



Signaleringswaarde: als waarde kleiner of gelijk is aan 35: risicoscore 0 , als waarde groter of gelijk is aan 35: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 7

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.4 *Veranderingen in het lerarenteam (mbo)*

Beschrijving /definitie: Percentage van leraren bestand in loondienst van het bestuur dat onderhevig is aan mutaties (personeel dat het bestuur heeft verlaten of juist bij het bestuur is komen werken). De kolom jaar refereert aan de start van het schooljaar (2020 gaat over schooljaar 20-21).

Berekeningswijze: ((aantal leraren dat bg heeft verlaten + aantal leraren dat nieuw is komen werken op bg) / (aantal leraren op bg + aantal leraren dat bg heeft verlaten))*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 35: risicoscore 0 , als waarde groter of gelijk is aan 35: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 7

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.5 *Gem. ervaringsjaren leraren op bg (po)*

Beschrijving /definitie: Het gemiddelde van het aantal jaren dat leraren in loondienst al werkzaam zijn binnen het bestuur. (hierbij worden jaren geregistreerd sinds 2006, omdat pas vanaf dat moment een dataset is opgebouwd). De kolom jaar refereert aan de start van het schooljaar (2020 gaat over schooljaar 20-21).

Berekeningswijze: Gemiddelde van het aantal ervaringsjaren

Signaleringswaarde: als waarde kleiner of gelijk is aan 6.91: risicoscore 1 , als waarde groter of gelijk is aan 6.91: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 3

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.6 *Gem. ervaringsjaren leraren op bg (so)*

Beschrijving /definitie: Het gemiddelde van het aantal jaren dat leraren in loondienst al werkzaam zijn binnen het bestuur. (hierbij worden jaren geregistreerd sinds 2006, omdat pas vanaf dat moment een dataset is opgebouwd). De kolom jaar refereert aan de start van het schooljaar (2020 gaat over schooljaar 20-21).

Berekeningswijze: Gemiddelde van het aantal ervaringsjaren

Signaleringswaarde: als waarde kleiner of gelijk is aan 5.42: risicoscore 1 , als waarde groter of gelijk is aan 5.42: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 3

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.7 *Gem. ervaringsjaren leraren op bg (vo)*

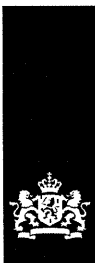
Beschrijving /definitie: Het gemiddelde van het aantal jaren dat leraren in loondienst al werkzaam zijn binnen het bestuur. (hierbij worden jaren geregistreerd sinds 2006, omdat pas vanaf dat moment een dataset is opgebouwd). De kolom jaar refereert aan de start van het schooljaar (2020 gaat over schooljaar 20-21).

Berekeningswijze: Gemiddelde van het aantal ervaringsjaren

Signaleringswaarde: als waarde kleiner of gelijk is aan 6.97: risicoscore 1 , als waarde groter of gelijk is aan 6.97: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 3

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x



4.6.8 *Aandeel onderwijzend personeel (po)*

Beschrijving /definitie: Aandeel onderwijsgevend personeel t.o.v. totaal personeel binnen de sector PO. Indiceert de mate waarin het bestuur personeel inzet op het primaire proces. Jaar refereert aan het boekjaar

Berekeningswijze: (Onderwijsgevend personeel/Totaal personeel in fte's)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 65.91: risicoscore 1, als waarde groter of gelijk is aan 65.91: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 5

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.9 *Aandeel onderwijzend personeel (so)*

Beschrijving /definitie: Aandeel onderwijsgevend personeel t.o.v. totaal personeel binnen de sector SO. Indiceert de mate waarin het bestuur personeel inzet op het primaire proces. Jaar refereert aan het boekjaar

Berekeningswijze: (Onderwijsgevend personeel/Totaal personeel in fte's)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 40.96: risicoscore 1, als waarde groter of gelijk is aan 40.96: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.10 *Aandeel onderwijzend personeel (vo)*

Beschrijving /definitie: Aandeel onderwijsgevend personeel t.o.v. totaal personeel binnen de sector VO. Indiceert de mate waarin het bestuur personeel inzet op het primaire proces. Jaar refereert aan het boekjaar

Berekeningswijze: (Onderwijsgevend personeel/Totaal personeel in fte's)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 66.49: risicoscore 1, als waarde groter of gelijk is aan 66.49: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.11 *Aandeel onderwijzend personeel (mbo)*

Beschrijving /definitie: Aandeel onderwijsgevend personeel t.o.v. totaal personeel binnen de sector MBO. Indiceert de mate waarin het bestuur personeel inzet op het primaire proces. Jaar refereert aan het boekjaar

Berekeningswijze: (Onderwijsgevend personeel/Totaal personeel in fte's)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 51: risicoscore 1, als waarde groter of gelijk is aan 51: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.12 *Aandeel parttime onderwijzend personeel (po)*

Beschrijving /definitie: Aandeel parttimers onderwijsgevend personeel t.o.v. totaal personeel binnen de sector PO. De veronderstelling is dat een groter aandeel parttimers onder het onderwijsgevend personeel, kan leiden tot risico's voor de onderwijskwaliteit. Een parttime dienstverband is een dienstverband van 0,8FTE of minder. Jaar refereert aan het boekjaar

Berekeningswijze: (Aantal FTE onderwijsgevend personeel met een parttime dienstverband/totaal onderwijsgevend personeel van een bestuur)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 59.77: risicoscore 0, als waarde groter of gelijk is aan 59.77: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x



4.6.13 *Aandeel parttime onderwijzend personeel (so)*

Beschrijving /definitie: Aandeel parttimers onderwijsgevend personeel t.o.v. totaal personeel binnen de sector SO. De veronderstelling is dat een groter aandeel parttimers onder het onderwijsgevend personeel, kan leiden tot risico's voor de onderwijskwaliteit. Een parttime dienstverband is een dienstverband van 0,8FTE of minder. Jaar refereert aan het boekjaar

Berekeningswijze: (Aantal FTE onderwijsgevend personeel met een parttime dienstverband/totaal onderwijsgevend personeel van een bestuur)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 52.64: risicoscore 0, als waarde groter of gelijk is aan 52.64: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.14 *Aandeel parttime onderwijzend personeel (vo)*

Beschrijving /definitie: Aandeel parttimers onderwijsgevend personeel t.o.v. totaal personeel binnen de sector VO. De veronderstelling is dat een groter aandeel parttimers onder het onderwijsgevend personeel, kan leiden tot risico's voor de onderwijskwaliteit. Een parttime dienstverband is een dienstverband van 0,8FTE of minder. Jaar refereert aan het boekjaar

Berekeningswijze: (Aantal FTE onderwijsgevend personeel met een parttime dienstverband/totaal onderwijsgevend personeel van een bestuur)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 34.56: risicoscore 0, als waarde groter of gelijk is aan 34.56: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 3

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.15 *Aandeel tijdelijk onderwijzend personeel (po)*

Beschrijving /definitie: Aandeel tijdelijk onderwijsgevend personeel t.o.v. totaal onderwijsgevend personeel binnen de sector PO. De veronderstelling is dat een groter aandeel tijdelijk onderwijsgevend personeel, kan leiden tot risico's voor de onderwijskwaliteit. Jaar refereert aan het boekjaar

Berekeningswijze: (Aantal FTE tijdelijk onderwijsgevend personeel/Totaal onderwijsgevend personeel in fte's)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 15.43: risicoscore 0, als waarde groter of gelijk is aan 15.43: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.16 *Aandeel tijdelijk onderwijzend personeel (so)*

Beschrijving /definitie: Aandeel tijdelijk onderwijsgevend personeel t.o.v. totaal onderwijsgevend personeel binnen de sector SO. De veronderstelling is dat een groter aandeel tijdelijk onderwijsgevend personeel, kan leiden tot risico's voor de onderwijskwaliteit. Jaar refereert aan het boekjaar

Berekeningswijze: (Aantal FTE tijdelijk onderwijsgevend personeel/Totaal onderwijsgevend personeel in fte's)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 16.97: risicoscore 0, als waarde groter of gelijk is aan 16.97: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.17 *Aandeel tijdelijk onderwijzend personeel (vo)*

Beschrijving /definitie: Aandeel tijdelijk onderwijsgevend personeel t.o.v. totaal onderwijsgevend personeel binnen de sector VO. De veronderstelling is dat een groter aandeel



tijdelijk onderwijsgevend personeel, kan leiden tot risico's voor de onderwijskwaliteit. Jaar refereert aan het boekjaar

Berekeningswijze: (Aantal FTE tijdelijk onderwijsgevend personeel/Totaal onderwijsgevend personeel in fte's)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 23.4: risicoscore 0 , als waarde groter of gelijk is aan 23.4: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 3

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.18 *Aandeel tijdelijk onderwijsgevend personeel (mbo)*

Beschrijving /definitie: Aandeel tijdelijk onderwijsgevend personeel t.o.v. totaal onderwijsgevend personeel binnen de sector MBO. De veronderstelling is dat een groter aandeel tijdelijk onderwijsgevend personeel, kan leiden tot risico's voor de onderwijskwaliteit. Jaar refereert aan het boekjaar

Berekeningswijze: (Aantal FTE tijdelijk onderwijsgevend personeel/Totaal onderwijsgevend personeel in fte's)*100%.

Signaleringswaarde: als waarde kleiner of gelijk is aan 22.08: risicoscore 0 , als waarde groter of gelijk is aan 22.08: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.19 *Percentage ziekteverzuim personeel (vo)*

Beschrijving /definitie: Het gewogen verzuimpercentage van personeel in de sector Voortgezet Onderwijs. De veronderstelling is dat een hoog percentage ziekteverzuim kan leiden tot risico's voor de onderwijskwaliteit, maar ook dat een hoog verzuim een gevolg kan zijn van problemen. Jaar refereert aan het boekjaar

Berekeningswijze: Het gemiddeld verzuimpercentage van het personeel per brin is vermenigvuldigd met het aantal fte's personeel. Dit is nodig om het gewogen gemiddeld verzuim te berekenen per bestuur.

Signaleringswaarde: als waarde kleiner of gelijk is aan 7.2: risicoscore 0 , als waarde groter of gelijk is aan 7.2: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 2

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.20 *Percentage ziekteverzuim personeel (so)*

Beschrijving /definitie: Het gewogen verzuimpercentage van personeel in de sector Speciaal Onderwijs. De veronderstelling is dat een hoog percentage ziekteverzuim kan leiden tot risico's voor de onderwijskwaliteit, maar ook dat een hoog verzuim een gevolg kan zijn van problemen. Jaar refereert aan het boekjaar

Berekeningswijze: Het gemiddeld verzuimpercentage van het personeel per vestiging is vermenigvuldigd met het aantal fte's personeel. Dit is nodig om het gewogen gemiddeld verzuim te berekenen per bestuur.

Signaleringswaarde: als waarde kleiner of gelijk is aan 9.48: risicoscore 0 , als waarde groter of gelijk is aan 9.48: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 3

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.21 *Percentage ziekteverzuim personeel (po)*

Beschrijving /definitie: Het gewogen verzuimpercentage van personeel in de sector Primair Onderwijs. De veronderstelling is dat een hoog percentage ziekteverzuim kan leiden tot risico's voor de onderwijskwaliteit, maar ook dat een hoog verzuim een gevolg kan zijn van problemen. Jaar refereert aan het boekjaar



Berekeningswijze: Het gemiddeld verzuimpercentage van het personeel per vestiging is vermenigvuldigd met het aantal fte's personeel. Dit is nodig om het gewogen gemiddeld verzuim te berekenen per bestuur.

Signaleringswaarde: als waarde kleiner of gelijk is aan 8.13: risicoscore 0, als waarde groter of gelijk is aan 8.13: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 5

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.22 *Gemiddelde leeftijd leraren (po)*

Beschrijving /definitie: Gemiddelde leeftijd onderwijsgevend personeel binnen de sector PO. Jaar refereert aan het boekjaar

Berekeningswijze: Gemiddelde leeftijd onderwijsgevend personeel

Signaleringswaarde: als waarde kleiner of gelijk is aan 46.1: risicoscore 0, als waarde groter of gelijk is aan 46.1: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 3

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.23 *Gemiddelde leeftijd leraren (so)*

Beschrijving /definitie: Gemiddelde leeftijd onderwijsgevend personeel binnen de sector SO. Jaar refereert aan het boekjaar

Berekeningswijze: Gemiddelde leeftijd onderwijsgevend personeel

Signaleringswaarde: als waarde kleiner of gelijk is aan 48.12: risicoscore 0, als waarde groter of gelijk is aan 48.12: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 3

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.24 *Gemiddelde leeftijd leraren (vo)*

Beschrijving /definitie: Gemiddelde leeftijd onderwijsgevend personeel binnen de sector VO. Jaar refereert aan het boekjaar

Berekeningswijze: Gemiddelde leeftijd onderwijsgevend personeel

Signaleringswaarde: als waarde kleiner of gelijk is aan 41.68: risicoscore 1, als waarde groter of gelijk is aan 41.68: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 2

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.6.25 *Gemiddelde leeftijd leraren (mbo)*

Beschrijving /definitie: Gemiddelde leeftijd onderwijsgevend personeel binnen de sector MBO. Jaar refereert aan het boekjaar

Berekeningswijze: Gemiddelde leeftijd onderwijsgevend personeel

Signaleringswaarde: als waarde kleiner of gelijk is aan 48.29: risicoscore 0, als waarde groter of gelijk is aan 48.29: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.7 Resultaten

4.7.1 *Opbrengsten (onvoldoende oordeel)*

Beschrijving /definitie: Laat zien of het bestuur ovt's onder zich heeft met een onvoldoende oordeel op een van de OR standaarden (oordelen vanaf sept 2017).

Berekeningswijze: Is er minstens 1 ovt onder het bestuur met een onvoldoende oordeel op een van de OR standaarden? Ja = 1, nee = 0.



Signaleringswaarde: als waarde kleiner of gelijk is aan 1: risicoscore 0 , als waarde groter of gelijk is aan 1: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: Alleen meest recente gegevens

4.7.2 *Percentage vsv (vo)*

Beschrijving /definitie: Het percentage vroegtijdig schoolverlaters binnen de sector VO van een bestuur.

Berekeningswijze: (Aantal VSV / aantal leerlingen in VSV-gerechtigde leeftijd) * 100%

Signaleringswaarde: als waarde kleiner of gelijk is aan 0.63: risicoscore 0 , als waarde groter of gelijk is aan 0.63: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 2

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.7.3 *Percentage vsv (mbo)*

Beschrijving /definitie: Het percentage vroegtijdig schoolverlaters binnen de sector MBO van een bestuur.

Berekeningswijze: (Aantal VSV / aantal leerlingen in VSV-gerechtigde leeftijd) * 100%

Signaleringswaarde: als waarde kleiner of gelijk is aan 6.81: risicoscore 0 , als waarde groter of gelijk is aan 6.81: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 6

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 3x, 2025: 2x

4.8 **Signalen**

4.8.1 *Risicoscore signalen*

Beschrijving /definitie: Laat zien of er meer signalen over het bestuur zijn binnengekomen dan verwacht zou kunnen worden o.b.v. aantal leerlingen. NB: Leerlingaantallen kunnen iets afwijken van andere indicatoren door gebruik van verschillende databronnen. De aantallen signalen uitgesplitst per kwaliteitsgebied in de aanvullende informatie kan optellen tot een hoger aantal dan het totaal van signalen omdat signalen meerdere kg-labels kunnen krijgen. Signalen op het onderwerp Corona (m.n. schoolsluitingen) zijn uit de dataset verwijderd.

Berekeningswijze: Per sector (per jaar) wordt eerst berekend wat het gemiddeld aantal signalen per leerling is. Vervolgens wordt per bestuur (per sector) berekend hoeveel signalen er verwacht zouden kunnen worden op basis van het aantal leerlingen. De risicoscore geeft weer in welke mate het daadwerkelijk aantal geregistreerde signalen per bestuur hoger is dan verwacht volgens de formule: $(\text{aantal_signalen_bestuur} - (\text{aantal_leerlingen_bestuur} * \text{kans_op_signaal})) / \sqrt{\text{aantal_leerlingen_bestuur} * \text{kans_op_signaal}}$

Signaleringswaarde: als waarde kleiner of gelijk is aan 1.25: risicoscore 0 , als waarde groter of gelijk is aan 1.25: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 8

Weging per jaar: Alleen meest recente gegevens

4.8.2 *Veiligheidsmonitor: percentage ovt's met risico m.b.t. tijdige aanlevering (po)*

Beschrijving /definitie: Schatting van het percentage van de ovt's onder het bestuur dat risico's laat zien m.b.t. het aanleveren van de monitor sociale veiligheid. Dat betekent doorgaans dat een of meerdere ovt's de monitor niet of te laat hebben ingeleverd, op het moment van productie van de Prestatiemonitors op onderliggend ovt-niveau (PM-NK niet meegenomen). Deze indicator is berekend aan de hand van de indicator monitor sociale veiligheid in de prestatimonitors op onderliggend ovt-niveau. Let op: als de monitor na oplevering van de PM op onderliggend ovt-niveau alsnog is aangeleverd kan er dus een verschil optreden tussen bronnen. Het betreft een schatting omdat scholen enige vrijheid hebben in het niveau waarop de monitor wordt aangeleverd (brin, brin-vest, etc).



Berekeningswijze: (Aantal ovt's met een risico op tijdige aanlevering/aantal ovt's) * 100.

Signaleringswaarde: als waarde groter of gelijk is aan 50: risicoscore 1, als waarde kleiner of gelijk is aan 50: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 4

Weging per jaar: Alleen meest recente gegevens

4.8.3 *Veiligheidsmonitor: percentage ovt's met risico m.b.t. tijdige aanlevering (so)*

Beschrijving /definitie: Schatting van het percentage van de ovt's onder het bestuur dat risico's laat zien m.b.t. het aanleveren van de monitor sociale veiligheid. Dat betekent doorgaans dat een of meerdere ovt's de monitor niet of te laat hebben ingeleverd, op het moment van productie van de Prestatiemonitors op onderliggend ovt-niveau. Deze indicator is berekend aan de hand van de indicator monitor sociale veiligheid in de prestatimonitors op onderliggend ovt-niveau. Let op: als de monitor na oplevering van de PM op onderliggend ovt-niveau alsnog is aangeleverd kan er dus een verschil optreden tussen bronnen. Het betreft een schatting omdat scholen enige vrijheid hebben in het niveau waarop de monitor wordt aangeleverd (brin, brin-vest, etc).

Berekeningswijze: (Aantal ovt's met een risico op tijdige aanlevering/aantal ovt's) * 100.

Signaleringswaarde: als waarde groter of gelijk is aan 50: risicoscore 1, als waarde kleiner of gelijk is aan 50: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: Alleen meest recente gegevens

4.8.4 *Veiligheidsmonitor: percentage ovt's met risico m.b.t. tijdige aanlevering (vo)*

Beschrijving /definitie: Schatting van het percentage van de ovt's onder het bestuur dat risico's laat zien m.b.t. het aanleveren van de monitor sociale veiligheid. Dat betekent doorgaans dat een of meerdere ovt's de monitor niet of te laat hebben ingeleverd, op het moment van productie van de Prestatiemonitors op onderliggend ovt-niveau (PM-EOA niet meegenomen). Deze indicator is berekend aan de hand van de indicator monitor sociale veiligheid in de prestatimonitors op onderliggend ovt-niveau. Let op: als de monitor na oplevering van de PM op onderliggend ovt-niveau alsnog is aangeleverd kan er dus een verschil optreden tussen bronnen. Het betreft een schatting omdat scholen enige vrijheid hebben in het niveau waarop de monitor wordt aangeleverd (brin, brin-vest, etc).

Berekeningswijze: (Aantal ovt's met een risico op tijdige aanlevering/aantal ovt's) * 100.

Signaleringswaarde: als waarde groter of gelijk is aan 50: risicoscore 1, als waarde kleiner of gelijk is aan 50: risicoscore 0. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 4

Weging per jaar: Alleen meest recente gegevens

4.8.5 *Percentage schorsingen (po)*

Beschrijving /definitie: Percentage schorsingen (ongeacht duur van de schorsing) t.o.v. het aantal leerlingen binnen de sector PO.

Berekeningswijze: (Totaal aantal schorsingen / totaal aantal leerlingen) * 100

Signaleringswaarde: als waarde kleiner of gelijk is aan 0.07: risicoscore 0, als waarde groter of gelijk is aan 0.07: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 4x, 2025: 0x

4.8.6 *Percentage schorsingen (so)*

Beschrijving /definitie: Percentage schorsingen (ongeacht duur van de schorsing) t.o.v. het aantal leerlingen binnen de sector SO.

Berekeningswijze: (Totaal aantal schorsingen / totaal aantal leerlingen) * 100

Signaleringswaarde: als waarde kleiner of gelijk is aan 1.69: risicoscore 0, als waarde groter of gelijk is aan 1.69: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie



Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 4x, 2025: 0x

4.8.7 *Percentage schorsingen (vo)*

Beschrijving /definitie: Percentage schorsingen (ongeacht duur van de schorsing) t.o.v. het aantal leerlingen binnen de sector VO.

Berekeningswijze: (Totaal aantal schorsingen / totaal aantal leerlingen) * 100

Signaleringswaarde: als waarde kleiner of gelijk is aan 1.11: risicoscore 0, als waarde groter of gelijk is aan 1.11: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 4x, 2025: 0x

4.8.8 *Percentage schorsingen (mbo)*

Beschrijving /definitie: Percentage schorsingen (ongeacht duur van de schorsing) t.o.v. het aantal leerlingen binnen de sector MBO.

Berekeningswijze: (Totaal aantal schorsingen / totaal aantal leerlingen) * 100

Signaleringswaarde: als waarde kleiner of gelijk is aan 0.12: risicoscore 0, als waarde groter of gelijk is aan 0.12: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 4x, 2025: 0x

4.8.9 *Percentage verwijderingen (vo)*

Beschrijving /definitie: Percentage verwijderingen t.o.v. het aantal leerlingen binnen de sector VO.

Berekeningswijze: (Totaal aantal Verwijderingen / totaal aantal leerlingen) * 100

Signaleringswaarde: als waarde kleiner of gelijk is aan 0.1: risicoscore 0, als waarde groter of gelijk is aan 0.1: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 4x, 2025: 0x

4.8.10 *Percentage verwijderingen (mbo)*

Beschrijving /definitie: Percentage verwijderingen t.o.v. het aantal leerlingen binnen de sector MBO.

Berekeningswijze: (Totaal aantal Verwijderingen / totaal aantal leerlingen) * 100

Signaleringswaarde: als waarde kleiner of gelijk is aan 0.02: risicoscore 0, als waarde groter of gelijk is aan 0.02: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: 2020: 1x, 2021: 2x, 2022: 3x, 2023: 4x, 2024: 4x, 2025: 0x

4.8.11 *Niet-melders via ISD (so)*

Beschrijving /definitie: Een of meerdere onderliggende brins heeft/hebben in de afgelopen drie jaar geen enkele melding van schorsing en verwijdering en aanvraag afwijking onderwijstijd via het ISD doorgegeven. Hierbij wordt alleen gekeken naar scholen die een dermate grote omvang hebben dat het zeer waarschijnlijk is dat er een melding zou zijn binnengekomen (vergeleken met andere scholen met vergelijkbare omvang).

Berekeningswijze: De indicator krijgt de waarde 1 als een of meerdere onderliggende brins geen enkele schorsing en verwijdering en aanvraag afwijking onderwijstijd via het ISD is gemeld.

Signaleringswaarde: als waarde kleiner of gelijk is aan 1: risicoscore 0, als waarde groter of gelijk is aan 1: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: Alleen meest recente gegevens



4.8.12 *Niet-melders via ISD (vo)*

Beschrijving /definitie: Een of meerdere onderliggende ovt's heeft/hebben in de afgelopen drie jaar geen enkele melding van schorsing en verwijdering, afwijkende wijze van examineren, artikel 5 van het Eindexamenbesluit en aanvraag afwijking onderwijstijd via het ISD doorgegeven.

Berekeningswijze: De indicator krijgt de waarde 1 als een of meerdere onderliggende ovt's geen enkele schorsing en verwijdering, afwijkende wijze van examineren, artikel 5 examenbesluit en aanvraag afwijking onderwijstijd via het ISD heeft gemeld.

Signaleringswaarde: als waarde kleiner of gelijk is aan 1: risicoscore 0, als waarde groter of gelijk is aan 1: risicoscore 1. Bij grensgevallen geldt de hoogste risicocategorie

Weegfactor indicator: 1

Weging per jaar: Alleen meest recente gegevens



Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

5 Feedback: vragen en opmerkingen

Voor vragen en opmerkingen over deze prestatiemonitor kunt u terecht bij **5.1.2.e** of **5.1.2.e**, **5.1.2.e** @owinsp.nl/ **5.1.2.e** @owinsp.nl.



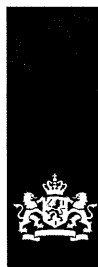
Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

56.

EVALUATIE VAN DE PRESTATIEMONITOR

EVALUATIE VAN DE KWALITEIT VAN DE RISICOSCHATTING EN DE BIAS

Oktober 2023



1 Inhoud

1	Inhoud.....	2
2	Samenvatting.....	3
3	Inleiding.....	4
4	Terugblik aanbevelingen Evaluatierapport 2021.....	6
4.1	Instellen centrale registratie JPRA.....	6
4.2	Inrichten van een JPRA voor besturen.....	6
5	Kwaliteit van de risicoschatting.....	8
5.1	Prestatiemonitors sectoren.....	8
5.2	Prestatiemonitor besturen.....	11
6	Evaluatie van de bias.....	14
6.1	Regio.....	15
6.2	Schoolgrootte.....	16
6.3	Leerlingen uit APC-gebied.....	17
6.4	Denominatie.....	18
6.5	Bestaansduur school, type eindtoets en vmbo-afdeling.....	19
7	Conclusies.....	21
7.1	Kwaliteit van de risicoschatting.....	21
7.2	Bias.....	22
8	Aanbevelingen.....	24



2 Samenvatting

In dit rapport evalueren we de kwaliteit van de risicoschatting en de aanwezigheid van mogelijke biases in de Prestatiemonitors.

Alle Prestatiemonitors lijken het grootste deel van de risicovolle ovt's (oordeel Onvoldoende of Zeer zwak, of ingeschat risico door analist) ook een hoge risicoscore te geven. Het is lastig om een globaal beeld te geven van de kwaliteit van de Prestatiemonitor ten opzichte van vorig jaar; sommige Prestatiemonitors presteren iets beter, andere presteren iets slechter. Ondanks dat we de afgelopen jaren de Prestatiemonitor hebben geprobeerd te verbeteren, zien we over de afgelopen drie jaar geen grote veranderingen in de kwaliteit van de risicoschatting.

Voor de Prestatiemonitor BO hebben we mogelijke oververtegenwoordiging van scholen op basis van regio, schoolgrootte, afgenomen type eindtoets, bestaansduur van de school en denominatie onderzocht. Binnen de groep scholen met een hoge risicoscore zien we een oververtegenwoordiging van scholen in Noord-Nederland en van kleine scholen. Wat betreft het type eindtoets en de bestaansduur van de school zien we kleine maar verklaarbare verschillen. Kijkend naar de denominatie zien we dat alleen openbare scholen vaker een hoge risicoscore krijgen dan op grond van de oordelen verwacht kan worden. Bij de Prestatiemonitor VO is gekeken naar oververtegenwoordigingen op grond van schoolgrootte, percentage leerlingen uit een APC-gebied en vmbo-afdelingen. We zien een oververtegenwoordiging van kleine scholen in de groep scholen met een hoge risicoscore. Wat betreft het percentage leerlingpopulatie uit APC-gebied en aanwezigheid van een vmbo-afdeling zien we een matige oververtegenwoordiging in de hoog-risicogroep, die echter grotendeels herkenbaar is zoals blijkt uit historische inspecteursoordelen.

Om de Prestatiemonitors en de manier van werken er omheen in de toekomst te verbeteren doen we een aantal concrete aanbevelingen: 1) De ontwikkelaars en sectorvertegenwoordigers moeten, gegeven de sterke oververtegenwoordiging van kleine scholen, onderzoeken of het wenselijk is om de Prestatiemonitor hierop aan te passen; 2) De ontwikkelaars moeten samen met vertegenwoordigers uit het toezicht van de Prestatiemonitor de indicatoren uitgebreider gaan documenteren; 3) We moeten de analyses over bias binnen de Prestatiemonitors uitbreiden naar andere Prestatiemonitors, en ook naar de onderwijsresultatenmodellen; 4) Bij de inventarisatie van te onderzoeken biases moeten we verder kijken dan inschattingen door sectorvertegenwoordigers.

Ten slotte kunnen we volgend jaar na de uitvoering van de steekproef KO's een nog beter beeld van de kwaliteit van de risicoschatting presenteren en aan de hand hiervan de Prestatiemonitor verder verbeteren.



3 Inleiding

Elk jaar voert de Inspectie van het Onderwijs (IvO) de jaarlijkse prestatie- en risicoanalyse uit (JPRA). Aan de hand van de JPRA wordt bepaald of er vanuit het oogpunt van risico's toezichtsactiviteiten noodzakelijk zijn. Om de JPRA uit te voeren, maken de analisten en inspecteurs onder andere gebruik van de verschillende Prestatiemonitors.

De Prestatiemonitors zijn informatieproducten waarmee risico's voor objecten van toezicht (hierna: ovt's) in kaart worden gebracht. Elke Prestatiemonitor bestaat uit een set aan risico-indicatoren. Enkele voorbeelden van indicatoren zijn het ziekteverzuim van het personeel, het aantal voortijdige schoolverlaters, en financiële kengetallen zoals solvabiliteit. Indien een indicator een kritieke grens overschrijdt, krijgt een school of bestuur voor de betreffende indicator risicopunten. Uiteindelijk wordt er voor elk object van toezicht een risicoscore berekend: het aantal behaalde risicopunten over alle indicatoren (gewogen naar het belang van de indicator) gedeeld door het maximaal te behalen aantal risicopunten. Op deze manier maakt de Prestatiemonitor een schatting van de risico's bij een object van toezicht.

De Prestatiemonitor ondersteunt hiermee de analisten en inspecteur bij het maken van een risicoprioritering. Omdat analisten en inspecteurs kunnen zien welke indicatoren als risicovol beschouwd worden, kunnen zij ook zien hoe de risicoscore opgebouwd is. Daarmee dragen de Prestatiemonitors bij aan zowel het risicogerichte toezicht, als aan de invulling van proportionaliteit (toezicht op maat), aangezien het handvatten biedt over waar de mogelijke risico's mee samenhangen.

Door de gestandaardiseerde manier waarop de risicoscore berekend wordt zijn de Prestatiemonitor algoritmes. Het is belangrijk om algoritmes te evalueren. De evaluatie van algoritmes laat ons namelijk zien of een algoritme risico's juist inschat. Ook kan uit een evaluatie blijken of er bepaalde groepen of instellingen zijn die vaker een hoge risicoscore krijgen dan andere, terwijl het algoritme niet geïnstrueerd is om dit te doen. Zo is het mogelijk dat bepaalde groepen of instellingen op een ernstige wijze benadeeld worden door een algoritme dat slecht werkt, zoals eerder zichtbaar is geworden bij de Belastingdienst en recent bij DUO¹. Het doel van evaluatie is om mogelijke oververtegenwoordiging van groepen in kaart te brengen. Daarnaast helpt de evaluatie bij het verbeteren van de kwaliteit van de Prestatiemonitor. We krijgen namelijk te zien welke indicatoren goed of slecht samenhangen met inschattingen en oordelen van analisten en inspecteurs en kunnen we de werking van de Prestatiemonitor daarop aanpassen. Mede gelet op de risico's die met algoritmes samenhangen, roepen de Rekenkamer en de Autoriteit Persoonsgegevens en andere (overheids)organisaties op om algoritmes zorgvuldig te evalueren.²

In dit rapport evalueren we de Prestatiemonitor op twee punten: de kwaliteit van de risicoschatting, en de bias. Met de evaluatie van de kwaliteit van de

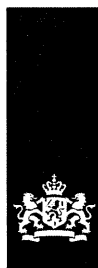
¹ <https://nos.nl/op3/artikel/2479700-studenten-met-migratieachtergrond-opvallend-vaak-peschuldigd-van-fraude-minister-wil-systeem-grondig-nagaan>

² Algemene Rekenkamer. (2022). Algoritmes getoetst. <https://www.rekenkamer.nl/publicaties/rapporten/2022/05/18/algoritmes-getoetst>
Autoriteit Persoonsgegevens. (2023). Rapportage algoritmerisico's Nederland. <https://www.autoriteitpersoonsgegevens.nl/documenten/rapportage-algoritmerisico's-nederland>



risicoschatting geven we antwoord op de vraag: "Hoe schatten analisten en inspecteurs de risico's in bij ovt's met een hoge risicoscore en hoe verhoudt de risicoscore zich tot de oordelen?". Het kan voorkomen dat inspecteurs en analisten risico's bij scholen en besturen met een hoge risicoscore totaal niet herkennen, of juist wel. Het is goed om zicht te krijgen op hoe vaak er overeenstemming is. Met de evaluatie van de bias geven we antwoord op de vraag: "Krijgen bepaalde typen scholen vaker een hoge risicoscore dan andere, wetende dat de Prestatiemonitor niet expliciet geïnstrueerd is om dit te doen?". Het kan mogelijk zijn dat bepaalde kenmerken van scholen sterk samenhangen met de indicatoren die in de Prestatiemonitor zitten. Ook al wordt er dan niet geselecteerd op basis van die kenmerken, de samenhang met de indicatoren zorgt er dan voor dat bepaalde typen scholen vaker een hoge risicoscore krijgen dan andere. Dat kan terecht of onterecht zijn, maar om dat te beoordelen is het eerst nodig om te onderzoeken welke typen scholen vaker een hoge risicoscore krijgen.

In hoofdstuk 4 kijken we naar de aanbevelingen van vorig jaar en bespreken we in hoeverre deze zijn opgevolgd. Hoofdstuk 5 bespreekt de evaluatie van de kwaliteit van de risicoschatting voor zowel de sectorspecifieke Prestatiemonitors, als de Prestatiemonitor Besturen. Hoofdstuk 6 bespreekt de evaluatie van de bias. In hoofdstuk 7 wordt vervolgens een conclusie gepresenteerd. Hoofdstuk 8 presenteert enkele actiepunten om de werking en evaluatie van de Prestatiemonitor in de toekomst te verbeteren.



4 Terugblik aanbevelingen Evaluatierapport 2021

In het evaluatierapport van de Prestatiemonitor van vorig jaar is uitsluitend gekeken naar de kwaliteit van de risicoschatting van de Prestatiemonitor. In dit rapport zijn enkele aanbevelingen gedaan, met als doel de Prestatiemonitor en het gebruik en de evaluatie daarvan te verbeteren. In dit hoofdstuk beschrijven we in hoeverre deze aanbevelingen zijn opgevolgd.

4.1 Instellen centrale registratie JPRA

De eerste aanbeveling had betrekking op het zorgdragen voor een registratie van alle activiteiten die uitgevoerd worden naar aanleiding van de gesignaleerde risico's. Deze aanbeveling werd gedaan omdat er naast risico-onderzoeken ook andere toezichtsactiviteiten bestaan, die veelal ook sectorspecifiek waren. Een besluit tot een lichtere vorm van toezichtsactiviteit, dus anders dan een risico-onderzoek, werd dus niet centraal geregistreerd. Ook werd er niet geregistreerd waarom een risico-onderzoek, waartoe initieel besloten was, uiteindelijk niet is uitgevoerd. Tevens werden risicovolle ovt's die niet vanuit de JPRA naar voren komen niet geregistreerd.

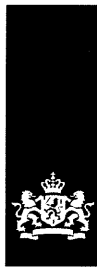
Een jaar later kunnen we constateren dat er veel verbeterd is met betrekking tot deze aanbeveling. We werken aan een centrale registratie van alle uitkomsten uit de JPRA. Om de verschillen tussen de sectoren zo veel mogelijk te stroomlijnen is er een gestandaardiseerd Excel-bestand ontwikkeld waarin alle sectoren, met uitzondering van MBO, het risicoanalyse proces kunnen vastleggen. In dit bestand is voorzien in risicodetectie vanuit meerdere bronnen (Prestatiemonitor, Onderwijsresultatenmodel, en kennis, signalen en info van de contactinspecteur). Ook voorziet dit bestand in een meer gedetailleerdere risico-categorisering, waarin het ook mogelijk is om lichtere toezichtsactiviteiten op te nemen.

Dit Excel-bestand was nog niet beschikbaar tijdens de evaluatie van dit jaar, maar veel wijzigingen waren al wel doorgevoerd in de sectorspecifieke registratiebestanden. Daarom zijn nog niet alle wijzigingen zichtbaar in dit rapport, echter met de invoering van het gestandaardiseerde registratiebestand is dit naar verwachting volgend jaar wel het geval. Over het algemeen kunnen we constateren dat de registratie sterk verbeterd is afgelopen jaar en dat verdient waardering.

4.2 Inrichten van een JPRA voor besturen

Vorig jaar constateerden wij dat veel risicovolle besturen in de bovenste 20% van de Prestatiemonitor terecht kwamen. Het is dus aannemelijk dat het loont om nader naar de groep besturen met de hoogste risicoscore te kijken. De kans lijkt groot dat de onvoldoende besturen oververtegenwoordigd zullen zijn in deze groepen, wat het toezicht gericht en efficiënter kan maken. In de praktijk bestaan er echter verschillen tussen sectoren wat betreft het gebruik van de Prestatiemonitor Besturen.

Naar aanleiding hiervan werd aanbevolen om een verkenning uit te voeren naar de mogelijkheid van een gezamenlijk JPRA-proces op bestuursniveau. In de praktijk blijkt echter dat er het afgelopen jaar onvoldoende capaciteit



Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

beschikbaar is geweest om een dergelijke lijn sectoroverstijgend in te voeren naast de 4JO's (en andere onderzoeken). Op het moment van schrijven van dit rapport loopt er echter een besluitvormingsproces op basis van input van de centrale planningswerkgroep over de verdeling van capaciteit op de korte termijn. Hierin is ook de proportionele inzet op besturen een optie. Mogelijk zullen er daarom op dit vlak op korte termijn ontwikkelingen zijn.



5 Kwaliteit van de risicoschatting

Dit hoofdstuk beantwoordt de vraag hoe analisten en inspecteurs risico's inschatten bij ovt's die een hoge risicoscore hebben gekregen. De groep scholen en besturen met de hoogste risicoscores zijn nader bekeken door analisten en inspecteurs tijdens de JPRA. Hun beoordelingen van deze scholen en besturen zijn vergeleken met de risicoscores die de Prestatiemonitor berekend heeft. Dit hoofdstuk beschrijft dus hoe de uitkomsten van de risicoschatting door analisten en inspecteurs en het huidige oordeel zich verhouden tot de positie van ovt's in de Prestatiemonitor.

Het is goed om te benadrukken dat deze analyse niet onderzoekt of de Prestatiemonitor het oordeel van een inspecteur of de inschatting van een analist kan voorspellen. De analist en inspecteur weten namelijk wat de risicoscore is op het moment dat zij een oordeel vormen, wat hen, bewust of onbewust, kan beïnvloeden. Omdat we dit niet kunnen uitsluiten, kan er niet zuiver van een "voorspelling" gesproken worden.

Twee Prestatiemonitors zijn niet opgenomen in deze analyse. De Prestatiemonitor MBO is niet meegenomen omdat deze uitgefaseerd is. Daarnaast is de Prestatiemonitor VO EOA (eerste opvang anderstaligen) niet opgenomen, aangezien deze Prestatiemonitor pas recent ontwikkeld is en relatief weinig ovt's bevat. In de analyse van volgend jaar zal deze Prestatiemonitor indien mogelijk wel meegenomen worden.

De resultaten van de evaluatie van de kwaliteit van de risicoschatting worden in twee delen besproken. Het eerste deel bespreekt de sectorspecifieke Prestatiemonitors. Het tweede deel bespreekt de resultaten voor de Prestatiemonitor Besturen.

5.1 Prestatiemonitors sectoren

In de jaarlijkse prestatie- en risicoanalyse (JPRA) worden binnen elke sector de ovt's met de hoogste risicoscores nader bekeken door analisten. Het aantal scholen dat nader bekeken wordt verschilt per sector, zowel in absolute als relatieve zin. Ook verschillen de manieren waarop de analisten de risico-inschatting documenteren, alhoewel dat steeds meer gestandaardiseerd wordt.

Figuur 5.1 toont per ovt-Prestatiemonitor de verdeling van de tijdens de JPRA geconstateerde risico's en de oordelen over de risicoscores die de Prestatiemonitor berekend heeft. De ovt's zijn als volgt opgedeeld in groepen, afhankelijk van de uitkomsten van de expertanalyse (EA) en de recente oordelen.

- **EA: Risico's.** De analist heeft tijdens de expertanalyse het object van toezicht bekeken en geconstateerd dat dit ovt risicovol is. Dit kan betekenen dat er een toezichtsactiviteit nodig is, maar dat is niet noodzakelijkerwijs het geval.
- **Oordeel: Onvoldoende of Zeer zwak".** Het ovt heeft recent het oordeel Onvoldoende of Zeer zwak gekregen. Hierbij wordt gekeken naar oordelen die na 1 september 2020 zijn gegeven.



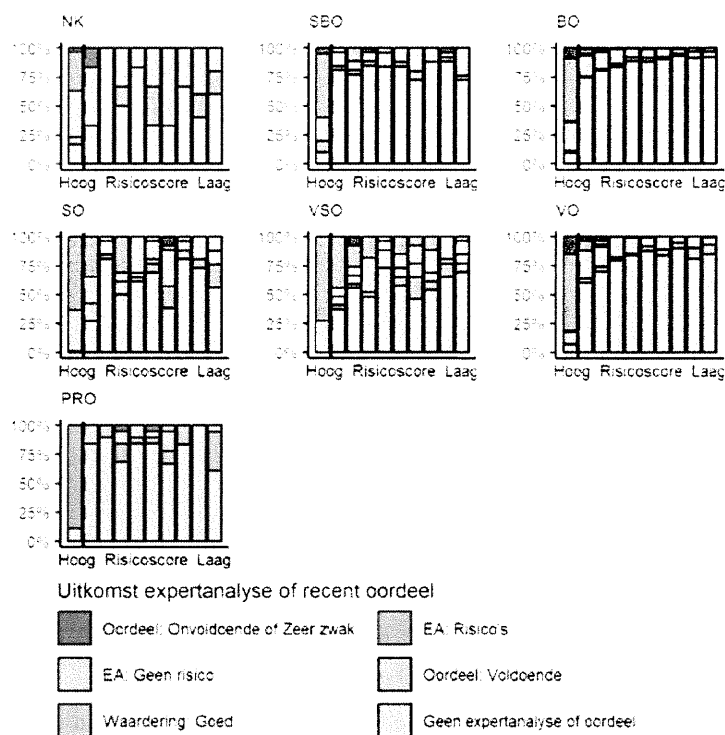
- **EA: Geen risico.** Het ovt is bekeken tijdens de expertanalyse, maar de analist constateert dat er geen risico's zijn.
- **Oordeel: Voldoende.** Het ovt heeft recent het oordeel Voldoende gekregen en is niet nader bekeken tijdens de expertanalyse.
- **Waardering: Goed.** Het ovt heeft recent de waardering Goed gekregen en is niet nader bekeken tijdens de expertanalyse.
- **Geen expertanalyse of recent oordeel.** Het ovt is niet bekeken tijdens de expertanalyse en heeft ook niet recent een oordeel gehad.

Hierbij is het goed om op te merken dat de categorieën mogelijk kunnen overlappen. Zo is het bijvoorbeeld mogelijk dat een ovt dat recent een oordeel Onvoldoende heeft gekregen, tijdens de expertanalyse ook als risicovol wordt beoordeeld. In dergelijke gevallen is er voor gekozen om de expertanalyse zwaarder te laten wegen aangezien dit een meer recente blik op het ovt betreft, en het ovt in de categorie in te delen met de uitkomsten van de expertanalyse. Dit voorbeeld-ovt zal dus geclassificeerd worden als "EA: Risico's" en niet als "Oordeel: Onvoldoende of Zeer zwak".

In Figuur 5.1 zijn de ovt's, gesorteerd op risicoscore, opgedeeld in groepen van gelijke omvang. De ovt's waarbij de Prestatiemonitor de risico's het hoogst inschat staan uiterst links, terwijl de ovt's met de laagste risicoscores uiterst rechts staan. De gestippelde balken geven de grens aan die de betreffende sector heeft gehanteerd bij het nader bekijken van de ovt's.

Bij een Prestatiemonitor die perfect alle risico's meet, zouden alle risicovolle ovt's en ovt's met een recent oordeel Onvoldoende of Zeer zwak uiterst links in het figuur belanden. In de praktijk is het echter niet haalbaar om dit optimale resultaat te realiseren, aangezien niet elk risico meetbaar is.

Figuur 5.1: Risicogroepen in de verschillende Prestatiemonitors afgezet door risico-inschattingen door analisten en inspecteurs





Figuur 5.1: Risicogroepen in de verschillende Prestatiemonitors afgezet door risico-inschattingen door analisten en inspecteurs

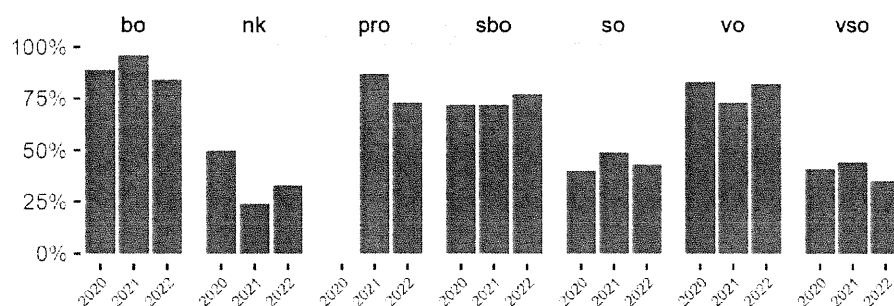
In het algemeen is het beeld dat Figuur 5.1 schetst overwegend positief. Veel van de ovt's die een hoge risicoscore hebben gekregen in de Prestatiemonitor, zijn ook door een analist als risicovol beoordeeld. Wel missen alle Prestatiemonitors nog een deel van de risicovolle ovt's.

Daarnaast geldt voor de meeste Prestatiemonitors ook dat de ovt's die de waardering Goed hebben gekregen juist vaak een relatief lage risicoscore krijgen. Dat is een goed signaal, aangezien het wenselijk is dat de Prestatiemonitor niet alleen risicovolle ovt's een hoge risicoscore geeft, maar ook niet-risicovolle ovt's een lage risicoscore. Bovendien verliep de selectie voor de waardering Goed niet via de JPRA, maar via aanmelding door scholen zelf. In die zin biedt het een enigszins onafhankelijke mogelijkheid tot validatie. Met die gedachte zou het onwenselijk zijn als de ovt's met de waardering Goed willekeurig over de risicoscores verdeeld zouden zijn. In plaats daarvan zien we dat deze ovt's vaak een relatief lage risicoscore krijgen. Vanwege het wegvallen van de waardering Goed zal deze vorm van validatie in de toekomst mogelijk niet meer mogelijk zijn.³

Voor alle Prestatiemonitors is eveneens zichtbaar dat voor een substantieel deel van de ovt's geen recent oordeel of een uitkomst uit de expertanalyse beschikbaar is. Dit toont ook één van de beperkingen van deze analyse: we weten nu dat een substantieel deel van de ovt's met hoge risicoscores ook als risicovol worden beoordeeld door analisten, maar aangezien we minder goed zicht hebben op met name de ovt's met een lage risicoscore, is het lastiger om te beoordelen hoe risicovol de ovt's daar zijn. Met de komst van de steekproef-KO's, waar inmiddels mee gestart is, wordt het mogelijk om daar meer zicht op te krijgen.

De kwaliteit van de risicoschatting van de Prestatiemonitors is vergeleken met de kwaliteit van de risicoschatting van de afgelopen twee jaren. Figuur 5.2 toont het percentage gedetecteerde risico-ovt's voor elke Prestatiemonitor voor de afgelopen drie jaren. Hierbij is in elke Prestatiemonitor gekeken naar de 15% ovt's met de hoogste risicoscores.

Figuur 5.2: Percentage gedetecteerde risico-ovt's met 15% risicogebied voor de verschillende prestatie-monitors



³ Een optie is nog om hiervoor de waarderingen op standaard-niveau te gebruiken.



Figuur 5.2: Percentage gedetecteerde risico-ovt's met 15% risicogebied voor de verschillende prestatie-monitors

Uit Figuur 5.2 blijkt dat de kwaliteit van de risicoschatting voor de meeste Prestatiemonitors redelijk constant is gebleven. Bij de Prestatiemonitor NK, SBO en VO zien we een lichte stijging in het percentage gedetecteerde risico-ovt's ten opzichte van vorig jaar. Het percentage gedetecteerde risico-ovt's in de Prestatiemonitors BO, PRO, SO en VSO is daarentegen lichtelijk gedaald ten opzichte van de Prestatiemonitor uit 2021. Over de drie jaren heen bekeken zien we dat de kwaliteit van de risicoschatting redelijk hetzelfde is gebleven, terwijl er wel pogingen zijn ondernomen om de Prestatiemonitor te verbeteren.

5.2 Prestatiemonitor besturen

De risicoscores van de Prestatiemonitor Besturen worden in dit deel afgezet tegen de recente oordelen die door inspecteurs zijn gegeven tijdens inspectiebezoeken op bestuursniveau. We kijken uitsluitend naar oordelen die gegeven zijn na 1 september 2020.

In onderstaand figuur zijn de besturen opgedeeld in verschillende groepen, afhankelijk van het recente oordeel: Onvoldoende, Voldoende of Goed en Geen oordeel. In de analyse bekijken we of één van de (B)KA-standaarden als onvoldoende beoordeeld is. De sector SO heeft daarnaast ook een risicoschatting van alle besturen gemaakt tijdens de expertanalyse. Dit is eveneens in het figuur meegenomen als de categorieën "Expertanalyse: Risico" en "Expertanalyse: Geen risico".

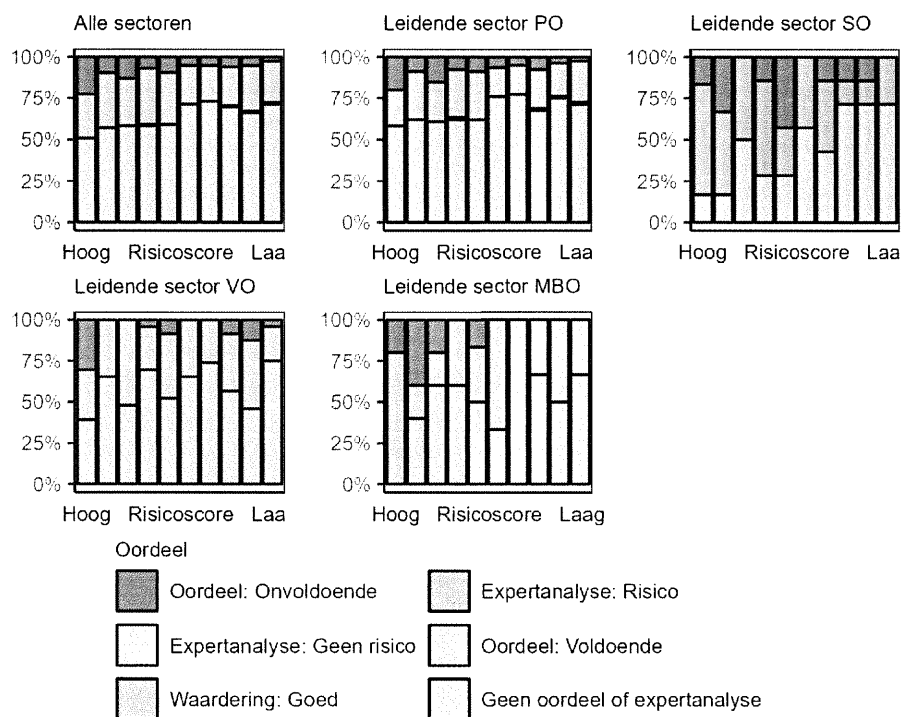
Op dit moment is er nog geen sprake van een sectoroverstijgende werkwijze op bestuursniveau met betrekking tot de expertanalyse. Hierdoor bestaat er ook geen duidelijke risicogrens. Aangezien vorig jaar de risicogrens van 20% is gehanteerd (dit was destijds het percentage dat de sector VO had gebruikt) en om een vergelijking met vorig jaar te kunnen uitvoeren, gebruiken we dit jaar wederom een grens van 20%.

Figuur 5.3 toont de verdeling van de recente oordelen en uitkomsten van de expertanalyse (SO) voor de besturen, uitgesplitst naar risicogroep in de Prestatiemonitor.

Figuur 5.3: Risicogroepen in de verschillende Prestatiemonitors afgezet door risico-inschattingen door analisten en inspecteurs



Figuur 5.3: Risicogroepen in de verschillende Prestatiemonitors afgezet door risico-inschattingen door analisten en inspecteurs



Uit Figuur 5.3 blijkt dat voor de meeste sectoren wel geldt dat het aantal besturen met het oordeel Onvoldoende toeneemt naarmate de risicoscore toeneemt. Ook voor de gehele Prestatiemonitor, dus over alle sectoren beschouwd, is dit het geval.

Er zijn echter wel sterke verschillen tussen de sectoren. Wat betreft de VO-besturen is te zien dat er een substantieel deel van de besturen met een Onvoldoende oordeel een lage risicoscore heeft gekregen.

Alhoewel het aandeel risicovolle besturen wel lijkt toe te nemen naarmate de risicoscore hoger is, is de verdeling veel vlakker dan bij de Prestatiemonitors op ovt-niveau zichtbaar was. Wat ook verschilt ten opzichte van de Prestatiemonitors op ovt-niveau is het feit dat er relatief weinig oordelen beschikbaar zijn voor besturen met een hoge risicoscore, maar relatief veel voor besturen met een lage risicoscore. Dit is het gevolg van het ontbreken van een JPRA en/of expertanalyse voor alle sectoren, met uitzondering van SO.

Vorig jaar viel 58% van de Onvoldoende-besturen in de top 20% besturen met de hoogste risicoscores. Dit jaar is dat 38%. Dit betekent dat de Prestatiemonitor er afgelopen jaar minder goed in geslaagd is om alle besturen met oordeel Onvoldoende een hoge risicoscore te geven. Om te onderzoeken welke risico's de Prestatiemonitor gemist heeft, kijken de ontwikkelaars naar inspectierapporten van besturen die het oordeel Onvoldoende hebben gekregen en tevens een lage risicoscore in de Prestatiemonitor hadden.



Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

Desalniettemin zijn er in de hoogste risicogroepen de meeste besturen met een onvoldoende oordeel te vinden. Hiermee kan de Prestatiemonitor besturen een bijdrage leveren aan het identificeren van risico's op bestuursniveau.



6 Evaluatie van de bias

In 2023 is de bias in de Prestatiemonitor voor het eerst onderzocht. Het is goed om te weten of er bepaalde typen scholen vaker een hoge of lage risicoscore krijgen in de Prestatiemonitor, terwijl de Prestatiemonitor niet expliciet geïnstrueerd is om deze scholen een hogere risicoscore te geven.

De analyse in 2023 is voor het eerst uitgevoerd voor de Prestatiemonitors BO en VO. Deze zijn gekozen omdat dit de grootste sectoren zijn (meeste ovt's per Prestatiemonitor). We zijn van plan om deze analyse volgend jaar voor alle Prestatiemonitors te herhalen. Ook al is de reikwijdte van deze analyse beperkt tot twee van de Prestatiemonitors, de resultaten zijn noemenswaardig genoeg om hier te vermelden. Vervolganalyses zijn nodig om de resultaten nader te interpreteren.

Om te onderzoeken welke typen scholen vaker een hoge risicoscore krijgen, is eerst bij de sectorvertegenwoordigers in de werkgroepen voor de Prestatiemonitors BO en VO geïnventariseerd welke typen scholen zij voor hun gevoel vaak terugzien in de hoog-risicogroepen. Het is bij deze selectie immers belangrijk om uit te gaan van de interessegebieden van de onderwijs-inhoudelijke experts. Aan de hand van deze inventarisatie zijn voor beide Prestatiemonitors enkele aspecten nader onderzocht. De volgende vermoedens van oververtegenwoordiging werden door de sectorvertegenwoordigers geopperd en zijn door ons nader onderzocht:

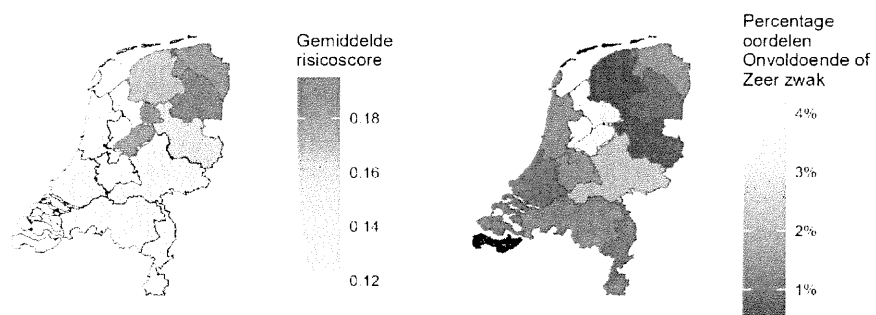
- **Prestatiemonitor BO:**
 - *Regio*: scholen in het noorden van Nederland zouden vaker een hoge risicoscore krijgen dan scholen in het zuiden van Nederland. Mogelijk zou dit ook verklaard kunnen worden door stedelijkheid.
 - *Schoolgrootte*: met name kleine scholen zouden vaker een hoge risicoscore krijgen.
 - *Bestaansduur van de school*: scholen met een relatief korte bestaansduur (korter dan 6 jaar) zouden vaker een hoge risicoscore krijgen dan scholen met een langere bestaansduur
 - *Type eindtoets*: scholen die bepaalde typen eindtoetsen afnemen zouden mogelijk een hogere risicoscore krijgen, omdat zij vanuit de overtuiging dat sommige eindtoetsen moeilijker zijn voor hun type leerlingpopulatie, bepaalde eindtoetsen kiezen.
 - *Denominatie*: risicoscores zouden mogelijk hoger zijn voor scholen van bepaalde denominaties.
- **Prestatiemonitor VO:**
 - *Percentage leerlingen uit armoedeprobleemcumulatiegebied (APCG)*: scholen met een grote populatie leerlingen met een APCG-achtergrond zouden vaker een hoge risicoscore krijgen dan scholen met een lager percentage leerlingen uit APC-gebied.
 - *Schoolgrootte*: met name kleine scholen zouden vaker een hoge risicoscore krijgen.
 - *Vmbo-afdeling*: vmbo-afdelingen zouden mogelijk een hogere risicoscore krijgen dan niet-vmbo-afdelingen.

Voor elke analyse die volgt zijn de ovt's in groepen van gelijke omvang opgedeeld, gesorteerd op risicoscore. In elke visualisatie staan de ovt's waarbij de Prestatiemonitor de risico's het hoogst inschat uiterst links, en de ovt's met

6.1 Regio

In het figuur zijn verschillen tussen de provincies met betrekking tot de risicoscore zichtbaar. In dit figuur is duidelijk te zien dat de gemiddelde risicoscore voor scholen in de provincies Drenthe, Groningen, Flevoland, Friesland en Overijssel hoger is dan de gemiddelde risicoscore voor scholen in de zuidelijke en westelijke provincies. Ook is te zien dat in de zuidelijke provincies de gemiddelde risicoscores juist lager zijn.

Figuur 6.1: Verschillen in gemiddelde risicoscore en percentage Onvoldoende/Zeer zwak van de oordelen tussen provincies



De mogelijke verklaring van het effect van stedelijkheid op de regionale verschillen tussen provincies is daarom onderzocht. Er is een splitsing gemaakt tussen scholen die in stedelijk gebied vallen en scholen die in niet-stedelijk gebied vallen. Daarbij hebben we de CBS-definitie van stedelijkheid gehanteerd. Vervolgens is wederom per provincie de gemiddelde risicoscore berekend. Uit deze vervolganalyse blijkt dat de verschillen tussen provincies blijven bestaan, ook nadat we rekening hebben gehouden met de verschillen in stedelijkheid.

Pagina 15 van 25



betreffende regio. Ook effecten die samenhangen met krimpregio's kunnen een verklaring zijn voor de verschillen in risicoscores tussen de provincies. Om de relatief hoge risicoscores in noordelijke provincies beter te kunnen begrijpen, zou vervolgonderzoek uitgevoerd moeten worden waarin de bovengenoemde factoren meegenomen worden.

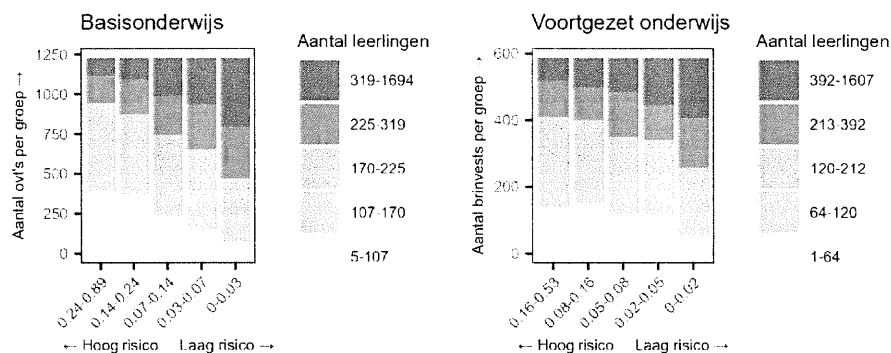
Tevens is onderzocht of er in de noordelijke provincies ook sprake is van relatief veel oordelen Onvoldoende of Zeer zwak. Dit zou suggereren dat de gemiddeld hoge risicoscores aansluiten bij verschillen in kwaliteit die inspecteurs observeren bij inspectie-onderzoeken. In het rechter paneel in Figuur 6.1 wordt het percentage van de gegeven oordelen in elke provincie dat Onvoldoende of Zeer zwak is gevisualiseerd. Met name de provincie Flevoland springt daar uit: daar is circa 4% van alle oordelen sinds 2017 onvoldoende. De noordelijke provincies waar we eerder relatief hoge risicoscores geconstateerd hebben, hebben gemiddeld gezien niet vaker een oordeel Onvoldoende of Zeer zwak. Dit verschil vraagt om verdere reflectie.

6.2 Schoolgrootte

De sectorvertegenwoordigers in zowel de werkgroep Prestatiemonitor BO als in de werkgroep Prestatiemonitor VO vermoedden dat kleinere scholen vaker een hoge risicoscore krijgen. Voor zowel de Prestatiemonitor BO als de Prestatiemonitor VO zijn de verdeling van de risicoscores afgezet tegen de grootte van de leerlingpopulatie. Voor zowel BO als VO is duidelijk te zien dat in de groep scholen met de hoogste risicoscores in de Prestatiemonitor er relatief veel scholen zijn met een kleine leerlingpopulatie. Over het algemeen lijkt hier te gelden: hoe hoger de risicoscore, hoe kleiner de leerlingpopulatie op de betreffende school.

Veel indicatoren in de Prestatiemonitor zijn percentages. Bij de berekening van percentages is het in het geval van kleine scholen al snel zo dat een percentage relatief hoog of laag is. De indicatoren in de Prestatiemonitor gaan af bij indicatoren met een relatief hoge of lage waarde, afhankelijk van wat de indicator probeert te meten. Kleine scholen hebben vaak een relatief hoge of lage waarde en daarom gaan er bij kleine scholen vaker indicatoren af.

Figuur 6.3: Schoolgrootte afgezet tegen risicoscore uit de Prestatiemonitor



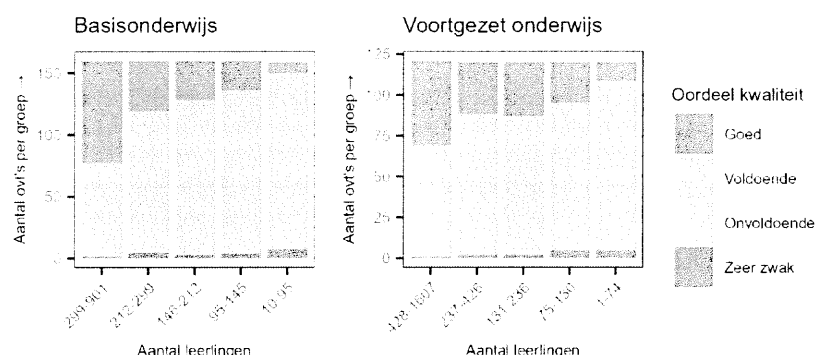
Om de oververtegenwoordiging van kleine scholen onder de scholen met hoge risicoscores te verklaren, is eveneens gekeken naar de relatie tussen oordelen en schoolgrootte. In Figuur 6.4 zijn de inspecteursoordelen afzet tegen de



schoolgrootte. Hierbij is het goed om op te merken dat de aantallen kleiner zijn dan in bovenstaand figuur. Dit komt omdat in Figuur 6.4 uitsluitend gekeken wordt naar inspecteursoordelen sinds augustus 2017, toen het Onderzoekskader 2017 in werking trad. Scholen die sindsdien geen oordeel hebben gekregen, zijn in dit figuur daarom niet meegenomen.

Uit Figuur 6.4 blijkt dat kleine scholen iets vaker het oordeel Onvoldoende of Zeer zwak krijgen dan grote scholen, echter niet in verhoudingen die verwacht worden op basis van de verdeling van de risicoscores in Figuur 6.3. De oververtegenwoordiging van kleine scholen in de hoog-risicogroepen wordt dus maar in beperkte mate verklaard door het feit dat kleine scholen iets vaker het oordeel Onvoldoende of Zeer zwak krijgen. Daarentegen is wel zichtbaar dat relatief grote scholen vaker de waardering Goed krijgen en dat zij daardoor relatief minder vaak het oordeel Voldoende krijgen.

Figuur 6.4: Oordelen afgezet tegen schoolgrootte



Op dit moment wordt in het overleg met ontwikkelaars van de Prestatiemonitors gesproken over mogelijke oplossingen voor deze oververtegenwoordiging. Er zijn meerdere opties die de oververtegenwoordiging van kleine scholen in de hoog-risicogroep kunnen verminderen, of daar meer rekenschap van kunnen geven, bijvoorbeeld scholing over de oververtegenwoordiging, of het toepassen van een statistische correctie. Elk van deze opties heeft voor- en nadelen, die tegen elkaar afgewogen dienen te worden.

6.3

Leerlingen uit APC-gebied

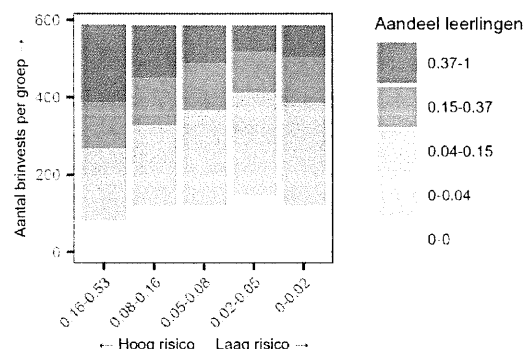
Op advies van de sector VO is eveneens gekeken naar een mogelijke oververtegenwoordiging van scholen met een relatief groot aandeel van de leerlingpopulatie afkomstig uit een armoedeprobleemcumulatiegebied (APC-gebied). Dit is een maat die het CBS opstelt in opdracht van het Ministerie van OCW. In een APC-gebied is er sprake van een relatief hoog aandeel huishoudens met een inkomen onder de lage inkomensgrens, een relatief hoog aandeel huishoudens met een uitkering en een relatief hoog aandeel huishoudens met een hoofdkostwinner met een niet-westerse migratieachtergrond. Bij het identificeren van APC-gebieden sluit het CBS studentenhuishoudens uit.

In Figuur 6.5 is de verdeling van het aandeel leerlingen uit APC-gebied over de risicogroepen zichtbaar. In dit figuur is zichtbaar dat er relatief veel scholen met een relatief hoog percentage leerlingen uit APC-gebied in de hoogste risicogroep vallen. Met uitzondering van de risicogroep met de hoogste risicoscores is het



aandeel scholen met geen tot zeer weinig (de groep 0 tot 0; deze bevat ook scholen met leerlingen uit APC-gebied, echter is dit aandeel zo laag dat dit in de afronding wegvalt) relatief stabiel over de verschillende risicogroepen. We zien echter wel dat het aandeel scholen waar 37% tot 100% van de leerlingpopulatie uit een APC-gebied afkomstig is, toeneemt naarmate de risicoscore hoger wordt.

Figuur 6.5: Aandeel leerlingen afkomstig uit APC-gebied afgezet tegen risicoscore uit de Prestatiemonitor



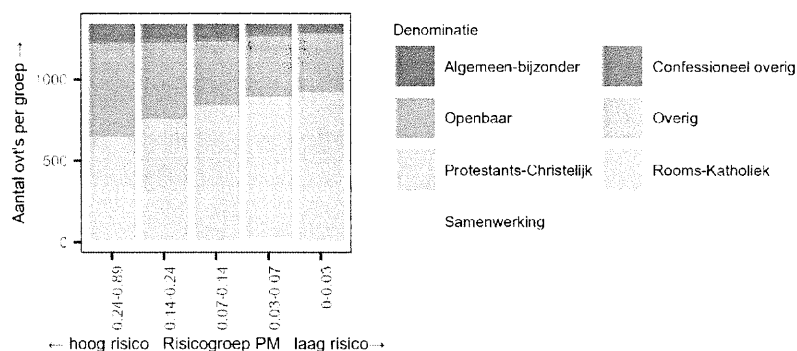
Om deze verschillen te verklaren is eveneens gekeken naar de samenhang tussen de oordelen en het percentage leerlingen uit APC-gebied (niet afgebeeld). Hieruit bleek dat scholen met het oordeel Onvoldoende relatief vaak een leerlingpopulatie hebben waar de meerderheid van de leerlingen uit APC-gebied komt. Voor de scholen met het oordeel Zeer zwak lijkt dit in mindere mate te gelden. De verschillen in percentages leerlingen afkomstig uit APC-gebied tussen de verschillende oordelen verklaart mogelijk ook de verschillen in de Prestatiemonitor.

6.4

Denominatie

Bij de Prestatiemonitor BO is ook de oververtegenwoordiging van bepaalde denominaties onderzocht. Figuur 6.6 toont de verdeling van de verschillende denominaties over de risicoscores uit de Prestatiemonitor. In dit figuur is zichtbaar dat de groep scholen met de hoogste risicoscores voornamelijk scholen met een algemeen-bijzondere of openbare denominatie zijn. Tevens is in dit figuur te zien dat de Protestants-Christelijke en Rooms-Katholieke scholen relatief vaker een lage risicoscore hebben.

Figuur 6.6: Verdeling van de verschillende denominaties over de risicoscores

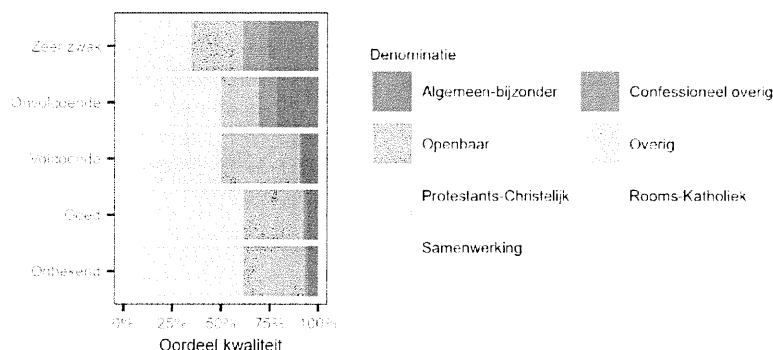




Figuur 6.6: Verdeling van de verschillende denominaties over de risicoscores

Om de oververtegenwoordiging van algemeen-bijzondere en openbare scholen in de hoog-risicogroep te verklaren is gekeken naar de verdeling van denominatie over de oordelen. Figuur 6.7 toont deze verdeling. Kijkend naar de oordelen valt op dat relatief veel algemeen-bijzondere scholen het oordeel Onvoldoende en Zeer zwak hebben. Scholen met een Protestants-Christelijke of Rooms-Katholieke achtergrond daarentegen hebben relatief vaak de waardering Goed. Beide komen overeen met het beeld dat zichtbaar was bij de verdeling van de risicoscores in Figuur 6.6. De verdeling van de openbare scholen over de oordelen daarentegen wijkt wel af. Waar openbare scholen relatief vaak een hoge risicoscore krijgen, zien we juist dat deze scholen relatief vaak het oordeel Voldoende krijgen.

Figuur 6.7: Verdeling van de verschillende denominaties over de oordelen



6.5

Bestaansduur school, type eindtoets en vmbo-afdeling

Op verzoek van de sectoren zijn ook oververtegenwoordiging op grond van bestaansduur van de school (PO), het afgenomen type eindtoets (PO) en vmbo-afdeling (VO) onderzocht.

Met betrekking tot de bestaansduur van de school zien we dat scholen met een korte bestaansduur vaker een hoge risicoscore krijgen dan scholen met een lange bestaansduur. Uit een analyse waarin bestaansduur van de school werd vergeleken met de inspectieoordelen, zagen we dat het oordeel Zeer zwak relatief vaak voorkomt onder scholen met een bestaansduur korter dan 5 jaar. Het oordeel onvoldoende komt daarentegen relatief vaak voor bij scholen die tussen de 5 en 10 jaar bestaan. De waardering Goed komt zelden voor bij scholen met een bestaansduur korter dan 5 jaar. Het beeld wat, kijkend naar de oordelen, naar voren komt is daarmee redelijk in overeenstemming met het beeld wat we bij de risicoscores waarnemen.

Het aandeel van elk type eindtoets lijkt niet aanzienlijk te verschillen tussen scholen met een relatief hoge dan wel lage risicoscore in de Prestatiemonitor. Deze verschillen blijken ook statistisch niet significant te zijn. Om deze reden is de samenhang van het afgenomen type eindtoets met de oordelen niet nader onderzocht en concluderen we dat er geen sprake is van oververtegenwoordiging onder de hoge risicoscores in de Prestatiemonitor op basis van het afgenomen type eindtoets.



Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

Kijkend naar de gemiddelde risicoscore van vmbo-afdelingen, blijkt deze maar marginaal hoger te zijn dan de gemiddelde risicoscore van niet-vmbo-afdelingen. Er lijkt slechts in beperkte mate sprake te zijn van een hoger aandeel vmbo-afdelingen in de hoogste risicogroepen. Omdat de verschillen erg gering zijn, is de oververtegenwoordiging op basis van vmbo-afdeling niet nader onderzocht.



7 Conclusies

In deze analyse is gekeken naar de kwaliteit van de risicoschatting en de biases in de Prestatiemonitor.

7.1 Kwaliteit van de risicoschatting

De meeste Prestatiemonitors slagen er in om een aanzienlijk deel van de ovt's te detecteren waar het oordeel Onvoldoende/Zeer zwak gegeven is, of waar een analist het ovt als risicovol heeft geïdentificeerd. Met betrekking tot de ontwikkeling van de kwaliteit van de risicoschatting ten opzichte van voorgaande jaren kunnen we concluderen dat het beeld wisselend is. Voor enkele sectorspecifieke Prestatiemonitors geldt dat de kwaliteit van de risicoschatting is toegenomen. Bij de Prestatiemonitor NK, SBO en VO zien we een lichte stijging in het percentage gedetecteerde risico-ovt's ten opzichte van vorig jaar. Voor andere Prestatiemonitors is het beeld minder positief. Het percentage gedetecteerde risico-ovt's in de Prestatiemonitors BO, PRO, SO en VSO is daarentegen lichtelijk gedaald. Over de langere termijn is het beeld constanter en zijn er geen grote verschillen zichtbaar. Ondanks pogingen om de Prestatiemonitor te verbeteren, zien we over tijd geen duidelijke verbetering.

De Prestatiemonitor Besturen is dit jaar in mindere mate in staat om de besturen met het oordeel Onvoldoende een hoge risicoscore te geven. Om de risicoschatting volgend jaar te verbeteren, hebben de ontwikkelaars van de Prestatiemonitor Besturen gekeken naar de inspectierapporten van de besturen met het oordeel Onvoldoende die een relatief lage risicoscore hebben gekregen. Hiermee wordt geprobeerd te achterhalen welke risico's door de Prestatiemonitor onopgemerkt zijn gebleven. Ook voor de andere Prestatiemonitors geldt dat het uitzoeken van gemiste risico's in inspectierapporten een zinvolle exercitie is voor de ontwikkelaars van de Prestatiemonitors om de risicoschatting te verbeteren.

In de JPRA bekijken analisten ovt's die een hoge risicoscore hebben nader. Zonder blik op de gehele populatie van ovt's kan dit echter leiden tot blinde vlekken, met name bij de ovt's die de Prestatiemonitor als niet-risicovol inschat. Om dit te onderzoeken, kunnen we in de toekomst gebruik maken van de uitkomsten van de Steekproef KO's, waarvan de eerste al 2023 zijn uitgevoerd. De uitkomsten van deze KO's geven een beter beeld over de hele linie, in aanvulling op de ovt's die de Prestatiemonitor risicovol acht.

Evenals vorig jaar benadrukken we dat het belangrijk blijft om naar ruimte voor verbetering voor elke Prestatiemonitor te blijven zoeken. Dit kan bijvoorbeeld door bij niet onderzochte ovt's te achterhalen welke risico's er gemist zijn. Ook kunnen de eerste resultaten van de Steekproef KO's helpen bij een betere risico-inschatting. Ook biedt een pilot met machine learning aanknopingspunten om de Prestatiemonitors te verbeteren. Een eerste pilot met machine learning algoritmen toont aan dat de kwaliteit van de risicoschatting van de indicatoren in de Prestatiemonitor BO slechts marginaal beter is dan de kwaliteit van de risicoschatting van het onderwijsresultatenmodel. Dit laat zien dat indicatoren in de Prestatiemonitor die op het onderwijsresultatenmodel gebaseerd zijn sterk voorspellend zijn. Op dit moment wordt het gebruik van machine learning verkend om hiermee de Prestatiemonitors verder te verbeteren, maar machine learning-algoritmen worden niet gebruikt in het toezicht. Ten slotte is de



verwachting dat het voor een grote verbetering in de kwaliteit van de risicoschatting waarschijnlijk noodzakelijk is om ook nieuwe informatiebronnen aan te boren die gebruikt kunnen worden om het risicogerichte toezicht effectiever te maken. Binnen het STO wordt in dit kader momenteel een project uitgevoerd.

7.2

Bias

Dit jaar beschrijven we ook voor de eerste keer de mogelijke biases in de Prestatiemonitors die onderzocht zijn. We hebben deze analyses beperkt tot de Prestatiemonitor BO en de Prestatiemonitor VO. In komende jaren wordt dit een structureel onderdeel van dit rapport om waakzaam te blijven voor vooringenomenheid die in de Prestatiemonitor mogelijk aanwezig is. We nemen dan ook de andere Prestatiemonitors mee.

Voor de bestaansduur van de school, het afgenomen type eindtoets, het percentage leerlingen uit APC-gebied en de aanwezigheid van een vmbo-afdeling zagen we wel oververtegenwoordiging in enkele risicogroepen. Omdat deze verschillen gering zijn, of omdat deze verschillen overeenkomen met een oververtegenwoordiging in de gegeven oordelen, wordt dit niet als zorgwekkend gezien met betrekking tot de werking van de Prestatiemonitor. Wel bestaat daarbij het risico van een *self-fulfilling prophecy* (onvoldoendes worden immers alleen daar afgegeven waar we inspecties uitvoeren). Maar omdat de Prestatiemonitors nog niet zo lang in gebruik zijn achten we dat risico op dit moment beperkt. Ook is dat risico beperkt omdat we tot enkele jaren geleden elke 4 jaar op een school kwamen. De SKO's (die uitgaan van willekeurige bezoeken) zullen ook op dit onderwerp vanaf de volgende evaluatie een zeer waardevolle bijdrage leveren aan de analyse.

Kijkend naar de denominatie van scholen, zien we dat er geen grote verschillen zijn in de verdeling van de risicoscores ten opzichte van ten opzichte van hetgeen we zien bij de oordelen. De uitzondering hierop zijn de openbare scholen, die relatief vaak een hoge risicoscore krijgen, terwijl zij niet vaker het oordeel Onvoldoende of Zeer zwak krijgen. Waar deze oververtegenwoordiging van openbare scholen in de hoog-risicogroep door komt, moeten we in vervolgonderzoek uitzoeken.

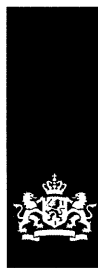
Twee andere vormen van afwijking komen minder duidelijk overeen met patronen in de oordelen (schoolgrootte en regio) en zijn daarom belangrijk om onder de aandacht te brengen. Met betrekking tot de schoolgrootte constateren we dat met name kleine scholen een hoge risicoscore krijgen. Dit is onder meer het resultaat van de manier waarop veel indicatoren in de Prestatiemonitor berekend worden, namelijk aan de hand van een grenswaarde waarbij kleine scholen al snel boven deze grenswaarde vallen. Dit probleem is reeds kenbaar gemaakt aan de ontwikkelaars en werkgroepen van de onderzochte Prestatiemonitors, zodat hiermee rekening gehouden kan worden. Zo is er inmiddels al een indicator uit de Prestatiemonitor BO verwijderd die sterk bijdroeg aan deze oververtegenwoordiging. Ook wordt er dit jaar in de JPRA van PO voor de 10% kleinste scholen een andere grens aangehouden dan voor andere scholen, om het aandeel kleine scholen in de JPRA te beperken. Er wordt daarnaast gewerkt aan een mogelijke technische oplossing in de Prestatiemonitors, waarbij we onder andere onderzoeken of er statistisch rekening gehouden kan worden met de grotere onzekerheid bij indicatoren bij kleine scholen. In algemene zin is het wenselijk om een discussie te houden hoe om te gaan met de grotere onzekerheid bij indicatoren van kleine scholen.



Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

Met betrekking tot de verschillen die tussen regio's geconstateerd worden, is het noodzakelijk om aanvullend onderzoek te doen om deze oververtegenwoordiging te verklaren. Zoals reeds in het rapport beschreven, zijn er een aantal aanknopingspunten waarin een mogelijke verklaring gevonden kan worden, echter kon dit niet voor het verschijnen van dit rapport meegenomen worden. Om de relatief hoge risicoscores in Noord-Nederland nader te verklaren, is vervolgonderzoek nodig.

Tenslotte was het belangrijkste doel van de evaluatie van bias in deze eerste evaluatie om oververtegenwoordiging van groepen scholen inzichtelijk en daarmee bespreekbaar maken. We hopen dat deze analyses aanleiding geven tot gesprek hierover tussen de ontwikkelaars en de sectoren.



8 Aanbevelingen

Ten opzichte van vorige rapportage over de kwaliteit van de risicoschatting van de Prestatiemonitor is er veel verbeterd. In het bijzonder is de registratie van de risicoschatting door analisten en de daaruit voortvloeiende toezichtsactiviteiten sterk verbeterd, wat een rijkere evaluatie oplevert. Ook is er veel voorbereidend werk geleverd voor de Steekproef KO's, die in de toekomst het zicht op met name ovt's met een lage risicoscore zal verbeteren.

Naar aanleiding van de evaluatie van de Prestatiemonitor doen we de volgende aanbevelingen:

- De evaluatie van de bias toont aan dat kleine scholen relatief vaak een hoge risicoscore krijgen. Dit is te verklaren aan de hand van de werking van de Prestatiemonitor. Aangezien kleine scholen relatief weinig leerlingen en personeelsleden hebben, zijn indicatoren die als percentages gescoord zijn al snel heel hoog of heel laag. Het is nodig om een visie te ontwikkelen hoe de inspectie met grotere onzekerheid bij indicatoren van kleine scholen moet omgaan. Is het terecht dat kleine scholen vaker een hoge risicoscore krijgen omdat de indicatoren bij kleine scholen vaker extreme waarden hebben? Enerzijds kunnen we beargumenteren dat toezicht op kleine scholen even belangrijk is als toezicht op grotere scholen. Anderzijds besteedt de inspectie mogelijk veel capaciteit aan kleine scholen (waar de risicoschatting onzekerder is), wat ten koste gaat van toezicht op grotere scholen. De ontwikkelaars moeten samen met het toezicht hier nader op reflecteren, beoordelen in hoeverre het nodig is om de Prestatiemonitor hierop aan te passen en zo nodig dit doorvoeren.
- De indicatoren die in de Prestatiemonitor opgenomen zijn, moeten nog beter onderbouwd worden. Momenteel worden de redenen waarom een indicator in een Prestatiemonitor worden opgenomen wel vastgelegd, echter is deze soms beperkt, en ontbreekt het aan een centrale registratie waarin de onderbouwing van elke indicator is opgenomen. Op deze manier kunnen de indicatoren beter inhoudelijk geduid worden. De onderbouwing van de indicatoren is een gedeelde taak voor de ontwikkelaars van de Prestatiemonitors en de sectorvertegenwoordigers.
- Veel indicatoren die in de Prestatiemonitor opgenomen zijn, hebben raakvlakken met de onderwijsresultatenmodellen. Het is daarom wenselijk om het onderwijsresultatenmodel op eenzelfde manier op biases en vooringenomenheid te evalueren als wij in dit rapport hebben gedaan. Daarbij moet wel opgemerkt worden dat het onderwijsresultatenmodel gegrond is in wettelijk vastgestelde regelingen. Hierover gaat de werkgroep Prestatiemonitors in overleg met de betrokkenen van het onderwijsresultatenmodel.
- Dit jaar is de evaluatie van bias voor het eerst uitgevoerd voor de Prestatiemonitor BO en de Prestatiemonitor VO. De onderzochte mogelijke vormen van vooringenomenheid zijn geïnventariseerd door middel van een uitvraag bij sectorvertegenwoordigers. Voor de analyse van komend jaar raden wij aan om breder te kijken dan de suggesties van sectorvertegenwoordigers, om te voorkomen dat mogelijke vormen van vooringenomenheid gemist worden. Denk hierbij aan biases die veelvuldig die nieuwsmedia halen, of die expliciet vermeld worden in diverse tools voor verantwoorde algoritmes. De *Handreiking non-discriminatie by design* van het Ministerie van Binnenlandse Zaken en het



Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

Toetsingskader Algoritmes van de Algemene Rekenkamer bieden hier handvatten.

Naast de aanbevelingen die we hierboven doen, worden er reeds diverse acties ondernomen om de kwaliteit van de Prestatiemonitor te verbeteren. Zo wordt er bij het updaten van de Prestatiemonitors inspiratie opgedaan uit andere bronnen. Zo kunnen inspectierapporten van Onvoldoende/Zeers zwakke ovt's met een lage risicoscore een indicatie geven van gemiste risico's. Tevens loopt er op dit moment een pilot met machine learning, waar gekeken wordt in hoeverre het mogelijk is om aan de hand van machine learning technieken de kwaliteit van de risicoschatting te verbeteren. Tenslotte kan er winst behaald worden door het aanboren van nieuwe informatiebronnen, al zal dat wel enige tijd kosten. Dergelijke bronnen en projecten kunnen aanknopingspunten bieden om de kwaliteit van de Prestatiemonitor verder te verbeteren. Ontwikkelaars van de individuele Prestatiemonitors kunnen hiermee aan de slag voor de Prestatiemonitor waar zij aan werken.