2018; Putnick & Bornstein, 2016).

Daarnaast zou een combinatie tussen het streven naar het rapporteren volgens COSMIN, en tegelijkertijd het rekening houden met moderne psychometrische modellen kunnen helpen bij het verkrijgen van een completer beeld van de kwaliteit (Hox & Bechger, 1998; Kline, 2016; Mokkink et al., 2024; Prinsen et al., 2018).

Verder is het wenselijk om bestaande instrumenten al dan niet opnieuw te valideren binnen de Nederlandse onderwijscontext. Kline (2016) benadrukt dat het regelmatig voorkomt dat onderzoekers op basis van eerdere studies ten onrechte aannemen dat een test betrouwbaar is. Hij noemt dit proces ook wel *reliability induction*, dit is problematisch aangezien psychometrische eigenschappen van een steekproef met een specifieke culturele achtergrond niet te generaliseren zijn naar een steekproef die een andere culturele achtergrond heeft. Aangezien de studies in deze review in verschillende landen en onderwijscontexten zijn uitgevoerd, is het niet vanzelfsprekend dat factorstructuur, betrouwbaarheid en interpretatie van de items te generaliseren zijn naar de Nederlandse context. Kline (2016) geeft aan dat lokale validatie zou kunnen bijdragen aan een betere generaliseerbaarheid en praktische toepasbaarheid binnen scholen, omdat dit minder kans heeft op het verkrijgen van misleidende resultaten.

Tot slot is het aan te bevelen dat toekomstig onderzoek onderzoekt hoe kwantitatieve vragenlijsten zich verhouden tot andere meetmethoden. Door bijvoorbeeld gebruik te maken van aanvullende observaties of mixed-methods zou de convergente validiteit verhoogd kunnen worden (Kline, 2016). Een combinatie van deze methoden kan helpen om autonomieondersteuning en behoeftebevrediging niet alleen als zelfrapportage te begrijpen, maar ook te verbinden aan feitelijk gedrag en de context in de klas. Door dit te combineren kan er mogelijk een vollediger beeld van autonomieondersteuning in de specifieke onderwijscontext worden geschetst (Deci & Ryan, 2000; Reeve, 2006).

Literatuurlijst

- Admiraal, W., Lockhorst, D., Post, L., & Kester, L. (2024). Effects of Students' Autonomy Support on their Self-Regulated Learning Strategies: Three Field Experiments in Secondary Education. *International Journal of Research in Education and Science*, 10(1), 1–20. <https://doi.org/10.46328/ijres.3343>
- Admiraal, W., Nieuwenhuis, G., Kooij, Y., Dijkstra, T., & Cloosterman, I. (2019). Perceived Autonomy Support in Primary Education in the Netherlands: Differences between Teachers and Their Students. *World Journal of Education*, 9(4), 1-12. <https://doi.org/10.5430/wje.v9n4p1>
- Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J., & Reeve, J. (2019). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Journal of Educational Psychology*, 111(3), 497–521. <https://doi.org/10.1037/edu0000293> *
- Ahn, I., Patrick, H., Chiu, M. M., & Levesque-Bristol, C. (2019). Measuring teacher practices that support student motivation: Examining the factor structure of the teacher as social context questionnaire using multilevel factor analyses. *Journal of Psychoeducational Assessment*, 37(6), 743–756. <https://doi.org/10.1177/0734282918791655> *
- Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, 84(6), 740–756. [https://doi.org/10.1002/1098-237X\(200011\)84:6<740::AID-SCE4>3.0.CO;2-3](https://doi.org/10.1002/1098-237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3)
- Burgueño, R., & Medina-Casaubón, J. (2021). Validity and reliability of the Interpersonal Behaviors Questionnaire in Physical Education With Spanish Secondary School Students. *Perceptual and Motor Skills*, 128(1), 522–545. <https://doi.org/10.1177/0031512520948286> *
- Burgueño, R., Abós, Á., Sevil-Serrano, J., Haerens, L., De Cocker, K., & García-González, L. (2023). A Circumplex Approach to (de)motivating Styles in Physical Education: Situations-In-School–Physical Education Questionnaire in Spanish Students, Pre-Service, and In-Service Teachers. *Measurement in Physical Education And Exercise Science*, 28(1), 86–108. <https://doi.org/10.1080/1091367x.2023.2248098> *

- Burgueño, R., Macarro-Moreno, J., & Medina-Casaubón, J. (2020). Psychometry of the Multidimensional Perceived Autonomy Support Scale in Physical Education With Spanish Secondary School Students. *Sage Open*, 10(1). <https://doi.org/10.1177/2158244019901253> *
- Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van Der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., & Verstuyf, J. (2014). Basic psychological need satisfaction, frustration, and need strength across four cultures. *Motivation And Emotion*, 39(2), 216–236. <https://doi.org/10.1007/s11031-014-9450-1>
- Cheon, S. H., Reeve, J., & Marsh, H. W. (2023). Autonomy-Supportive Teaching Enhances Prosocial and Reduces Antisocial Behavior via Classroom Climate and Psychological Needs: A Multilevel Randomized Control Intervention. *Journal Of Sport and Exercise Psychology*, 45(1), 26–40. <https://doi.org/10.1123/jsep.2021-0337>
- Cheon, S. H., Reeve, J., & Song, Y. (2016). A Teacher-Focused Intervention to Decrease PE Students' Amotivation by Increasing Need Satisfaction and Decreasing Need Frustration. *Journal Of Sport and Exercise Psychology*, 38(3), 217–235. <https://doi.org/10.1123/jsep.2015-0236>
- Conesa, P. J., & Duñabeitia, J. A. (2021). The Basic Psychological Needs in the Classroom Scale (BPN-CS). *Behavioral Sciences*, 11(7), 96. <https://doi.org/10.3390/bs11070096> *
- COSMIN. (z.d.). *Consensus-based Standards for the selection of health Measurement Instruments*. <https://www.cosmin.nl/about/>
- COTAN. (z.d.). *Beoordelingssysteem – COTAN-documentatie online*. Nederlands Instituut van Psychologen. <https://www.cotandocumentatie.nl/cotan/beoordelingssysteem/>
- Deci, E.L., & Ryan, R.M. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Dima, A. L. (2018). Scale validation in applied health research: tutorial for a 6-step R-based psychometrics protocol. *Health Psychology and Behavioral Medicine*, 6(1), 136–161. <https://doi.org/10.1080/21642850.2018.1472602>

- Elsman, E. B. M., Mekkink, L. B., Terwee, C. B., Prinsen, C. A. C., Boers, M., de Vet, H. C. W., ... & Roorda, L. D. (2024). Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. *Quality of Life Research*, 33, 2029–2046. <https://doi.org/10.1007/s11136-024-03634-y>
- Grolnick, W., & Ryan, R. M. (1987). Autonomy in children's learning: an experimental and individual difference investigation. *Journal of Personality and Social Psychology*, 52, 890e898. <https://doi.org/10.1037/0022-3514.52.5.890>.
- Heijnen, J. Koopman, L., van Schaik, Monique., & Troost, C. (2013) Autonomiebehoefte en Autonomieondersteuning op Zwarte en Witte Scholen.
- Hornstra, L., van der Veen, I., Peetsma, T., & Volman, M. (2013). Developments in motivation and achievement during primary school: A longitudinal study on group-specific differences. *Learning and Individual Differences*, 23, 195-204. <https://doi.org/10.1016/j.lindif.2012.09.004>
- Hornstra, L., Weijers, D., Peetsma, T., & van der Veen, I. (2016). Motiverend lesgeven. Hoe kan je als docent condities scheppen waarbinnen leerlingen zichzelf kunnen motiveren? Handleiding voor docenten. Universiteit Utrecht.
- Hox, J. J. & Bechger, M. T. (2002). An Introduction to Structural Equation Modeling. *Family Science Review*, 11, 354-373.
- Inman, R. A., Costa, P. J. C., & Moreira, P. A. S. (2023). Psychometric properties of the Portuguese Adolescent Students' Basic Psychological Needs at School Scale (ASBPNSS) and evidence of differential associations with indicators of subjective wellbeing. *Journal of Psychoeducational Assessment*, 41(1), 100–119. <https://doi.org/10.1177/07342829221125843> *
- Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: it's not autonomy support or structure, but autonomy support and structure. *Journal of Educational Psychology*, 102, 588e600. <https://doi.org/10.1037/a0019682>
- John Marshall Reeve & Sung Hyeon Cheon (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice, *Educational Psychologist*, 56:1, 54-77, <https://doi.org/10.1080/00461520.2020.1862657>

- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.
- Koçak, A., Mouratidis, A., Alp Christ, A., Michou, A., & Sayıl, M. (2025). Basic Psychological Needs Satisfaction and Frustration Scale (BPNSFS): A Turkish adaptation study. *Psychology in the Schools*, 62(7), 1987–2000.
<https://doi.org/10.1002/pits.23449> *
- Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: age differences and academic correlates. *Journal of Educational Psychology*, 97, 184e196. <https://doi.org/10.1037/0022-0663.97.2.184>
- Li, J., Shao, S., Huebner, E. S., & Tian, L. (2025). Preliminary Development and Validation of the Students' Basic Psychological Needs Frustration at School Scale. *Psychology in The Schools*. <https://doi.org/10.1002/pits.70111> *
- Liu, J.-D., & Chung, P.-K. (2017). Factor structure and measurement invariance of the need-supportive teaching style scale for physical education. *Perceptual and Motor Skills*, 124(4), 864–879. <https://doi.org/10.1177/0031512517712803> *
- Iyengar, S. S., & Lepper, M. R. (1999). Rethinking the value of choice: A cultural perspective on intrinsic motivation. *Journal of Personality and Social Psychology*, 76(3), 349–366.
<https://doi.org/10.1037/0022-3514.76.3.349>
- Mammadov, S., & Schroeder, K. (2023). A meta-analytic review of the relationships between autonomy support and positive learning outcomes. *Contemporary Educational Psychology*, 75, 102235. <https://doi.org/10.1016/j.cedpsych.2023.102235>
- Manuel, C. D., Magalhães, C. R., Huber, C. M., Smerek, L., Costa, A. F., & Alves, J. R. (2024). Cross-Cultural Adaptation of a Questionnaire Measuring Organizational Citizenship Behavior towards the Environment. *Administrative Sciences*, 14(3), 57.
<https://doi.org/10.3390/admsci14030057>
- Miao, J., & Ma, L. (2023). Teacher Autonomy Support Influence on Online Learning Engagement: The Mediating Roles of Self-Efficacy and Self-Regulated Learning. *SAGE Open*, 13(4). <https://doi.org/10.1177/21582440231217737>
- Minnaert, A., & Odenthal, L. (Eds.) (2018). Motivatie is een werkwoord. CPS.
<https://research.rug.nl/en/publications/motivatatie-is-een-werkwoord>

- Mokkink, L. B., Elsmann, E. B., & Terwee, C. B. (2024). COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research*, 33, 2929–2939. <https://doi.org/10.1007/s11136-024-03761-6>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, 19, 539–549. <https://doi.org/10.1007/s11136-010-9606-8>
- Ntumi, S., & Antwi-Agyakwa, K. T. (2022). A Systematic Review of Reporting of Psychometric Properties in Educational Research. *Mediterranean Journal Of Social & Behavioral Research*, 6(2), 53–59. <https://doi.org/10.30935/mjosbr/11912>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews*. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin*, 134(2), 270–300. <https://doi.org/10.1037/0033-2909.134.2.270>
- Patzak, A., & Zhang, X. (2025). Blending Teacher Autonomy Support and Provision of Structure in the Classroom for Optimal Motivation: A Systematic Review and Meta-Analysis. *Educational Psychology Review*, 37(1). <https://doi.org/10.1007/s10648-025-09994-2>
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27, 1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- PRISMA. (2024). *PRISMA: Transparent reporting of systematic reviews and meta-analyses*. <https://www.prisma-statement.org/> Published by: Taylor & Francis, Ltd. Stable.
- Putnick, D. L., & Bornstein, M. H. (2016). *Measurement Invariance Conventions and Reporting*. (PMC/open access). [PMC](https://pubmed.ncbi.nlm.nih.gov/27000000/)

- Ramsay, G., Williams, M., Marks, E., & Morgan, G. (2025). A COSMIN systematic review of the psychometric properties of instruments that measure climate change-related distress. *Cogent Mental Health*, 4(1), 1–27.
<https://doi.org/10.1080/28324765.2025.2449878>
- Reeve, J. (2006). Teachers as Facilitators: What Autonomy-Supportive Teachers Do and Why Their Students Benefit. *The Elementary School Journal*, 106(3), 225–236.
<https://doi.org/10.1086/501484>
- Reeve, J., Jang, H., Carrell, D., Jeon, S. & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion*, 28(2), 147-169.
<https://doi.org/10.1023/B:MOEM.0000032312.95499.6f>
- Rijksuniversiteit Groningen. (z.d.). *Systematic reviews – LibGuides*.
<https://libguides.rug.nl/systematic-reviews>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860.
<https://doi.org/10.1016/j.cedpsych.2020.101860>
- Ryan, R.M., & Deci, E.L. (2000). Self-determination Theory and the Facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78.
<https://doi.org/10.1037/0003-066X.55.1.68>
- Scholtes, V. A., Terwee, C. B., & Poolman, R. W. (2011). What makes a measurement instrument valid and reliable? *Injury*, 42(3), 236–240.
<https://doi.org/10.1016/j.injury.2010.11.042>
- Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21–44). Springer.
https://doi.org/10.1007/978-1-4614-2018-7_2
- Sturm, D. J., Bachner, J., Haug, S., & Demetriou, Y. (2020). The German Basic Psychological Needs Satisfaction in Physical Education Scale: Adaption and Multilevel Validation in a Sample of Sixth-Grade Girls. *International Journal of Environmental Research and Public Health*, 17(5), 1554. <https://doi.org/10.3390/ijerph17051554> *

- Sung Hyeon Cheon, Johnmarshall Reeve, and Herbert W. Marsh (2023). Autonomy-Supportive Teaching Enhances Prosocial and Reduces Antisocial Behavior via Classroom Climate and Psychological Needs: A Multilevel Randomized Control Intervention, *Journal of Sport and Exercise Psychology*, 2023, 45, 26-40, <https://doi.org/10.1123/jsep.2021-0337>
- Tsai, Y.-M., Kunter, M., Lüdtke, O., Trautwein, U., & Ryan, R. M. (2008). What makes lessons interesting? The role of situational and individual factors in three school subjects. *Journal of Educational Psychology*, 100(2), 460–472. <https://doi.org/10.1037/0022-0663.100.2.460>
- van Aart, I., Hartman, E., Elferink-Gemser, M., Mombarg, R., & Visscher, C. (2017). Relations among basic psychological needs, PE-motivation and fundamental movement skills in 9–12-year-old boys and girls in Physical Education. *Physical Education and Sport Pedagogy*, 22(1), 15–34. <https://doi.org/10.1080/17408989.2015.1112776> *
- Van Burg, J. C. (2011). Kwaliteitscriteria voor ontwerpgericht wetenschappelijk onderzoek. In J. van Aken, & D. Andriessen (editors), *Handboek ontwerpgericht wetenschappelijk onderzoek: Wetenschap met effect* (blz. 146-164). Boom Lemma. <https://research.vu.nl/ws/portalfiles/portal/2981134/Van%20Burg%202011%20Kwaliteitscriteria%20voor%20ontwerpgericht%20wetenschappelijk%20onderzoek.pdf>
- Vansteenkiste, M., & Soenens, B. (2015). *Vitamines voor groei: Ontwikkeling voeden vanuit de Zelf-Determinatie Theorie*. https://www.accoutgeverij.nl/sites/default/files/book_pdfs/9789462922860_fragm.pdf
- Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A., Aelterman, N., Haerens, L., & Beyers, W. (2012). Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learning and Instruction*, 22(6), 431–439. <https://doi.org/10.1016/j.learninstruc.2012.04.002>
- Verbeeck, K. (2010). *Op eigen vleugels. Autonomie voor kinderen in het basisonderwijs*. 's-Hertogenbosch: KPC Groep in opdracht van het ministerie van OCW.
- Vlachopoulos, S. P., & Michailidou, S. (2009). Development and Initial Validation of a Measure of Autonomy, Competence, and Relatedness in Exercise: The Basic

Psychological Needs in Exercise Scale. *Measurement in Physical Education and Exercise Science*, 10(3), 179–201. https://doi.org/10.1207/s15327841mpee1003_4

Yang, D., Chen, P., Wang, H., Wang, K., & Huang, R. (2022). Teachers' autonomy support and student engagement: A systematic literature review of longitudinal studies. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.925955>

Bijlagen

Bijlage A.

Overzicht zoekstrings en databases.

Database	Zoekstring	Filters	Datum	Hits
ERIC	(autonomy OR "autonomy support*" OR "autonomy need*" OR "autonomy supportive teaching" OR "support for autonomy" OR "autonomy frustration*") AND (student* OR pupil* OR learner* OR "primary education" OR "secondary education" OR school*) AND (questionnaire* OR scale* OR instrument* OR survey* OR measurement* OR test*) AND (psychometric* OR reliability OR validity OR "measurement property*")	Peer-reviewed, Engels/Nederlands, ≥2016	29-9-2025	163
Web of Science	TS=(autonomy OR "autonomy support*" OR "autonomy need*" OR "autonomy supportive teaching" OR "support for autonomy" OR "autonomy frustration*") AND TS=(student* OR pupil* OR learner* OR "primary education" OR "secondary education" OR school*) AND TS=(questionnaire* OR scale* OR instrument* OR survey* OR measurement* OR test*) AND TS=(psychometric* OR reliability OR validity OR "measurement property*")	Peer-reviewed, Engels/Nederlands, ≥2016	29-9-2025	564

Scopus	TITLE-ABS-KEY(autonomy OR "autonomy support*" OR "autonomy need*" OR "autonomy supportive teaching" OR "support for autonomy" OR "autonomy frustration*") AND TITLE-ABS-KEY(student* OR pupil* OR learner* OR "primary education" OR "secondary education" OR school*) AND TITLE-ABS-KEY(questionnaire* OR scale* OR instrument* OR survey* OR measurement* OR test*) AND TITLE-ABS-KEY(psychometric* OR reliability OR validity OR "measurement property*")	Peer-reviewed, Engels/Nederlands, ≥2016	29-9-2025	438
PubMed	(autonomy[tiab] OR "autonomy support*"[tiab] OR "autonomy need*"[tiab] OR "autonomy supportive teaching"[tiab] OR "support for autonomy"[tiab] OR "autonomy frustration*"[tiab]) AND (student*[tiab] OR pupil*[tiab] OR learner*[tiab] OR "primary education"[tiab] OR "secondary education"[tiab] OR school*[tiab]) AND (questionnaire*[tiab] OR scale*[tiab] OR instrument*[tiab] OR survey*[tiab] OR measurement*[tiab] OR test*[tiab]) AND (psychometric*[tiab] OR reliability[tiab] OR validity[tiab] OR "measurement property*"[tiab])	Peer-reviewed, Engels/Nederlands, ≥2016	29-9-2025	128

Noot. Overzicht van de databases waarin is gezocht en welke zoekstrings en filters zijn gebruikt in overeenstemming met de PRISMA- en COSMIN-richtlijnen.