



# Udvikling af talblindheds- test til 4. klasse og pædagogiske indsatser

Slutrapport  
Februar 2022 (rev marts 2023)

# Indholdsfortegnelse

1.	Indledning.....	1
2.	Grundlaget for projektet.....	2
2.1	Testbatteriet.....	2
2.2	Projektets organisering .....	3
2.3	Tidligere rapporter i projektet .....	4
2.4	Projektets tidsplan og datagrundlag.....	4
3.	Definition af talblindhed .....	6
3.1	Omfang af talblindhed i Danmark.....	6
3.2	Udvikling af numerisk kognition - en model i 4 trin .....	8
3.3	Identifikation af talblindhed og ordblindhed .....	8
3.4	Diagnosticering af talblindhed og målingers validitet.....	9
3.5	Matematikvanskeligheder og talblindhed .....	10
4.	Ekspertgruppen .....	13
5.	Rekruttering af skoler til afprøvning.....	15
5.1	Tilmelding til projektets aktiviteter .....	15
5.2	Rekruttering af kerneskolers og reserveskolers .....	16
6.	Tragtmodellen .....	18
6.1	Det oprindelige design: Tragtmodellens fire trin .....	18
6.2	Den praktiske udformning af afprøvningen.....	18
6.3	Trin 1: Observationsguide.....	19
6.3.1	Scoring og feedback .....	20
6.4	Trin 2: Den digitale test.....	20
6.4.1	Opsætning af den digitale test .....	20
6.4.2	Opsætning til opgaveafprøvning .....	21
6.4.3	Opsætning til demotest.....	21
6.4.4	Digital test, den egentlige afprøvning, runde 1 og 2 .....	22
6.4.5	Motivation og selvtillid .....	22
6.5	Trin 3: Samtaletest.....	23
6.5.1	Opsætning af samtaletesten .....	23
6.5.2	Rapportering af resultater .....	24
7.	Opgaveafprøvning – digital test.....	26
7.1	Opgaveafprøvning på pilotskoler .....	26
7.2	Analyser af data fra opgaveafprøvning .....	26
7.3	To scoringsmodeller for hver skala .....	29

7.4	Raschanalyser – valg af opgaver til digital test.....	29
7.4.1	Skala 1: Sammenligning af mængder på tid.....	29
7.4.2	Skala 2: Sammenligning af talsymboler .....	33
7.4.3	Skala 3: Identifikation af små mængder (Subitizing) .....	35
7.4.4	Skala 4: Sammenligning af tal og mængder.....	37
7.4.5	Skala 5: Addition af tal/mængder .....	38
7.5	Udvælgelsen af items og scoringsmodel til runde 1 og 2 .....	39
8.	Resultater og validering.....	41
8.1	Observationsguide – tjekliste og it-baseret udgave .....	41
8.1.1	Grundlag for analyse af responser .....	41
8.1.2	Raschanalyse af data fra observationsguiden.....	43
8.1.3	To dimensioner i observationsguiden.....	43
8.2	Den digitale test .....	44
8.2.1	Præsentation af resultater og læsevejledning til lærerne.....	45
8.2.2	Resultater fra de to afprøvningsrunder .....	46
8.2.3	Motivation og selvtillid .....	48
8.3	Samtaletesten.....	49
8.3.1	Analyse af data .....	49
8.3.4	Kriterievalidering med RD4 .....	52
8.4	Samspillet mellem den digitale test og samtaletesten .....	55
8.4.1	Fordele og ulemper ved hver type af test .....	56
8.4.2	Sammenhæng mellem resultater i digital test og samtaletest.....	56
8.4.3	Estimation af Cutscores .....	57
9.	Samarbejde med kerneskolere .....	59
9.1	Udvikling og afprøvning af pædagogiske indsatser .....	59
9.1.1	Kortlægning .....	60
9.1.2	Tilbage melding af resultater og udvælgelse af elever .....	61
9.1.3	Udvikling og afprøvning af indsatser .....	64
9.2	Erfaringer fra samarbejdet med kerneskolerne .....	69
9.2.1	Kan man undervise sig ud af talblindhed?.....	69
9.2.2	Koordinatorernes erfaringer .....	70
9.3	Kerneskolernes afprøvning af talblindhedstesten.....	73
9.3.1	De første tre trin i tragtmodellen.....	73
9.3.2	PPR i tragtmodellen .....	74
10.	Spørgeskemaundersøgelse til lærere .....	76
10.1	Lærernes baggrund og rolle på skolen.....	76
10.2	Skolens deltagelse i projektets aktiviteter.....	76

10.3	Vurdering af projektets aktiviteter og testresultater .....	78
10.3.1	Lærernes erfaringer med observationsguiden.....	78
10.3.2	Lærernes erfaringer med den digitale test.....	80
10.3.3	Lærernes erfaringer med samtaletesten .....	81
10.3.4	Lærernes vurdering af talblindhedstesten .....	82
11.	Effektvurdering .....	84
11.1	Undersøgelser- og kontrolgrupper.....	85
11.2	Effektvurdering med den digitale test.....	85
11.3	Effekten på motivation og selvtillid .....	87
12.	Konklusion.....	88
12.1	Tragtmodellens elementer .....	88
12.1.1	Observationsguiden.....	88
12.1.3	Samtaletesten.....	89
12.1.4	PPR.....	90
12.2	Pædagogiske indsatser .....	90
12.2.1	Samarbejdet med kerneskolerne .....	90
12.2.2	Effekten af pædagogiske indsatser .....	90
13.	Anbefalinger .....	92
13.1	Anbefalinger om talblindhedstestens status .....	92
13.2	Anbefalinger til videre udvikling af testens elementer.....	92
13.3	Anbefalinger til øvrige tiltag.....	93
13.4	Anbefalinger til udførelsen af en eventuel fase 2.....	94
13.4.1	Overordnede anbefalinger vedr. fase 2.....	95
13.4.2	Detaljerede anbefalinger vedr. fase 2 .....	95

## 1. Indledning

Med den stigende vægt på matematik i folkeskolen, fx i forhold til optagelse på ungdomsuddannelser, er der behov for en styrket indsats på feltet. Dyskalkuli (talblindhed) er et centralt men underbelyst emne, hvor det bl.a. er uklart, hvor stor udbredelse fænomenet har. Forskellige dele af forskningen vurderer, at 1-10 pct. af en årgang har talblindhed. Talblindhed og ordblindhed er begge specifikke udviklingsforstyrrelser (jf. WHO's klassifikation), og udover at der er centrale lighedspunkter, er der et stort antal elever, der har begge dele (komorbiditet). I dette projekt fokuseres alene på talblindhed.

Hovedsigtet med dette projekt er at følge op på "Talblindhedsprojektet", som i regi af DPU, Aarhus Universitet, har tilvejebragt et samlet testbatteri til vurdering af talblindhed, der i dette projekt er blevet afprøvet, valideret og tilrettet/forbedret.

I henhold til opgavebeskrivelsen skal projektet ses som den første af to faser i afprøvningen. Der er lagt op til, at en evt. fase 2 kan bestå af et eksperiment i stor skala med henblik på at afprøve anvendeligheden af test, testresultater og vejledningsmateriale samt måle forskellige pædagogiske indsatsers effekt.

I nærværende slutrapport redegøres for projektets design, forløb og resultater – herunder en vurdering af kvaliteten af de enkelte dele af det afprøvede testbatteri og det tilhørende vejledningsmateriale.

Endelig gives som afslutning en række anbefalinger, som vedrører:

- Hvilken status testen kan få
- Videre udvikling – herunder udførelsen af en eventuel fase 2
- Øvrige tiltag der kan understøtte en indsats for talblindhed af høj kvalitet

## 2. Grundlaget for projektet

Projektet har haft til formål at afprøve og undersøge kvaliteten af de enkelte dele af et testbatteri bestående af en observationsguide og to test til undersøgelse af talblindhed hos elever i 4. klasse. Testene er udviklet i et forudgående projekt om talblindhed, der blev påbegyndt i 2014 i et samarbejde mellem DPU, Aarhus Universitet og Professionshøjskolen Absalon med bistand fra Københavns Professionshøjskole og selvstændig konsulent Pernille Pind (efterfølgende betegnet Udviklingsprojektet)<sup>1</sup>.

Udviklingsprojektet blev sat i gang af Undervisningsministeriet på baggrund af en aftale fra 2013 om et fagligt løft af folkeskolen. Det blev her aftalt, at der skal sikres en tidligere indsats for ord- og talblinde elever, og at der derfor skal udvikles en test til identifikation af talblindhed til brug for skolerne. Testen skal sammen med faglige vejledninger og metoder stilles gratis til rådighed for skolerne<sup>2</sup>.

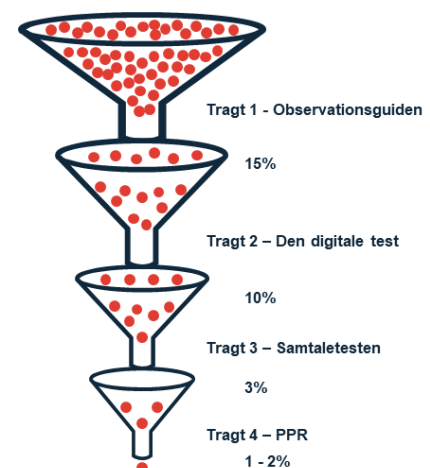
### 2.1 Testbatteriet

I rapporten fra udviklingsprojektet beskrives den samlede talblindhedstest som et batteri bestående af fire redskaber – dette batteri kaldes tragtmodellen. Denne er illustreret i Figur 1.

Modellen består af fire tragte:

- 1. Observationsguiden** består af en række udsagn, som læreren vurderer, hvorvidt passer til eleven. På baggrund heraf beregnes en score og de elever, der ligger under en fastsat grænseværdi, vil ikke betragtes som værende i risiko for talblindhed. For de resterende elever fortsættes forløbet til tragt 2.
- 2. Den digitale test** består af en række opgaver, som eleven skal løse selvstændigt på en computer. På baggrund af elevernes besvarelser og svarhastighed beregnes en score, som afgør om forløbet fortsættes til tragt 3.
- 3. Samtaletesten** afvikles som en dialog med eleven om en række konkrete tal- og regneopgaver. På baggrund af dialogen registrerer læreren elevens svar, og der beregnes en score. Hvis eleven på denne baggrund vurderes at være i risiko for talblindhed, fortsættes til tragt 4.
- 4. Den sidste tragt** består af en vurdering fra **PPR** af, om elevens udfordringer skyldes talblindhed eller i højere grad kan tilskrives andre udfordringer af fx personlig, social eller kognitiv karakter.

Figur 1: Tragtmodellen



Tragtmodellen er altså baseret på en frasortingsmekanisme, hvor hvert trin har til formål at reducere antallet af elever, der vurderes at være i risiko for talblindhed. I testdesignet, som ligger til grund for afprøvningen i dette projekt, er fastlagt, hvor stor en frasortering der skal ske ved hver tragt i testbatteriet. Den procentvise andel af elever, der frasorteres ved tragtmodellen, er fastlagt

<sup>1</sup> Lindenskov, Kirsted, Allerup og Lindhardt (2019) Talblindhedsprojektet – Rapport om udvikling af talblindhedstest og vejledningsmateriale, DPU, Aarhus Universitet og Professionshøjskolen Absalon: [https://emu.dk/sites/default/files/2019-09/Talblindhedsprojektet\\_endelig%20april%202019.pdf](https://emu.dk/sites/default/files/2019-09/Talblindhedsprojektet_endelig%20april%202019.pdf)

<sup>2</sup> Se aftaletekst her: <https://www.uvm.dk/-/media/filer/uvm/udd/folke/pdf/14/okt/141010-endelig-aftaletekst-7-6-2013.pdf>

på grundlag af international forskning i talblindhed og andre landes erfaringer. I afprøvningen er der således tale om en normbaseret tilgang<sup>3</sup> ved fastlæggelsen af grænseværdier.

## 2.2 Projektets organisering

Projektet består overordnet af delopgaverne:

1. Opdatering af vidensgrundlag samt vurdering af test, vejledningsmateriale og indsatser
2. Afprøvning og tilpasning af test og tilhørende vejledningsmateriale
3. Udvikling og afprøvning af lovende pædagogiske indsatser

Projektet har lidt over 800 deltagende skoler, hvoraf alle er bedt om at udvælge de 25 pct. af 4. klasseeleverne, der er mest udfordret i matematikfaget. Ud af de deltagende skoler er otte udvalgt som projektets kerneskolers, dvs. skoler, der deltager i alle dele af projektet (se afsnit 5 om rekruttering af skoler og kerneskolers). De otte kerneskolers fordeler sig på fire kommuner: København, Odense, Lyngby-Taarbæk og Århus.

Som en del af løsningen på delopgave 1 er der udarbejdet en kortlægning af forskningsbaseret materiale vedr. pædagogiske indsatser. På baggrund heraf blev det klart, at en forskningsorienteret tilgang til kortlægning af materiale vedr. pædagogiske indsatser ikke dækker eksisterende undervisningsmaterialer. Der er derfor indarbejdet et kortlægningselement i delopgave 3. Kortlægningen er derved suppleret med viden og erfaringer fra praktikere (lærere, matematikvejledere mv.) samt dialog med eksperter (se afsnit 9.1.1).

Delopgave 2 består overordnet af tre hovedaktiviteter: Afprøvning af items til den digitale test (pilotafprøvning) og to runder for afprøvning af denne. Ved begge testrunder har observationsguiden og samtaletesten stået tilgængelig for samtlige deltagende skoler og elever. I tilfælde, hvor elever på baggrund af den digitale test er vurderet til at være i risiko for talblindhed, har vi opfordret skolerne til at gennemføre samtaletesten. Der blev yderligere lanceret en demoudgave af den digitale test med det formål at give lærere og elever mulighed for at blive fortrolige med testens formater. Data fra demotesten har ikke dannet grundlag for analyse.

Delopgave 3 er i særlig grad fokuseret på kerneskolerne. I samarbejde med eksperter og matematiklærere/-vejledere fra kerneskolerne justeres/udvikles pædagogiske indsatser til understøttelse af elever med tegn på talblindhed. Indsatserne er afprøvet på kerneskolerne. Runde 1 og 2 af testafprøvningen er planlagt hhv. før og efter afprøvning af pædagogiske indsatser med henblik på at skabe grundlag for en effektvurdering.

---

<sup>3</sup> **Normbaseret** betyder, at elevers resultater holdes op mod en norm, dvs. sammenlignes med en fordeling som afspejler den gruppe, der anvendes som reference (fx landsfordelingen af testresultater på klassetrinnet). Ved en 'ægte normbaseret' opdateres normen jævnlige (fx årlige), således at gennemsnittet/normen, der refereres til, altid afspejler det aktuelle niveau. Ofte fastlægges en norm som en gennemsnitsværdi (også kaldet standardværdi) på et tidspunkt, som kan danne reference for fremtidige målinger. Alternativet til normbaseret er kriteriebaseret, hvilket i udgangspunktet henviser til, at elevernes resultater holdes op imod faglige kriterier typisk fastsat af eksperter/fagfolk, der ideelt er objektivt konstaterbare velbeskrevne observationer eller mål. Men begrebet 'kriteriebaseret' er ikke et entydigt defineret og velafgrænset begreb, og i forskningslitteraturen på området findes en række forskellige forståelser af begrebet. Begrebet 'criteria based' behandles ofte som synonymt med 'standard referenced', hvor der defineres en cutscore eller en andel rigtige (fx procentscoring), der betegnes som kriteriet (jf. fx Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*). Der er altså i praksis ofte en flydende grænse mellem norm- og kriteriebaseret.

## 2.3 Tidligere rapporteringer i projektet

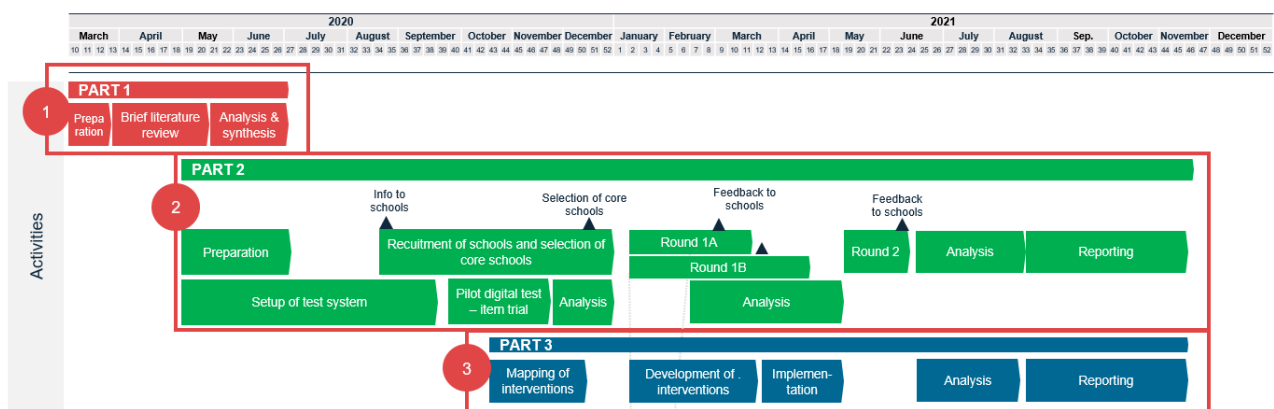
Der er knyttet en række delleverancer til projektet, blandt hvilke følgende fire skal fremhæves:

1. **Forskningsoversigt om talblindhed**, september 2020 (Offentliggjort på EMU<sup>4</sup>). Oversigten indeholder et opdateret vidensgrundlag med oversigter over aktuelle definitioner, nyudviklede test og virkningsfulde pædagogiske indsatser, redskaber mv.
2. **Skriftlig midtvejsstatus**, Udvikling af talblindhedstest til 4. klasse og pædagogiske indsatser, februar 2021 (internt dokument). Midtvejsstatus indeholder en revideret tids- og projektplan, oversigt over rekrutterede skoler, oversigt over opsætning og plan for afprøvning for alle dele af testen, analyser af resultater fra opgaveafprøvningen af den digitale test og vejledningsmateriale samt midtvejsstatus for samarbejde med Kerneskolers, herunder anvendelse af RD4-testen og udvikling af pædagogiske interventioner.
3. **Metodenotat om test**, Notat om talblindhedstestenes psykometriske egenskaber, september 2021 (internt notat). Heri redegøres for den gennemførte afprøvning af opgaver, justeringer og vurdering af test og testmaterialer. Det sker gennem en detaljeret beskrivelse af opgaveafprøvning samt en grundig gennemgang af analyseresultater og argumenter for forslag til justeringer. Yderligere vurderes samspillet mellem den digitale test og samtaletesten.
4. **Metodenotat om pædagogiske indsatser**, september 2021 (internt notat). I notatet redegøres for projektets kortlægning, justering/udvikling af design for samt beskrivelse og afprøvning af pædagogiske interventioner/indsatser.

I nærværende slutrapport er indeholdt en sammenskrivning af alle væsentlige analyseresultater fra de interne rapporteringer, som er nævnt (midtvejsstatus samt metodenotater om hhv. test og pædagogiske indsatser). Der er taget afsæt i resultaterne fra forskningsoversigt om talblindhed (fx ved valg af pædagogiske indsatser), men da denne rapport er offentliggjort i sin helhed, er resultaterne herfra ikke refereret.

## 2.4 Projektets tidsplan og datagrundlag

Figur 2: Oversigt over projektets organisering, præsenteret ved ekspertgruppemøde 2



Af Figur 2 fremgår en oversigt over projektets organisering, som senest er opdateret i marts 2021 i forbindelse med præsentation for projektets ekspertgruppe (se afsnit 4).

Delopgave 1 blev gennemført efter den oprindelige tidsplan og afsluttet i juni 2020.

<sup>4</sup> [https://emu.dk/sites/default/files/2020-11/Forskningsoversigt%20om%20talblindhed\\_2020.pdf](https://emu.dk/sites/default/files/2020-11/Forskningsoversigt%20om%20talblindhed_2020.pdf)



Delopgave 2 blev igangsat i maj 2020 med forberedende arbejde og opsætning af systemer for de tests, der skulle afprøves. Som det ses i oversigten, er der gennemført en pilotafprøvning af den digitale tests opgaver forud for den første egentlige afprøvning af testene (runde 1A og B). Pilotafprøvningen er afviklet som en separat afprøvning af items på de tilmeldte skoler, i alt ca. 4.000 elever, og har til formål at undersøge, om opgaverne i den digitale test passer til de forskellige delskalaer og frasortere eventuelt dårligt fungerende opgaver. På baggrund af en analyse af besvarelserne fra pilotafprøvningen af den digitale test samt frasorteringen af dårligt fungerende opgaver, blev runde 1 af afprøvningen påbegyndt i januar 2021. Ved testafprøvningens frist i marts måned blev der lavet en foreløbig afrapportering af resultater på baggrund af elevbesvarelserne (runde 1A). Dog var mange skoler udfordrede af Corona-restriktioner og nåede derfor ikke at afvikle testene inden fristen – de udfordrede skoler fik en forlængelse af fristen for testafvikling på ca. en måned, hvorefter der blev lavet endnu en afrapportering af resultater (runde 1B).

Sideløbende med delopgave 2, har delopgave 3 været afviklet siden oktober 2020. Det indledende arbejde omhandlede i særlig grad kortlægning af pædagogisk materiale til elever med talblindhed (og svære matematikvanskeligheder) og afklaring af kerneskolernes rolle i projektet. Kerneskolerne har yderligere haft til opgave at gennemføre en hollandsk samtaletest (RD4 - oversat fra hollandsk til engelsk) med de udvalgte elever (se afsnit 8.1). RD4-testen blev afviklet efter kortlægningen, inden runde 1 af testafprøvningen blev sat i gang. Data fra RD4-testen danner grundlag for en validering af den danske samtaletest (jf. afsnit 8.3.4).

Som en central del af samarbejdet med kerneskolerne og delopgave 3 har skolerne udvalgt eller udviklet pædagogiske indsatser til støtte for egne elever med tegn på talblindhed. Udviklingsarbejdet blev igangsat sideløbende med runde 1 af afprøvningen af den digitale test i januar 2021. Indsatserne tager udgangspunkt i kerneskolernes vurdering af, hvad der passer til deres elevgrupper og de udfordringer, der har vist sig bl.a. igennem den hollandske samtaletest og den danske talblindhedstest (se afsnit 9.1.2). Projektgruppen har bidraget med løbende sparring og facilitering af dialog med ekspertgruppen. De planlagte indsatser er afviklet parallelt på de to skoler i hver af de fire kommuner således, at fire forskellige indsatser afprøves mellem 1. og 2. runde af testafprøvningen. Resultater fra de to afprøvninger er sammenholdt med henblik på at afdække eventuelle målbare effekter af indsatserne (se afsnit 11).

I Tabel 1 ses en samlet oversigt over projektets datagrundlag.

**Tabel 1: Oversigt over projektets datagrundlag**

		<b>Elevniveau</b>	<b>Lærer-/koordinatorniveau</b>
Data fra tilmeldings-survey (skoleoplysninger og deltagelsens omfang)	Alle skoler		X
	Kerneskoler	X	X
Observationsguide		X	X
Digital test, Runde 1 og 2 i testafprøvningen	Samlet score	X	
	De fem delskalaer	X	
	Motivation og selvtillid	X	
Samtaletest		X	X
Spørgeskemaundersøgelse til lærere (Lærersurvey)			X

### 3. Definition af talblindhed

Et fundamentalt spørgsmål, som der ikke er enighed om blandt forskere og eksperter, er, hvordan talblindhed afgrænses og defineres – nogle betvivler ligefrem talblindhed som et selvstændigt fænomen. I udviklingsprojektet er der arbejdet med en definition, som man kan kalde den rene talblindhed. Denne definition er også udgangspunktet for aktiviteterne i dette projekt:

*”Talblindhed er en læringsudfordring, der er påvirket af en specifik neurologisk udviklingsforstyrrelse, som kan have forskellige udtryk, men som ikke primært kan forklares på baggrund af generelle indlæringsvanskeligheder, mangelfuld undervisning, psykologiske eller sociologiske årsager. Talblindhed omfatter vanskeligheder ved at automatisere tal, antal og størrelser samt fastholde og anvende aritmetiske færdigheder.” (Udviklingsprojektets definition)*

Forskningen peger på, at der er visse medfødte og tillærte basale færdigheder, som elever med talblindhed kan have vanskeligheder med. Det drejer sig blandt andet om at kunne vurdere størrelsen af mængder uden at tælle elementerne i dem og om at kunne oprette en mental tallinje med ordnede tal og dermed kunne regne med tal.

Den talblindhed, der er neurologisk betinget, benævnes i den engelsksprogede forskning som ”Developmental Dyscalculia” (forkortet DD)<sup>5</sup>. DD betragtes i mange sammenhænge som det, der kan betegnes som ”en specifik indlæringsvanskelighed, der påvirker den normale tilegnelse af aritmetiske færdigheder”<sup>6</sup>. DD forekommer ofte i kombination med andre indlæringsvanskeligheder (især fremhæves i forskningen dysleksi), hvilket kan medvirke til at forstærke vanskelighederne med at lære matematik<sup>7</sup>. DD er registreret på WHO’s liste over neurologiske udviklingsforstyrrelser ([ICD-11, gruppe 6, Mental, behavioral and neurodevelopmental disorders](#)) med koden 6A03.2 (Developmental dyscalculia, Developmental learning disorder with impairment in mathematics) lige efter ordblindhed med koden 6A03.0 (Developmental dyslexia, Developmental learning disorder with impairment in reading).

#### 3.1 Omfang af talblindhed i Danmark

Forskningens definition af talblindhed og omfanget heraf varierer en del fra undersøgelse til undersøgelse. De estimater, der har været normsættende for tragtmodellen, bygger på erfaringerne fra udviklingsprojektet, der havde til formål at udvikle talblindhedstesten. Her redegøres for baggrundene for, at andelen af talblinde anslås til 1–2 pct. i dette projekt. Procentsatsen er begrundet i den ”smalle angivelse” på 1,8 pct. hos von Aster<sup>8</sup>. Det anføres i rapporten fra udviklingsprojektet, at andelen kunne have været på mellem 3,6 og 8,4 pct., hvis ikke ”brede afgrænsninger var blevet afvist som upræcise”. I flere større kvantitative undersøgelser angives, at 3–6 pct. af en befolkning kan formodes at være talblinde (Butterworth, England; Shalev, Israel; von Aster, Tyskland/Schweitz). Men der er også undersøgelser, hvor det findes, at udbredelsen af talblindhed (prævalensen) er

<sup>5</sup> Von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental medicine & child neurology*, 49(11), 868-873.

<sup>6</sup> Shavel definerer DD således: ”Developmental dyscalculia is a specific learning disability affecting the normal acquisition of arithmetic skills”, Shalev, R. S., & Gross-Tsur, V. (2001). *Developmental dyscalculia. Pediatric neurology*.

<sup>7</sup> Malmer, G. (2000). Mathematics and dyslexia—an overlooked connection. *Dyslexia*, 6(4), 223-230.

<sup>8</sup> Det som von Aster benævner ”ren talblindhed” (pDD). I artiklen anslår von Aster, at der er 6 pct. talblinde hvis dem, der både er tal- og ordblinde, medregnes.

den samme som for ordblindhed, som regel i intervallet 5-11 pct. af en befolkning (fx Wilson 2015<sup>9</sup>).

I rapporten fra udviklingsprojektet anføres, at den skønnede andel talblinde også har afsæt i undersøgelser, som blev gennemført i regi af udviklingsprojektet. Det fremgår dog ikke direkte om, og i givet fald hvordan, disse undersøgelser understøtter, at man kan finde frem til 1-2 pct. talblinde af en årgang fra 4. klassetrin, men det kan formentlig ses som et minimumsniveau<sup>10</sup>. Hvis elever i indskolingen inddrages i opgørelsen, vil andelen antageligt øges: Som det fremgår af von Asters udviklingsmodel (jf. Tabel 2, nedenfor), udvikles talforståelse og aritmetiske færdigheder med alderen. Derfor vil en digital talblindhedstest med stor sandsynlighed finde et større antal elever med tegn på talblindhed, hvis man – som von Aster – arbejder med en målgruppe, der dækker 1.-4. klassetrin.

Et af de screeningsredskaber til talblindhed, som har stor udbredelse og for mange udgør en reference, er den test, som Butterworth har udviklet til identifikation af tendens til talblindhed<sup>11</sup>. Butterworths datagrundlag for vurdering af forekomsten af talblinde i en befolkning er ganske stort (>15.000 personer) og øges stadig. Et af de senere estimater lyder på 5-7 pct. talblinde af en årgang<sup>12</sup>. I de senere års internationale forskning<sup>13</sup> har opfattelsen af prævalensen af talblindhed været, at den ligger på omkring 6 pct., hvori komorbiditeten, fx ift. ordblindhed, indregnes.

De mange forskellige skøn dækker ikke nødvendigvis over uenighed eller fejl i de statistiske oplysninger, men måske snarere, at der ikke er en skarp grænse for talblindhed – talblindhed kan manifestere sig i forskellige niveauer og grader<sup>14</sup>. De praktiske opgørelser af antallet af elever med talblindhed beror på resultater og grænseværdier i forskellige test.

Det samme gør sig gældende for ordblindhed. I vejledningen til ordblindetesten fremgår, at det tidligere Dansk Videnscenter for Ordblindhed vurderede, at 2-5 pct. af befolkningen er ordblinde – hvilket er i overensstemmelse med Socialstyrelsens angivelse. Ordblindeforeningen anser 3-7 pct. af befolkningen som værende ordblinde, mens grænseværdien i Ordblindetesten er sat ved 8 pct. Grænseværdien i Ordblindetesten ligger således over de ellers anvendte grænseværdier for ordblindhed. Hermed sikres angiveligt, at personer med både svær og mild grad af ordblindhed identificeres med Ordblindetesten.

Det kan således tyde på, at omfanget af talblindhed og ordblindhed er i samme størrelsesorden, og at mange talblinde også er ordblinde (komorbiditet)<sup>15</sup>.

<sup>9</sup> Wilson, A. J., Andrewes, S. G., Struthers, H., Rowe, V. M., Bogdanovic, R., & Waldie, K. E. (2015). Dyscalculia and dyslexia in adults: cognitive bases of comorbidity. *Learning and individual differences*, 37, 118-132.

<sup>10</sup> Det fremhæves i rapporten fra udviklingsprojektet, at der skete en ressourcebaseret tilpasning af cutscoren fra 10 pct. til 7 pct., og at det medførte, at det identificerede antal elever som følge heraf blev reduceret fra 80 til 30. Det beskrives, at der derfor er en risiko for, at "der på denne måde i udviklingsprocessen var falsk negative, altså talblinde elever, der ikke diagnosticeredes" S85 (se: [https://emu.dk/sites/default/files/2019-09/Talblindhedsprojektet\\_endelig%20april%202019.pdf](https://emu.dk/sites/default/files/2019-09/Talblindhedsprojektet_endelig%20april%202019.pdf))

<sup>11</sup> Fx Butterworth, B. (2003). *Dyscalculia screener*. London, UK: GL Assessment.

<sup>12</sup> Butterworth, B. (2018). *Dyscalculia: From science to education*. Routledge.

<sup>13</sup> Se fx Haberstroh, S., & Schulte-Körne, G. (2019). The diagnosis and treatment of dyscalculia. *Deutsches Ärzteblatt International*, 116(7), 107.

<sup>14</sup> Kucian, K., & von Aster, M. (2015). Developmental dyscalculia. *European journal of pediatrics*, 174(1), 1-13.

<sup>15</sup> Shalev, R. S., Auerbach, J., Manor, O. H. A. D., & Gross-Tsur, V. (2000). Developmental dyscalculia: prevalence and prognosis. *European child & adolescent psychiatry*, 9(2), S58-S64.

### 3.2 Udvikling af numerisk kognition - en model i 4 trin

DD (Developmental Dyscalculia, herefter benævnt Talblindhed) skal således ses som en specifik udviklingsforstyrrelse, der påvirker udviklingen i tilegnelsen af numeriske og aritmetiske færdigheder. Elever med talblindhed tilegner sig også numeriske og aritmetiske færdigheder, men tilegnelsen bliver forsinket og besværliggjort i forhold til elever uden talblindhed.

Med afsæt i forskning på området (von Aster 2007) kan den generelle udvikling beskrives således: Evnen til at sammenligne mængder udvikles normalt i den tidlige barndom (trin 1), hvilket giver betydning til 'antal', hvilket igen er en forudsætning for at tilegne sig talord, tallenes mundtlige repræsentation (trin 2), og senere de visuelle arabiske talsymboler (trin 3), mens en voksende arbejdshukommelse muliggør udvikling af en ekspanderende mental tallinje i skoleårene (trin 4). Den almindelige udvikling for ikke-talblinde elever er illustreret i Tabel 2.

**Tabel 2: Normal udvikling af numerisk kognition, en model i 4 trin**

Udviklingstrin	Alder	Repræsentation
<b>Trin 1:</b> Basal mængdeforståelse, sammenligning (konkrete mængder)	Vuggestue (-> 3 år)	Kardinalitet 
<b>Trin 2:</b> Det verbale talsystem (tallene som lyde)	Børnehave (3-5 år)	 En, to, tre . . . .
<b>Trin 3:</b> Talsystemet og skriftlig aritmetik (de arabiske tal)	0. klasse (5-6 år)	. . . 8, 9, 10, 11, . . .
<b>Trin 4:</b> Overslagsregning og aritmetisk tænkning (positionssystem, regningsarter m.v.)	Indskolingen (7år+)	Ordinalitet (den mentale tallinje) 

Kilde: von Aster, 2007

Modellen beskriver altså en seriel udvikling i fire trin. Forudsætningen for at komme videre til næste trin uden vanskeligheder er, at færdighederne fra det foregående trin er opnået. En afledt virkning heraf er ifølge forskerne bag modellen, at indsigt i, på hvilket trin eleven befinder sig, kan give et vigtigt input til udformning af passende uddannelsesinterventioner.

Som det ses af modellen i Tabel 2, forventes trin 1-3 at være på plads i starten af indskolingen, således at færdighederne kan siges at være tilnærmelsesvist automatiserede. Og i 4. klasse har alle elever (også langt de fleste talblinde elever) tilegnet sig færdigheder, der sætter dem i stand til at svare korrekt på lette opgaver på disse niveauer, fx opgaverne i projektets digitale test.

Projektets digitale test er en test med 5 skalaer (jf. afsnit 6.4) – de første fire af disse skalaer er hovedsageligt designet til at fokusere på færdighederne i de første tre trin i modellen, herunder deres automatisering, mens skala 5 i den digitale test og samtaletesten (jf. afsnit 6.5) har hovedfokus på trin 4 i modellen i Tabel 2.

### 3.3 Identifikation af talblindhed og ordblindhed

Forskning i talblindhed er stadig på et tidligt stadie, hvilket antageligt hænger sammen med, at talblindhed er en meget forskelligartet forstyrrelse, som har mange fremtrædelsesformer, og derfor er vanskelig at beskrive udtømmende – og dermed også at identificere. Dette i modsætning til fx ordblindhed, der i øvrigt på mange områder har lighedspunkter med talblindhed. Mest markante forskelle og lighedspunkter kan beskrives således:

- Både talblindhed og ordblindhed er neurologiske udviklingsforstyrrelser, som man gennem observation af adfærd (fx usikker oplæsning, tællen på fingrene) eller med en god test kan finde tegn på (lav score).
- Ordblindhed består kort sagt i *vanskeligheder med at lære skriftens lydprincipper*, dvs. at afkode skrevne ord<sup>16</sup>, med andre ord at omsætte skrifttegn til lyd. En tests normale funktion er at afdække en latent egenskab, dvs. noget som ikke kan observeres med det blotte øje, fx ordblindhed. Det gøres ved at afprøve noget, der lader sig teste og/eller observere, fx evnen til at afkode ord. Herefter sammenfattes resultaterne af en række opgavebesvarelser til en samlet score, målt på en éndimensionel skala. Pointen er, at en god afkodningstest kan tages som en sikker indikator på fonologiske vanskeligheder, dvs. som er et sikkert tegn på ordblindhed (givet visse forudsætninger om elevens sproglige forståelse).
- Talblindhed er vanskeligere at identificere, bl.a. fordi fænomenet har mange fremtrædelsesformer jf. fx von Asters model præsenteret i Tabel 2. Talblindhed kan fremtræde som problemer med basal mængdeforståelse, fx ANS (*approximate number sense*, dvs. sammenligning af to mængder) eller subitisme (genkendelse af mængder op til fire), det kan fremtræde som problemer med at forbinde talsymboler med talord og/eller mængder, men det kan også fremtræde som problemer med basale operationer i forhold til regningsarterne og tallinjer.

Talblindhed kan altså have forskellige fremtrædelsesformer, hvilket også afspejler sig i, at alene den digitale test består af fem underskalaer, der måler noget forskelligt (se afsnit 6.4.1). Hertil kommer indholdet af samtaletesten, som tilfører yderligere dimensioner til målingen af talblindhed. Talblindhed er således et væsentligt mere komplekst fænomen end ordblindhed, uanset at der er mange lighedspunkter mellem de to neurologiske forstyrrelser. Dette er utvivlsomt en af årsagerne til, at der hersker en større forskningsmæssig uafklarethed ift. en entydig definition af talblindhed og identifikation af effektive metoder til pædagogiske indsatser med henblik på at afhjælpe virkningerne af talblindhed.

### 3.4 Diagnosticering af talblindhed og målingers validitet

Når man skal diagnosticere talblindhed, er der få ting, alle forskere er enige om. Én af disse er, at det ikke er nok at se på besvarelserne af en test/prøve. Der er elever uden talblindhed, der klarer sig meget dårligt i almindelige matematikprøver, ligesom der er elever med talblindhed, som har udviklet strategier til at omgå problemerne og svare rigtigt på mange typer regneopgaver – fx ved at tælle på fingrene og/eller bruge meget lang tid herpå.

Der findes screeningsredskaber til talblindhed, som har en bred anerkendelse, fx screeningsredskabet udviklet af Butterworth. I daglig tale siges det ofte, at testen skal kunne ”diagnosticere talblindhed”, men det er en misfortolkning af testens virkemåde, hvilket Butterworth selv understreger, når han siger, at hans screener foretager ”diagnosticering af *tegn* på talblindhed”<sup>17</sup>. Dermed er det ikke udviklingsforstyrrelsen talblindhed, der diagnosticeres, men snarere *tegn herpå*, som en test leverer. Testen er dermed kun leverandør af data til en mere dybtgående personlig vurdering af eleven. Butterworth oplyser, at man med hans screener kan identificere 3-6 pct. af en årgang (bl.a. afhængigt af klassetrin) som elever med tegn på talblindhed.

Skal man identificere talblindhed, er det nødvendigt ikke blot at se på **testresultater**, men også at analysere **de processer, der fører frem til resultatet**. Svarprocessen får man normalt ikke meget

<sup>16</sup> Elbro, C. (2006). Seminarierien. Læsning og læseundervisning. Gyldendal Uddannelse.

<sup>17</sup> Fx Butterworth, B. (2003). Dyscalculia screener. London, UK: GL Assessment.

at vide om ved at se på resultatet af fx en regne- eller matematiktest – her vil typisk skulle anvendes nogle observations- og/eller dialog-/samtalebaserede teknikker (fx en samtaletest), hvor man kan observere adfærd, der kan være tegn på talblindhed.

Dog er der information om svarprocesser, som vil kunne opfanges maskinelt i forbindelse med en digital test med avancerede redskaber (fx eye-tracking, analyse af tastaturintegration, analyse af følelser afspejlet i ansigtsudtryk, biometriske målinger<sup>18</sup>). Men der er et mere enkelt mål, der viser en væsentlig forskel på de lavt og normalt præsterende elever – nemlig svarhastigheden for den enkelte korrekt besvarede opgave. Hvor hurtigt eleven svarer (på opgaver hvor eleven svarer rigtigt), kan tages som et mål for automatiseringsgraden. Svarhastigheden er inddraget i både Butterworths dyscalculia screener og i den oprindeligt udviklede digitale test. Ved at kombinere oplysninger om svarhastighed og -rigtighed kan der tegnes et mere nuanceret billede af elevernes dygtighed og deres sandsynlige svarstrategi. Sammenhængen mellem elevdygtighed, svarhastighed og rigtighed er illustreret i Tabel 3.

**Tabel 3: Rigtighed og svartid på en let opgave**

		Typisk svartid for eleverne i 4. klasse med forskellig dygtighed		
		Langsomt	Normalt → Hurtigt	Meget hurtigt
Besvarelse	Korrekt	De lavt præsterende	De normalt & højt præsterende	De heldige & meget dygtige
	Forkert	De meget lavt præsterende	De lavt præsterende	De uengagerede

For elever hvis opgavebesvarelser i betydelig grad hører til i de røde felter i tabellen (forkerte og/eller langsomme), kan besvarelsen fortolkes som tegn på talblindhed. Elever med udelukkende grønne felter tyder på, at eleven ikke er talblind, mens besvarelse i det hvide felt kan tages som udtryk for, at eleven ikke har gjort et forsøg på at finde det rigtige svar. Der bør derfor ses bort fra nogle af de meget hurtige besvarelser ved scoring af testen, hvilket kan gøres rent teknisk (jf. afsnit 7.2).

I fortolkning af data forudsættes som regel, at elevens besvarelser er retvisende for elevens dygtighed – dvs. at scoren er valid. Det er imidlertid ikke nok, at testen fungerer som den skal, dvs. at testen i sig selv leverer en valid måling givet responserne<sup>19</sup>. Det er fx en trussel mod målingens validitet, hvis eleven har modtaget dårlig eller misvisende instruktion, hvis eleven har modstand eller modvilje mod testen og/eller af anden grund mangler koncentration eller engagement – eller blot har en dårlig dag.

Det kan altså spille en helt central rolle for elevens tilgang til testen, hvordan testen og dens formål er blevet præsenteret. Derfor er retvisende vejledning, instruktion og feedback en væsentlig forudsætning for testresultatets validitet.

### 3.5 Matematikvanskeligheder og talblindhed

Ofte beskrives talblindhed som en form for matematikvanskelighed. Og vanskeligheder med at løse simple aritmetiske opgaver kaldes ofte for talblindhed. Der bliver med andre ord sat lighedstegn mellem visse matematikvanskeligheder og talblindhed. Det er en nærliggende forveksling, for der er et betydeligt overlap: Hovedreglen er, at talblinde elever har matematikvanskeligheder, men

<sup>18</sup> Reckase, M. D. (2017). A tale of two models: Sources of confusion in achievement testing. ETS Research Report Series, 2017(1), 1-15.

<sup>19</sup> Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological bulletin, 52(4), 281.

der er både elever med matematikvanskeligheder, der ikke er talblinde, og talblinde elever der ikke har matematikvanskeligheder. Således er der talblinde elever, som har udviklet og trænet effektive strategier til at omgå problemerne og derfor svarer hurtigt og rigtigt på de opgaver, som talblinde normalt har svært ved. Ligeledes er der mange ikke-talblinde elever, der *ikke* har opnået basale aritmetiske og matematiske færdigheder (fx i forbindelse med andre læringsudfordringer, dårlig/manglende undervisning i regning, sociale og/eller sproglige problemer mv.).

Følgende to elevprofiler illustrerer denne kompleksitet og mulige fejlkilder (eksemplerne er konstruerede med inspiration fra observationer i projektet):

### Profil 1

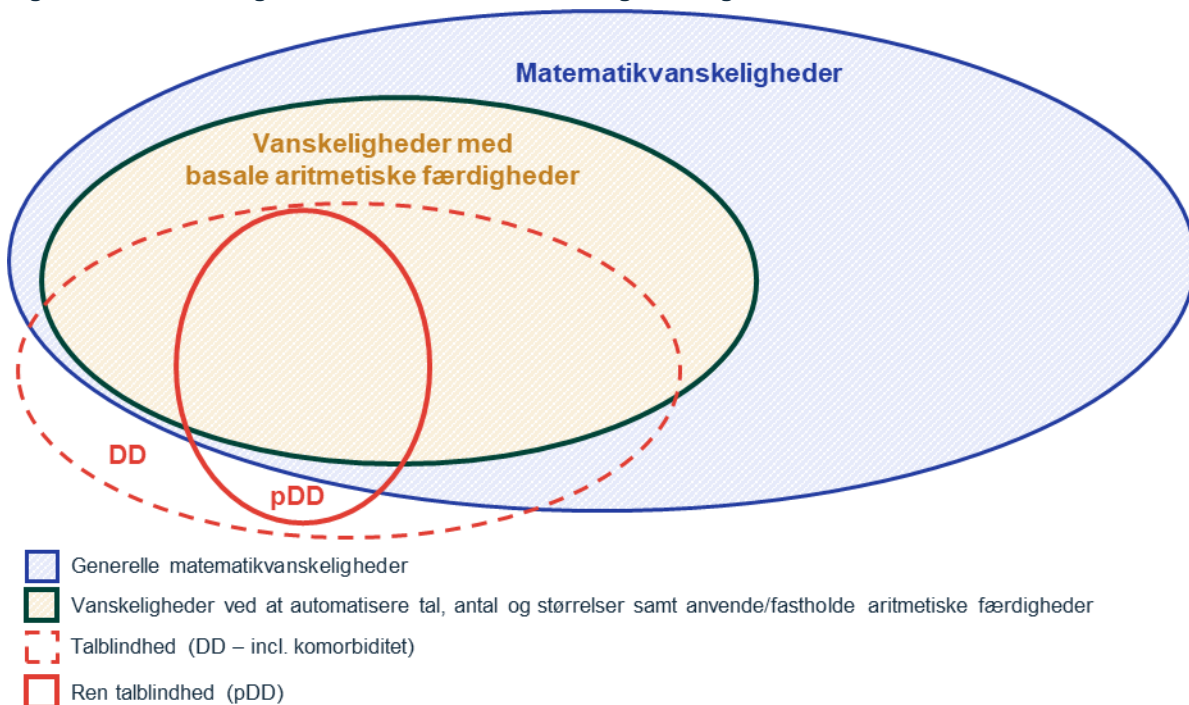
*En computerbaseret aritmetiktest viser, at en elev laver mange fejl i opgaver, der kræver basale aritmetiske færdigheder, mens andre former for evaluering viser, at han godt kan løse opgaver, der kræver avanceret matematisk ræsonnement. Resultatet fortolkes som tegn på, at eleven kunne være talblind. Nøjere undersøgelser viser dog, at eleven ikke er talblind, men klarede opgaverne dårligt af andre årsager (fx koncentrationsproblemer, ADHD).*

### Profil 2

*En anden elev i klassen klarer alle opgaverne i den samtalebaserede talblindhedstest hurtigt og uden fejl. Men hendes løsningsstrategier (fx tæller hun på fingrene) viser, at hun faktisk har udviklingsforstyrrelsen dyskalkuli/talblindhed. Grunden til at hun alligevel klarer sig så godt er, at hun har udviklet strategier (gennem flere års målrettet indsats i form af undervisning og træning) til hurtigt og sikkert at løse matematikopgaver, som de fleste andre talblinde har meget svært ved.*

Sammenhænge mellem matematikvanskeligheder og talblindhed er søgt illustreret i Figur 3.

Figur 3: Sammenhænge mellem matematikvanskeligheder og talblindhed



Figuren illustrerer følgende sammenhænge:

- De elever, der har problemer med at løse basale opgaver i aritmetik, udgør en ægte delmængde af elever med matematikvanskeligheder.
- De fleste talblinde elever (DD) har matematikvanskeligheder – med særlig vægt på basale aritmetiske færdigheder.

- Der er elever med talblindhed, som ikke har matematikvanskeligheder (fx som følge af effektiv støtte, træning og undervisning) og derfor ikke har behov for yderligere indsatser ift. deres talblindhed.
- De 'rent' talblinde elever (pDD – dvs. elever, der ikke har andre vanskeligheder end talblindhed), udgør en ægte delmængde og sandsynligvis et mindretal af de talblinde elever (DD).

Selvom der er et meget stort overlap mellem talblinde elever og elever med matematikvanskeligheder, bør der principielt skelnes systematisk mellem på den ene side matematikvanskeligheder (herunder vanskeligheder med tal, antal, størrelser og aritmetiske færdigheder) og på den anden side talblindhed (DD), herunder den rene talblindhed (dvs. pDD – pure Developmental Dyscalculia.).



## 4. Ekspertgruppen

Epinion og DPU har tilknyttet førende danske og internationale ressourcepersoner til projektet i en ekspertgruppe, der mødes tre gange i løbet af projektperioden med henblik på at bidrage til hver af de tre delopgaver. Ekspertgruppen består af seks eksperter/forskere og en række praktikere.

De seks eksperter er **Lena Lindenskov**, lektor i matematikkens didaktik, afdeling for fagdidaktik, DPU, Aarhus Universitet. Lena er formand for ekspertgruppen, og hun har stået i spidsen for udviklingen af den samlede testmodel, som afprøves i indeværende projekt. I tillæg hertil indgår **Bent Lindhardt**, lektor i matematik ved Professionshøjskolen Absalon, som har omfattende erfaring i arbejdet med talblindhed. Blandt andet har han sammen med Lena Lindenskov stået i spidsen for udviklingen af den model for test af talblindhed, der anvendes som afsæt for dette projekt. **Steen Polk**, PPR i Svendborg, er inddraget som ekspert i det praktiske pædagogisk psykologiske arbejde på talblindhedsområdet. Steen Polk har en omfattende viden og erfaring med det pædagogisk psykologiske arbejde (PPR) med elever med talblindhed og diagnosticering heraf. Endvidere har vi tilknyttet professor **Mieke van Groenestijn** som har 30 års erfaring inden for området og som har stået for udvikling og validering af den hollandske RD4-test for talblindhed og matematikvanskeligheder. **Steve Chinn** er professor ved University of Derby i England, hvor han beskæftiger sig med matematikvanskeligheder, herunder talblindhed. Han er yderligere forfatter til en række fagbøger om matematikundervisning og -læring for elever med vanskeligheder og talblindhed. Endelig indgår **Angela Verschoor** fra CITO i Holland. Hun er en af de førende internationale eksperter i praktisk psykometrisk udvikling af it-baserede test og har en omfattende praktisk og teoretisk ekspertise i konstruktion og implementering af test i mange forskellige sammenhænge.

**Praktikerne** udgøres af koordinatore (som udgangspunkt matematikvejledere) fra de otte skoler, der blev udvalgt som kerneskolers til projektet. Der var 1-2 koordinatore pr. skole. Udvælgelsesmetoden og kerneskolernes opgave i projektet er uddybet i afsnit 5.2.

Den valgte organisering sikrer, at vi i videst muligt omfang inddrager international ekspertise og forskning for at skabe et bredt forskningsmæssigt vidensgrundlag for testen, vurderingen af dens kvalitet, samt hvilken status testen kan få. Ligeledes sikrer organiseringen et samarbejde med videnspersoner og praktikere med indsigt i matematikdidaktik og test, hvilket understøtter og kvalificerer projektets gennemførelse og resultater samt understøtter sektoren med relevant viden og handlemuligheder om talblindhed.

Arbejdet i ekspertgruppen har været organiseret som en møderække, og der har været afholdt tre ekspertgruppemøder i løbet af projektet. Nedenfor fremgår en oversigt over hovedindholdet af disse møder.

**Tabel 4: Ekspertgruppens møderække**

<b>Møde</b>	<b>Emner for mødet</b>
Ekspertgruppe-møde 1 25.6.2020	<ul style="list-style-type: none"> <li>• Forudgående samt nuværende projekt v. DPU &amp; Epinion</li> <li>• Resultatet af forskningskortlægningen v. Epinion</li> <li>• Den hollandske model (RD4-testen) v. Mieke van Groenestijn</li> <li>• Talblindhedstestens setup og scoringsmekanismer v. Epinion</li> <li>• Pædagogiske indsatser v. Steve Chinn</li> </ul>
Ekspertgruppe-møde 2 5.3.2021	<ul style="list-style-type: none"> <li>• RD4 anvendelse på kerneskolerne v. Mieke van Groenestijn</li> <li>• Foreløbige designs af pædagogiske indsatser v. kerneskoler &amp; Lena Lindenskov</li> <li>• Metode og resultater fra den digitale test v. Epinion</li> <li>• Samtaletesten og tragtmodellen v. Epinion</li> <li>• Validering af samtaletesten v. Epinion &amp; Mieke van Groenestijn</li> <li>• Effektmålingens design v. Epinion</li> </ul>
Ekspertgruppe-møde 3 23.9.2021	<ul style="list-style-type: none"> <li>• Afprøvede pædagogiske indsatser v. Epinion</li> <li>• Resultater fra samtaletest v. Epinion</li> <li>• Validering af samtaletest v. Mieke van Groenestijn</li> <li>• Skolens samarbejde med PPR v. Steen Polk</li> <li>• Resultater fra digital test (runde 1 og 2) v. Epinion</li> <li>• Resultater fra effektmåling v. Epinion</li> <li>• Resultater fra lærersurvey v. Epinion</li> <li>• Mødet blev afrundet med en generel diskussion, råd og anbefalinger til næste fase i projektet.</li> </ul>

Foruden de planlagte ekspertgruppemøder blev der afholdt ugentlige møder mellem kerneskolernes koordinatører og Mieke van Groenestijn i perioden uge 44-49 med henblik på instruktion i anvendelse og fortolkning af den hollandske samtaletest, RD4, som er yderligere beskrevet i afsnit 8.3.4 (Kriterievalidering med RD4).

I tillæg hertil efterspurgte kerneskolernes koordinatører møder med eksperterne i forbindelse med udvikling af pædagogiske interventioner med fokus på feedback, råd og anbefalinger. Der blev således faciliteret to møder mellem eksperterne og kerneskolernes koordinatører: ét med de dansksprogede eksperter, Lena Lindenskov, Bent Lindhardt og Steen Polk og ét engelsksproget møde med Steve Chinn og Mieke van Groenestijn.

## 5. Rekruttering af skoler til afprøvning

For at opnå et tilstrækkeligt datagrundlag har der været behov for at invitere alle landets grundskoler til projektet. Følgende kriterier har været gældende for, at skoler er inviteret til deltagelse:

1. Skolen er folkeskole, friskole eller specialskole
2. Skolen har elever i 2. eller 4. klasse (2018-2019)<sup>20</sup>

Dette giver en bruttoliste på 1.756 skoler, der blev inviteret til at deltage i projektet. Rekrutteringen er foregået ved en række aktiviteter jf. nedenstående Tabel 5.

**Tabel 5: Tidsplan og aktiviteter ifm. rekruttering**

Dato	Afsender	Aktivitet
16. juni 2020	STUK	Orienteringsbrev til skoleledere
19. juni 2020	STUK	Orienteringsbrev til kommunaldirektører
10. august 2020	Epinion	Invitation til alle relevante grundskoler og offentliggørelse af hjemmeside <sup>21</sup>
13. august 2020	STUK	Orienteringsbrev til børn- og ungedirektører
18. august 2020	Epinion	Udarbejdelse af nyhedsartikel i folkeskolen.dk <sup>22</sup>
9. september 2020	Epinion	Påmindelse til skoler, der ikke svarede på første henvendelse

Derudover har Epinion i hele projektperioden haft en hotline (e-mail + telefon), hvor skolerne kunne henvende sig med spørgsmål til projektet. Som led i alle skriftlige henvendelser til skolerne er der desuden henvist til hjemmesiden <https://epinionglobal.com/en/talblindhed>, hvor projektets formål, tidsplan og øvrige relevante materialer løbende er gjort tilgængeligt.

### 5.1 Tilmelding til projektets aktiviteter

Selve tilmeldingen af den enkelte skole er foregået ved udfyldning af et online survey. Heri har skolen bl.a. skulle angive oplysninger om:

- Kontaktoplysninger på skolens tovholder/kontaktperson – typisk matematikvejlederen
- Tilsagn om deltagelse i projektets dele ud over den digitale test, dvs. pilotafprøvning, observationsguide og samtaletest
- Skolen forventer at deltage med alle relevante elever eller kun udvalgte elever
- Skolens interesse i at deltage som kerneskole

Tilmelding til projektet var åben i perioden 10. august til og med 30. december 2020. Pr. 1. januar 2021 var der i alt 817 tilmeldte skoler<sup>23</sup>. Blandt dette meget brede udsnit af de danske grundskoler er repræsenteret skoler fra 96 af landets kommuner.

I Tabel 6 fremgår antallet af deltagende skoler på tværs af regioner.

<sup>20</sup> Tal om elevbestand er fra skoleåret 2018-2019. Hvis fx en skole kun har elever i 2. klasse, vil de således være i 4. klasse i skoleåret 2020-2021.

<sup>21</sup> <https://epinionglobal.com/en/talblindhed>

<sup>22</sup> <https://www.folkeskolen.dk/1850210/nu-sker-der-noget-talblindhedstesten-klar-til-foerste-afproevning-til-januar>

<sup>23</sup> I opgørelsen er fratrukket 38 skoler, der har afmeldt sig igen. Størstedelen af disse skoler har framellet sig, da de ikke har haft elever, som læreren vurderede, var relevante at teste for talblindhed.

**Tabel 6: Deltagende skoler i projektet pr. 1. januar 2021**

Region	Deltagende skoler	Inviterede skoler	Andel deltagende skoler
Region Hovedstaden	164	411	40%
Region Midtjylland	214	422	51%
Region Nordjylland	110	210	52%
Region Sjælland	132	292	45%
Region Syddanmark	197	421	47%
<b>I alt</b>	<b>817</b>	<b>1756</b>	<b>47%</b>

## 5.2 Rekruttering af kerneskolers og reserveskolers

Som en del af projektet er der rekrutteret en række skoler, som indgår i en større del af projektets aktiviteter end de øvrige deltager-skoler – disse kaldes kerneskolers i projektet. Hovedformålet for deres deltagelse er, at...

- ... få afprøvet alle dele af den samlede test omhyggeligt efter projektets retningslinjer
- ... opnå reflekteret og gennemtænkt feedback på alle dele af den samlede test i direkte dialog med ekspertgruppen
- ... udvælge/tilpasse/designe fire forskellige eksemplariske undervisningsforløb og afprøve dem på to skoler hver.

Tilmeldingerne blev foretaget medio september 2020. Umiddelbart efter tilmelding udvalgte Epinion en række skoler med henblik på dialog om deltagelse som kerneskolers. Skolerne blev kontaktet for at afdække de ovenfor nævnte rammer. Primo oktober blev udvalgt de otte skoler, der bedst opfyldte kravene og som på kvalitative parametre skønnedes at være de bedst egnede til at bidrage til projektets mål.

Som led i identifikation af kerneskolers identificeredes for hvert udvalgte sæt af kerneskolers de mest velegnede 'reserveskolers' (og 'reservekommuner'). Disse kunne aktiveres såfremt udvalgte kerneskolers faldt fra inden opstart af forløbet.

Ved udformningen af kriterier for udvælgelsen blev danske eksperter konsulteret. Udgangspunktet for projektets definition af talblindhed er, at der fokuseres på den rene neurologisk baserede talblindhed ("pure Developmental Dyscalculia", jf. von Aster 2007) i modsætning til en forståelse af dyskalkuli som lav præstation på generelle matematiktest ud fra læseplaner. Det blev på grundlag af denne dialog besluttet at forsøge at rekruttere skoler, der i mindst muligt omfang har ressourceproblemer og/eller belastede elevgrundlag, fx en stor andel elever med sociale og sproglige problemer.

Udvælgelsen af kerneskolers foregik i henhold til følgende seks kriterier, som alle i videst mulige omfang skulle være opfyldt:

1. Kerneskolersne udvælges 'i par' på kommuneniveau – dvs. to skoler fra fire kommuner
2. Der skal være mindst to spor på 4. klassetrin med hver deres fagligt stærke matematiklærer
3. Lav andel af elever hvor eventuelle matematikvanskeligheder skyldes andre forhold end talblindhed (sprogvanskeligheder, generelle læringsvanskeligheder, hyppige lærerskift, sociale problemer, utilstrækkelig undervisning)
4. Stærkt matematikfagligt miljø omkring 4. klasseeleverne på skolen

5. Skolens kontaktperson skal være en fagligt stærk og erfaren matematikfaglig vejleder, der kan fungere som koordinator internt på skolen og som kan repræsentere skolen i forhold til projekt/ekspertgruppen
6. Projektet, lærerne og kontaktpersonen skal have en ledelsesmæssig stærk opbakning til deltagelse i arbejdet med at implementere projektets initiativer og bidrag til ekspertgruppen, da de erfaringsmæssigt afgørende faktorer for en succesfuld indsats vil være den seriøsitet og professionalisme samt det engagement som skolerne går til opgaven med

## 6. Tragtmodellen

Som beskrevet i afsnit 2.1 er testbatteriet bygget op på baggrund af en frasorteringsmekanisme bestående af fire dele, kaldet tragtmodellen. De 4 tragte eller trin benævnes 1) Observationsguiden, 2) Den digitale test, 3) Samtaletesten og 4) PPR.

Tragtmodellens oprindelige design er udviklet forud for indeværende projekt, og af hensyn til projektets praktiske gennemførelse har det været nødvendigt at lave nogle tilpasninger af det oprindelige design til dette projekts formål.

### 6.1 Det oprindelige design: Tragtmodellens fire trin

1. **Observationsguiden** er en vejledning til lærerobservation og et registrerings-skema med 39 udsagn, der besvares for hver elev. Ideen er, at afgrænsningen er så bred, at alle i risiko for talblindhed bliver fanget ind. Herved identificeres en bruttogruppe, som i rapporten anslås til **15 pct. af en årgang**, der har problemer med matematik af forskellige årsager. Disse 15 pct. gennemfører den digitale test, der identificerer dem, der har målbare tegn i deres besvarelse, som peger på talblindhed.
2. I **den digitale test** er der fem forskellige opgavetyper (sub-konstrukter):
  - a. Sammenligning af to sæt prikker, hvor er der flest?
  - b. Sammenligning af to tal, hvilket er størst?
  - c. I 10 sæt prikker: Marker de sæt, hvor der er fire prikker.
  - d. Sammenligning af antal prikker med et tal, hvad er størst?
  - e. Identificér samlet antal prikker: Hvor svarer det til et angivet tal?
3. Testtageren besvarer opgaver af hver type i 2 minutter – i alt 10 minutter. Ved den første opgavetype vises de to sæt prikker ganske kort (for at forhindre eleven i at tælle), men der er ikke andre tidsrestriktioner på opgavebesvarelsen end de 2 minutter pr. opgavetype. Ved denne digitale test forventes det, at der kan identificeres yderligere 5 procentpoint, der kan udelukkes, så de potentielle elever med talblindhed **afgrænses til 10 pct. af en årgang**.
4. **Samtaletesten** gennemføres for disse 10 pct. af en årgang. Den består af 19 opgaver, som afprøves individuelt med en testafholder (lærer/matematikvejleder) over ca. 30 minutter. Samtaletesten kan afsløre procesproblemer, som ikke fanges af den digitale test (fx tøven eller tællen på fingre), hvorved gruppen af elever, der muligvis har talblindhed, kan reduceres til **3 pct. af en årgang**.
5. **PPR** skal herefter identificere dem, der har den rene talblindhed, ved gennem en pædagogisk psykologisk udredning at udelukke elever med andre eller flere problemer. I projektet skønnes det, at ca. halvdelen af de 3 pct. har generelle indlæringsvanskeligheder og derfor ikke er i målgruppen. Herefter resterer **1-2 pct. af en årgang**, som har talblindhed uden andre vanskeligheder.

### 6.2 Den praktiske udformning af afprøvningen

**Udvælgelse af elever:** Alle projektets godt 800 skoler blev bedt om at udvælge de 4.-klasseselever, der er mest udfordrede i matematikfaget. De kunne anvende observationsguiden på alle elever i klassen eller tage afsæt i de elever, som i de nationale test (DNT) i matematik scorer "jævn" eller derunder, hvilket identificerede *en bruttomålgruppe* svarende til de 25 pct. lavest præsterende elever. De, som ikke havde aktuelle DNT-resultater, kunne anvende en forlagsfremstillet prøve. Denne udvælgelse baseres på en vurdering af, at det er inden for denne gruppe, at man ville finde så godt som alle elever med talblindhed.

Tilmeldingen til projektet indebar, at skolen gav tilsagn om at lade den digitale test anvende af alle disse elever. Skolen kunne frit vælge om og hvor mange elever, der skulle anvende observationsguiden og samtaletesten. Dette gjaldt dog ikke kerneskoler, som på forhånd havde givet tilsagn om at anvende alle redskaber på alle elever.

**Tilpasning og videreudvikling af testen:** Erfaringer fra den indledende afprøvning i udviklingsprojektet gav anledning til justeringer af virkemåden for særligt den digitale test – herunder re-design af items for at kunne måle svarhastighed samt integration af videovejledninger i testen. Yderligere blev designet af både den digitale test og samtaletesten drøftet på ekspertgruppemødet og anbefalinger hertil blev implementeret. På det grundlag identificeredes elever, hvor en samtaletest kunne give mening.

## 6.3 Trin 1: Observationsguide

Observationsguiden har i den oprindelige udgave af tragtmodellen til formål at støtte læreren i udvælgelsen af elever, der er relevante at teste yderligere for talblindhed. Observationsguiden skal identificere den gruppe elever (anslået 15 pct. af en årgang), der har matematikvanskeligheder, og som det derfor anbefales at lade gennemføre den digitale test.

**Figur 4: Del af spørgsmålsbatteriet i observationsguiden**

?
OBS: Du kan tilføje flere elever, når du har udfyldt tjeklisten for den første elev.
✕

**Observer/tjek om eleven...**

	Aldrig eller sjældent	Af og til	Ofte eller altid	Ved ikke
<b>TAL OG TÆLLING</b>				
.. har svært ved at tælle antal større end 10	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. ikke kan se et antal op til 4 uden at tælle	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved at anvende tal til at angive overslag på antal elementer i større mængder	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved at læse to- og trecifrede tal højt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved at ordne et- og tocifrede tal, herunder placere dem rimeligt på tallinje	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved at huske "vigtige" tal, fx eget husnummer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved at anvende titalssystemet, fx at 35 består af 3 tiere og 5 enere	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved ikke-kendte talfølger fx 1 - 3 - 5 - 7 eller 14 - 24 - 34 - 44	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. let lader sig forvirre ved læsning og skrivning af tal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. taster forkert ind på lommeregner	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved at tælle baglæns, fx fra 35 til 27	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<b>ADDITION OG SUBTRAKTION</b>				
.. har svært ved fortsat tælling	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved at fremkalde plustabellen inden for området 0 - 5 uden at tælle	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. mangler regnestrategier ved addition ud over fortsat tælling	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. mangler regnestrategier ved subtraktion ud over at tælle baglæns	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. har svært ved at vurdere om et resultat er næsten rigtigt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
.. anvender konsekvent tælling med brug af fingrene ved beregninger	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Den i projektet anvendte version af observationsguiden (en videreudvikling af observationsguiden fra udviklingsprojektet) består af 34 spørgsmål, der er opsat som et online spørgeskema. Som orientering til læreren er guiden også opsat som en tjekliste i PDF. Læreren logger ind med UNI-login

og angiver oplysninger om skole, sted (sidder de hjemme eller på skolen) og elevens UNI-login. Der indgår spørgsmål om:

- Tal og tælling
- Addition og subtraktion
- Multiplikation
- Orientering
- Tid
- Tal og størrelser i hverdagen
- Elevadfærd, oplevelser og holdning

### 6.3.1 Scoring og feedback

Når spørgsmålene er besvaret, beregnes automatisk en score og et resultat i form af en anbefaling. Scoren går fra 0 til 200, og jo højere score des større risiko for talblindhed. Beregningen sker således:

- Der skal svares på over halvdelen af spørgsmålene, før der kan beregnes en score. Er der svaret på under halvdelen af spørgsmålene, gives følgende resultat: **"Der er svagt grundlag for at vurdere eleven – hvis elevens testresultater i matematik i øvrigt er lave, bør du lade eleven tage den digitale test"**
- En score på 0 vil være et tegn på, at der ikke er tale om talblindhed. Resultat: **"Der er ikke grund til at formode eleven har talblindhed"**
- Ved en score på 1-99 vil der være grund til at overveje, om det skal undersøges nærmere (jo højere score des større risiko). Resultat: **"Hvis du selv har en mistanke om talblindhed, bør du lade eleven tage den digitale test"**
- En score på over 100 kan bestyrke en mistanke om talblindhed, som bør undersøges nøjere. Resultat: **"Du bør lade eleven tage den digitale test"**

En oversigt over besvarelser og resultat for den enkelte elev kan herefter hentes som PDF og gemmes eller printes. Principperne for scoring af observationsguiden er yderligere beskrevet i afsnit 8.1.

## 6.4 Trin 2: Den digitale test

Opsætning af den digitale test anvendes i tre forskellige sammenhænge: Opgaveafprøvning, demotest og den egentlige afprøvning. I tillæg hertil er tilføjet to non-kognitive skalaer med henblik på måling af hhv. motivation og selvtillid. Disse indgår ikke i selve testbatteriet og grundlag for vurdering af elevens talblindhed, men giver oplysninger som supplement til den digitale test.

### 6.4.1 Opsætning af den digitale test

Opsætningerne er ens bortset fra antallet og rækkefølgen af opgaverne i testen. Opsætningen er tilrettelagt, så der indledes med en kort forklaring, der logges på med UNI-login, der vises en lille introvideo til hver af de fem deltest, og herefter kan eleven gå i gang med besvarelsen. Skalaerne til de fem deltest er:

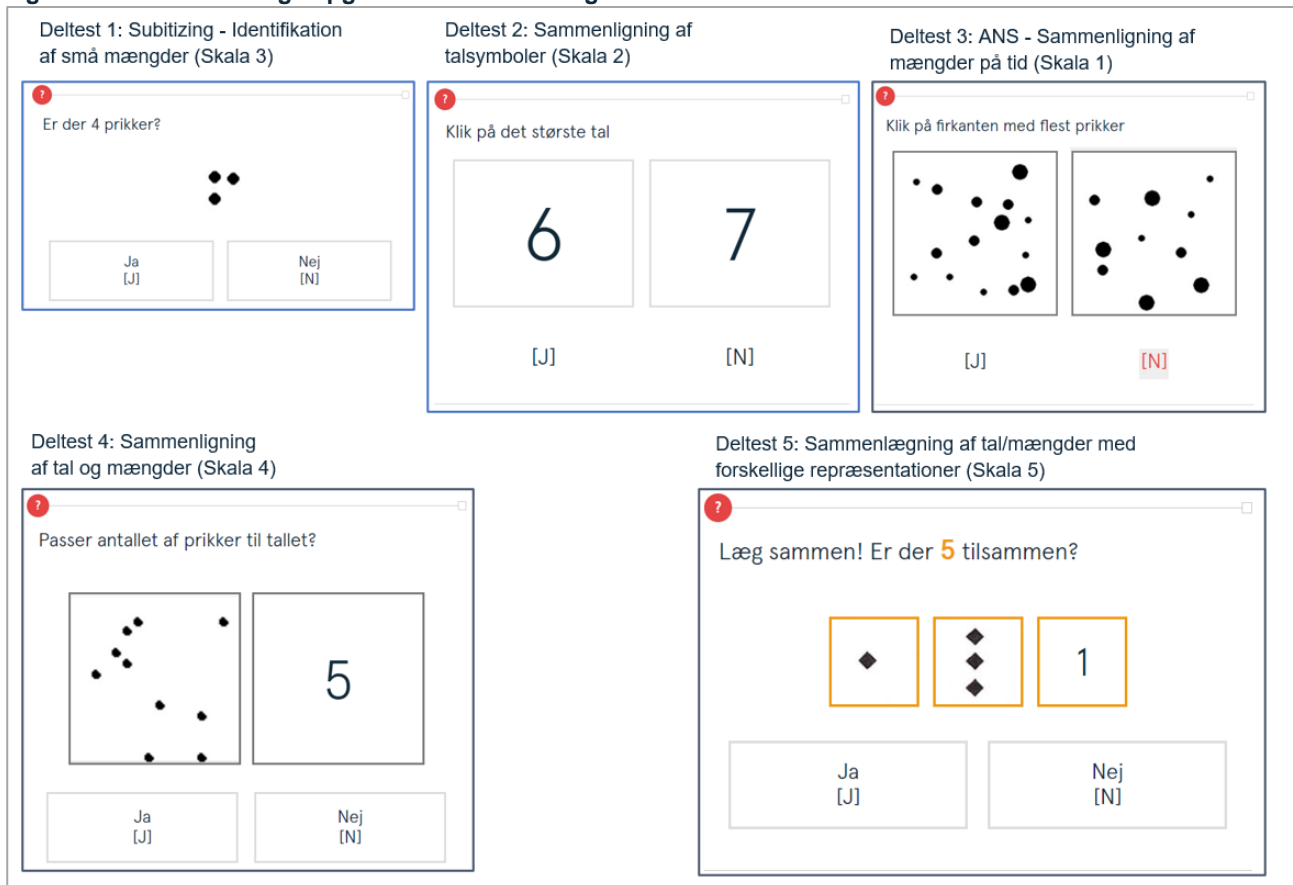
- Deltest 1: Subitizing – Identifikation af små mængder (Skala 3)
- Deltest 2: Sammenligning af talsymboler (Skala 2)



- Deltest 3: ANS – Sammenligning af mængder på tid (Skala 1)
- Deltest 4: Sammenligning af tal og mængder (Skala 4)
- Deltest 5: Sammenlægning af tal/mængder med forskellige repræsentationer (Skala 5)

Skalanummereringen er foretaget ud fra den oprindelige rækkefølge i den første afprøvning (som opsat til DPU/Absalon i 2015). Rækkefølgen i deltestene blev justeret i projektet (skala 1 og 3 byttede plads), da erfaringer viser, at det er bedst at få eleverne i gang med de lettest mulige opgaveformater. Af Figur 5 fremgår eksempler på de fem forskellige opgaveformater.

**Figur 5: De fem forskellige opgaveformater i den digitale test**



### 6.4.2 Opsætning til opgaveafprøvning

Ved opgaveafprøvningen starter eleven med tre lette opgaver i fast rækkefølge i hver deltest, så eleven kan få en overkommelig start. Herefter kommer opgaverne i randomiseret orden, hvor alle opgaver har samme udtrækssandsynlighed. Formålet hermed er at sikre en ligelig afprøvning af opgaverne. Eleverne kunne kun besvare testen én gang. Ved svigtende forbindelse, utilsigtet nedlukning af browser mv. kunne eleven logge på svarforløbet igen og fortsætte, hvor elevens forbindelse faldt ud.

### 6.4.3 Opsætning til demotest

Til brug for at gøre lærere og elever bekendte med test-formatet blev der lanceret en demotest. Den fungerer fuldstændig som den rigtige test med UNI-login og videoinstruktioner. Der er dog væsentlige forskelle, idet der kun er fem opgaver pr. skala, og der er mulighed for at tage testen flere gange.

## 6.4.4 Digital test, den egentlige afprøvning, runde 1 og 2

Her opsættes testen, så den fra elevens side fremstår ligesom ved opgaveafprøvningen og demotesten. Rent teknisk fungerer det, så der i hver skala startes med en begrænset række velfungerende opgaver (jf. afsnit 7.4) i fastlagt rækkefølge for alle elever, startende med én meget let og derefter med stigende sværhedsgrad til den sværeste opgave. Herefter kommer resten af opgaverne i skalaen i randomiseret orden.

Når testen af de fem skalaer er afsluttet, undersøges både i runde 1 og 2 elevernes motivation for og selvtillid i forhold til matematik og aritmetikfærdigheder gennem to mindre batterier af spørgsmål.

## 6.4.5 Motivation og selvtillid

Det er pointeret i mange sammenhænge, at elevernes motivation for og selvtillid i forhold til matematik og aritmetikfærdigheder kan spille en væsentlig rolle for deres udvikling. Derfor inddrages en analyse heraf i den digitale test. Non-kognitive færdigheder som motivation og selvtillid kan undersøges på flere måder.

Dels kan man spørge *direkte* (fx "Hvor motiveret er du for matematik?"), hvilket har den klare fordel, at det er en meget lettilgængelig undersøgelsesmetode. Ved effektvurdering kan man spørge retrospektivt (fx "Er du blevet mere motiveret?" eller "Har du fået mere selvtillid?"). Det kan have den pædagogiske fordel, at respondenterne gennem indsatsen bliver opmærksomme på sit eget udviklingspotentiale, hvilket kan opleves som øget selvtillid og motivation. Måleteknisk er problemet med denne måde at spørge på, at verden ofte ser væsentligt anderledes ud i bakspejlet, og opgørelserne bliver subjektive og usammenlignelige.

**Tabel 7: Spørgsmål til vurdering af indre motivation og selvtillid i matematik, regne- og talfærdigheder**

Indre motivation (6 items)	Selvtillid (6 items)
<b>Instruktion:</b>	
<i>Tænk på undervisningen og opgaverne i matematik og regning. Er du enig eller uenig i følgende:</i>	
Jeg kan godt lide svære regne- og talopgaver, så jeg kan lære nye ting	Jeg tror, at jeg vil klare mig rigtigt godt i regning og matematik
Jeg kan godt lide regne- og talopgaver, der gør mig nysgerrig, også selvom det er svært	Jeg er sikker på, at jeg kan forstå selv de sværeste ting i regning og matematik
Jeg kan godt lide regne- og talopgaver, jeg lærer meget ved	Jeg er sikker på, at jeg kan lære det matematik, jeg bliver undervist i
De mest tilfredsstillende opgaver i matematik er dem, som jeg udvikler mig mest ved	Jeg tror, at jeg vil få et rigtigt godt resultat i en matematiktest
Jeg kan bedst lide regne- og talopgaver, som jeg lærer af, også selvom jeg ikke får et godt resultat	Jeg er sikker på, at jeg vil klare mig godt i regne- og talopgaverne i skolen
Jeg kan bedst lide meget svære regne- og talopgaver, så jeg kan lære rigtigt meget	Jeg er sikker på, at jeg kan få bedre matematikresultater end de fleste andre elever

Det, vi gør i denne undersøgelse, er at spørge til *en række indikatorer* på denne færdighed (fx "Jeg kan godt lide opgaver, jeg lærer meget ved") og efterfølgende analysere svarene og beregne en skalascore. Fordelen ved denne metode er, at de scorer, der beregnes for de enkelte elever, er ab-

solutte og sammenlignelige, samt at skalascoren kan bruges til at vurdere udvikling. Dette er særligt interessant i forbindelse med effektvurdering i forhold til en indsats ved før- og eftermåling med de samme spørgsmål.

Man skal dog være opmærksom på, at det ofte vil give helt andre resultater, end hvis der spørges retrospektivt til en direkte vurdering af effekten. Af Tabel 7 fremgår de spørgsmål, der lægges til grund for beregning af scorer i indre motivation og faglig selvtillid. Eleven bliver instrueret i at tænke på undervisningen og opgaverne i matematik og regning. Dernæst bliver eleven spurgt, om han/hun er enig eller uenig i en række indikatorudsagn med svarmulighederne *Helt enig, Nærmest enig, Hverken enig eller uenig, Nærmest uenig og Helt uenig*.

Som grundlag for spørgsmålene vedr. motivation og selvtillid benyttes et tidligere valideret dansk-sproget spørgebatteri<sup>24</sup> rettet mod generel motivation og selvtillid. Dette batteri er valideret på danske elever fra 4.-8. klasse. I denne udgave er spørgsmålene – i samråd med forskerne bag batteriet – blevet målrettet *motivation og selvtillid i matematik, regne- og talfærdigheder*.

I det originale batteri var der seks indikatorspørgsmål (items) til måling af hver af de to færdigheder (indre motivation og faglig selvtillid) og deres overordnede tilgang er videreført i det tilpassede batteri, der anvendes i denne sammenhæng.

## 6.5 Trin 3: Samtaletest

På baggrund af resultaterne fra den digitale test vil læreren/vejlederen få en indikation af, hvilke elever det er relevant at gennemføre samtaletesten med. I samtaletesten svarer eleven på en række opgaver, men det er læreren, der besvarer selve spørgeskemaet på baggrund af en vurdering af elevens svar på opgaverne og eventuelle vanskeligheder ved opgaveløsningen. Fx indgår opgaver, hvor læreren skal observere elevens adfærd og reaktioner (fx om eleven tæller på fingrene).

### 6.5.1 Opsætning af samtaletesten

Samtaletesten er opsat som et online opgavesæt, hvor læreren som det første logger ind med UNI-login. Ligesom ved observationsguiden er der en introduktion til læreren på første side. Læreren angiver desuden oplysninger om skole og elevens UNI-login. Læreren skal også svare på, om vedkommende sidder hjemme eller til stede fysisk på skolen, og på den baggrund kan hentes en PDF til elevens opgavesæt direkte i samtaletesten. Der er to udgaver, som benyttes afhængigt af, om læreren sidder fysisk sammen med eleven (19 opgaver/49 items) eller fysisk adskilt (18 opgaver/41 items).

I samtaletesten er der en side for hver opgave. Til højre vises opgaven fra elevmaterialet, og til venstre er lærerens instruktion – herunder hvad læreren skal være opmærksom på, hvad der skal læses højt for eleven samt vejledning til at besvare spørgsmålet.

---

<sup>24</sup> Makransky, G., Wandall, J., Madsen, S. R., Hood, M., & Creed, P. (2020). Development and Validation of the UiL-Scales for Measurement of Development in Life Skills—A Test Battery of Non-Cognitive Skills for Danish School Children. *Scandinavian Journal of Educational Research*, 64(4), 612-627.

Figur 6: Eksempel på opgave i samtaletesten

OBS: Dette er lærerens visning – bør ikke deles med eleven. Kopi af elevark

?

*Opgave 5: POSITIONSSYSTEM OG ORDNING*

**Instruktion:**  
Hvis eleven ikke forstår pointsystemet, må det gerne forklares yderligere.

**Læs op:**  
"Her er en skydeskive. Man får point efter hvor man rammer. De røde prikker viser, hvor Viktor har ramt. Hvor mange point har Viktor fået i alt?"

**Scoring:**  
Sæt kryds ud for det, som passer bedst til elevens svar.  
Skriv evt. en kommentar til besvarelsen.

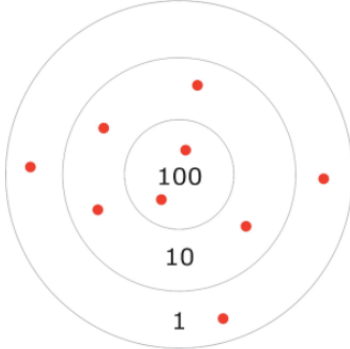
	Sæt ét X
Eleven svarer korrekt: 243	<input type="radio"/>
Eleven forstår princippet, men svarer forkert fx pga en tællefejl	<input type="radio"/>
Eleven svarer ikke eller uden at tage hensyn til prikkernes værdier	<input type="radio"/>

Kommentar:

Tilbage
Næste

opgave 5

Hvor mange point har Viktor fået i alt?



## 6.5.2 Rapportering af resultater

I modsætning til den digitale test indgår rapportering af resultater som en del af opsætningen ved samtaletesten (ligesom ved observationsguiden). Efter gennemførelse af testen vises en resultat-side med oversigt over alle besvarelser. Nederst på siden fremgår start- og sluttidspunkt og samlet tid brugt på testen. Endeligt beregnes automatisk et resultat (ligesom ved observationsguiden). Følgende indikatorer vises:

- Score (antal rigtige)
- Antal opgaver besvaret
- Max-score (max antal rigtige)

Beregningen af resultat sker som følger:

- Scoren er ikke gyldig, hvis færre end 50 pct. af spørgsmålene er besvaret: **"Der er ikke nok data til at give feedback"**
- Hvis gyldig score og over ca. 50 pct. rigtige: **"Der ser ikke ud til at være grundlag for mistanke om talblindhed"**
- Hvis gyldig score og under ca. 50 pct. rigtige: **"Der kan være mistanke om talblindhed"**

Oversigten med alle besvarelser og resultatet for den enkelte elev kan hentes som PDF og gemmes eller printes. Eksempel er vist i figuren nedenfor.

Figur 7: Nederste del af resultatside fra samtaletesten

13	Du skal regne denne opgave. Forklar, hvordan du regner. (62 - 28)	0	1	1
14	Oskar vil stege bøf til 10 mennesker. En pakke kød er nok til 4 mennesker. Hvor mange pakker kød skal Oskar købe?	1	1	1
15	Camilla og Hugo har 12 kroner tilsammen. Camilla har 5 kroner. Hvor mange penge har Hugo?	0	1	1
16	Du skal gætte, hvor lang pinden er	0	1	1
17	Gæt hvor lang badebroen er.	1	1	1
18	Du skal sige, hvor mange penge der er på billedet.	0	1	1
19	Du skal sige, hvad klokken er på de tre ure	2	3	3
	<b>Sum af kolonne</b>	<b>22</b>	<b>49</b>	<b>49</b>
	<b>Andel rigtige besvarelser</b>	<b>45%</b>		

Besvarelse start: **08:08:29**

Besvarelse slut: **08:45:59**

Tidsforbrug (brutto): **37 minutter 30 sekunder**

Resultat: **Der kan være mistanke om talblindhed**

Forklaring af resultatberegning	Scoren er ikke gyldig, hvis færre end 9 opgaver eller færre end 25 spørgsmål er besvaret: <b>Der er ikke nok data til at give feedback</b>
	Hvis gyldig score og over 50% rigtige: <b>Der ser ikke ud til at være grundlag for mistanke om talblindhed</b>
	Hvis gyldig score og under 50% rigtige: <b>Der kan være mistanke om talblindhed</b>

**PRINT**

Det var indledningsvist anslået, at hvis over halvdelen af opgaverne var besvaret korrekt, ville det svare til, at eleven ikke kunne anses for at være i risikozonen. Efterfølgende ville der imidlertid skulle beregnes en mere præcis cutscore. Det viste sig dog, at de 50 pct. var tæt på det teoretisk korrekte resultat.

## 7. Opgaveafprøvning – digital test

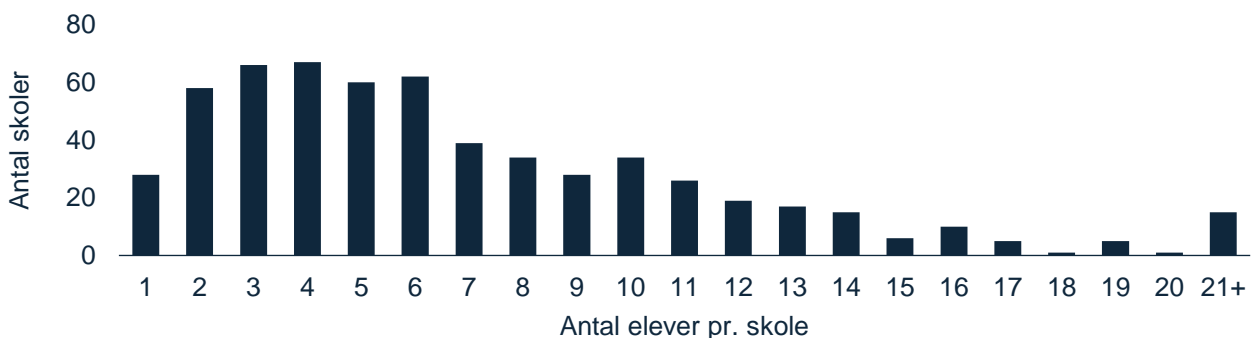
I det oprindelige design var det hensigten, at opgaveafprøvningen skulle ske i forbindelse med første gennemløb af testen (runde 1). Grundet de forsinkelser som Covid-19 medførte, blev det imidlertid nødvendigt at udskille opgaveafprøvningen som en særskilt dataindsamling.

### 7.1 Opgaveafprøvning på pilot-skoler

Der blev afprøvet i alt 556 opgaver fordelt på 5 skalaer. I pilotafprøvningen af opgaverne deltog 597 skoler med i alt 4.288 elever, der har gennemført afprøvningen af alle 5 skalaer. Som det fremgår af Figur 8 deltog et begrænset antal elever fra langt de fleste pilot-skoler (godt 7 elever pr. skole i gennemsnit – altså 2-3 elever i gennemsnit pr. klasse). Enkelte respondents svar er blevet kasseret – det drejer sig om besvarelser, der ikke kom fra elever i 4. klasse eller fra lærere, samt besvarelser hvor afprøvningen ikke blev fuldført.

Pilot-skolerne modtog en e-mail med en beskrivelse af testforløbet d. 8. oktober 2020. Selve pilot-testen startede d. 19. oktober 2020, hvor skolerne modtog en invitationsmail med et link, hvorfra undersøgelsen kunne tilgås. Der blev sendt en påmindelse til de skoler, hvor ingen elever endnu havde deltaget i testen d. 2. november 2020 samt en sidste påmindelse med information om en forlængelse af svarfristen d. 6. november. Undersøgelsen lukkede endeligt for svar d. 16. november 2020.

Figur 8: Antal elever (i alt 4.288) på skoler (597), der deltog i afprøvningen af opgaver til den digitale test



### 7.2 Analyser af data fra opgaveafprøvning

Den analysemodel, som anvendes til at beregne elevernes dygtighed (beskrives i afsnit 7.3), kan være følsom over for, at nogle elever i hele eller dele af skemaet gætter tilfældigt på et svar fremfor at svare efter bedste evne. Vi har derfor anvendt en særlig fremgangsmåde til at identificere og slette besvarelser, som med en vis sandsynlighed er gæt.

Vi identificerer opgaver, hvor andelen af korrekt svar nærmer sig 50 pct. for de allerhurtigste elever og sletter dernæst svarene fra disse meget hurtige elever. På den måde udnytter vi to særlige kendetegn ved gæt: 1) At elever som gætter kan svare på et spørgsmål meget hurtigt, fordi de ikke behøver dvæle ved spørgsmålet, og 2) at et gæt vil have lige så stor sandsynlighed for at være rigtigt som forkert (50 pct.), fordi alle opgaverne kun har to svarmuligheder.

Helt nøjagtigt slettes de 3 pct. hurtigste besvarelser fra opgaver, hvor andelen af korrekte svar (for de 3 pct. hurtigste) ligger på mellem 45 og 55 pct. Det vil altså sige, at 3 pct. af de afgivne besvarelser på hver af disse opgaver slettes. Samlet set blev 6.211 responser fordelt på 1.357 elever og

100 opgaver slettet. Det svarer til 0,5 pct. af de afgivne svar, og der er således tale om en ubetydelig del af det samlede datamateriale.

Halvdelen af de opgaver, hvor vi identificerede gætteadfærd, ligger indenfor deltest 1, som handler om sammenligning af to sæt prikker (jf. Figur 5). I denne opgavetype går det ud på at identificere det største antal prikker af varierende placering og størrelse. Efter ganske kort tid (1-2 sekunder) forsvinder prikkerne. Eleven skal nu klikke på den side af skærmen, hvor der var det største antal. Noget tyder på, at denne opgavetype inviterer nogle elever til hurtige og ukvalificerede gæt. En mulig forklaring kan være, at nogle elever bliver frustrerede over ikke at kunne opfatte svaret og derfor på et tidspunkt opgiver at prøve seriøst.

Også i deltest 5 er en betydelig andel responser blevet kasseret. Det kan hænge sammen med, at denne deltest er den sværeste og sidste, hvor enkelte elever kan være blevet trætte og/eller ukoncentrerede.

**Tabel 8: Opgaver med gætteadfærd fordelt på deltest**

Deltest	Antal opgaver	Heraf opgaver med ukvalificeret gæt <sup>(1)</sup>	Andel opgaver med ukvalificeret gæt
1	164	50	30%
2	103	2	2%
3	160	11	7%
4	63	17	3%
5	66	20	30%
I alt	556	100	18%

**Note:** Opgaver besvaret med ukvalificerede gæt er opgaver, hvor 45-55 pct. af de 3 pct. hurtigst afgivne besvarelser er korrekte.

Der er tendens til, at enkelte elever er ansvarlige for en større andel af de samlede besvarelser. De 68 elever, som gættede ukvalificeret ved mere end 20 opgaver, afgav 35 pct. af de svar, som vi identificerede som ukvalificerede gæt. Der er med andre ord nogle elever, som har svaret meget hurtigt på en stor del af de udvalgte opgaver. Det betyder dog ikke, at disse elever som strategi klikker sig hurtigt igennem hele testen. Der er således ikke tale om, at vi her har slettet hele elevbesvarelser: Den elev, hvor flest svar blev klassificeret som ukvalificerede gæt, fik slettet ca. 15 pct. af sin samlede besvarelse. Der er tale om en finjustering, som måske ikke rykker meget på elevniveau, men som er vigtig, når det drejer sig om vurderingen af de enkelte opgavers egenskaber.

Der er også mange elever, for hvem gættene snarere må siges at være udtryk for enkelte hændelser. Således afgav 1.048 elever mellem 1 og 4 gæt i løbet af afprøvningen.

**Tabel 9: Slettede besvarelser**

Antal slettede svar	Antal elever	Samlet antal slettede svar	Andel af det samlede antal slettede svar
1-4	1.048	1.668	27%
5-9	136	885	14%
10-14	63	747	12%
15-19	42	718	12%
20-29	34	785	13%
30-39	12	392	6%
40-79	22	1.016	16%
I alt	1.357	6.211	100 %

De 1.048 elever i første række har hver afgivet mellem 1 og 4 svar identificeret som ukvalificerede gæt. Det svarer til i alt 1.668 slettede svar og udgør 27 % af det samlede antal slettede svar.

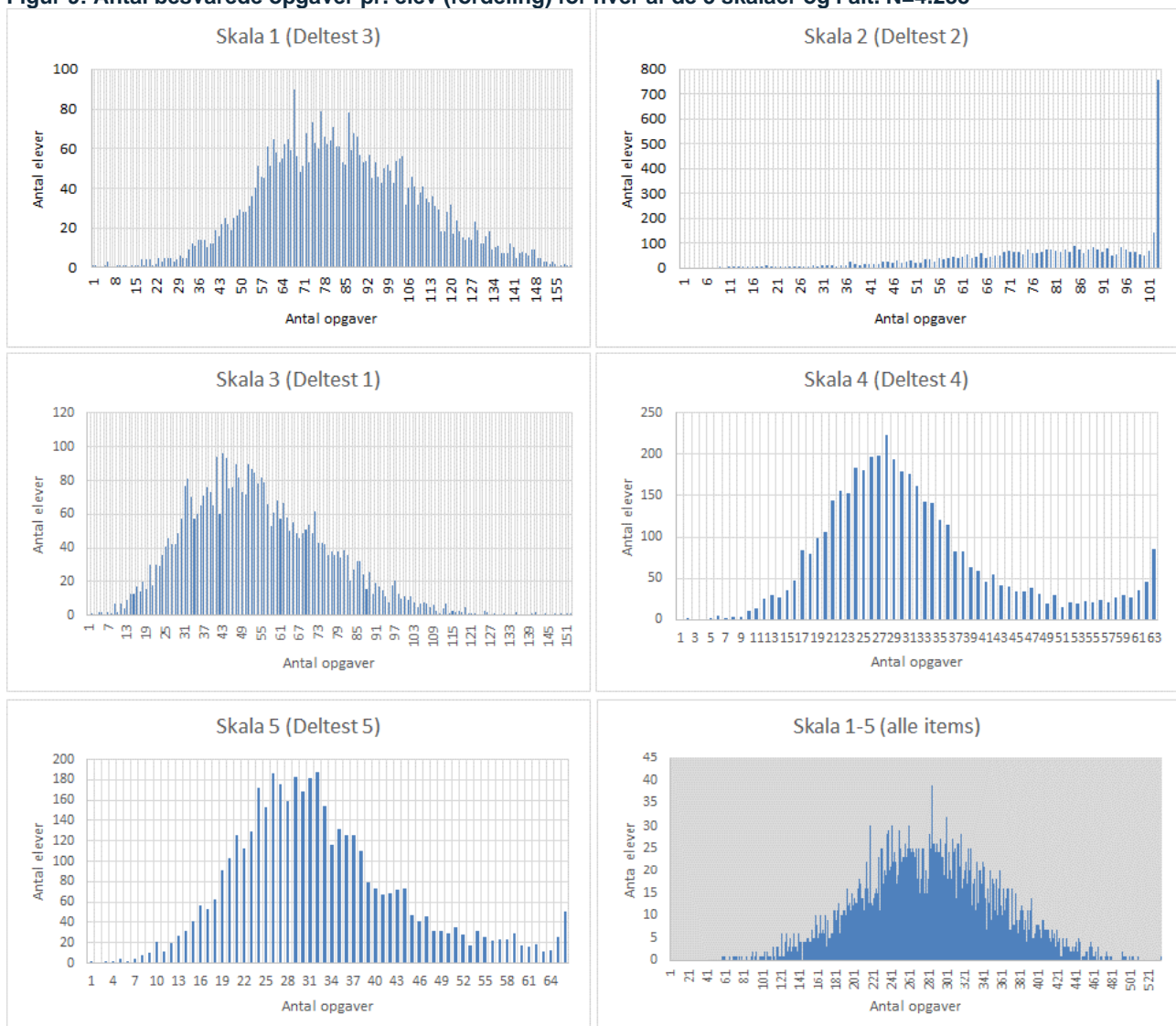
Efter frasortering af ukomplette (dårlige) responser kan de 4.288 elevers afprøvning af de 556 items opgøres, som det fremgår i Tabel 10, hvor antallet af items og deres besvarelse er opgjort. Der er i gennemsnit lidt over dobbelt så mange opgaver til rådighed, som der er behov for. Kun på skala 2, hvor opgaverne gennemgående er meget lette og hurtige at besvare, bliver opgaverne ofte besvaret alle sammen.

**Tabel 10: Antallet af items og generel information om besvarelsen ved afprøvningen**

	Skala 1	Skala 2	Skala 3	Skala 4	Skala 5	I alt, de fem skalaer
Antal items	164	103	160	63	66	556
Antal besvarede pr. elev	84	80	54	32	32	281
Antal rigtige pr. elev	74	77	52	28	29	260
Andel rigtige pr. elev	88%	96%	96%	91%	91%	92%
Item, som det største antal elever har afprøvet (antal elever)	2.227	3.344	1.524	2.128	2.002	3.344
Item, som det mindste antal elever har afprøvet (antal elever)	2.039	3.209	1.368	1.937	1.813	1.368

Fordelingen af antallet af besvarede opgaver pr. elev fremgår af Figur 9.

**Figur 9: Antal besvarede opgaver pr. elev (fordeling) for hver af de 5 skalaer og i alt. N=4.288**





## 7.3 To scoringsmodeller for hver skala

En væsentlig del af at have praktisk anvendelige matematikfærdigheder består i, at visse basale funktioner er automatiserede – det vil fx sige at kunne skønne over tal, antal og størrelser og sammenligne uden at skulle tælle. Mange af opgaverne er så lette, at alle (selv elever med svære matematikvanskeligheder og/eller talblindhed) vil kunne løse dem korrekt, hvis de blot bruger den fornødne tid. Den hastighed, med hvilken testens opgaver løses, kan dermed være en vigtig indikator på, hvor operationelle elevens færdigheder er, og derfor er hastigheden blevet registreret i forbindelse med besvarelsen.

For at kunne inddrage dette aspekt i modellen, vil der i det følgende blive arbejdet med to alternative scoringsmodeller for den samme besvarelse af opgaven. Scoringsprincipperne kan beskrives således:

1. Dikotom scoring (0=forkert, 1=korrekt)
2. Polytom scoring (0=forkert, 1=korrekt men langsomt, 2=korrekt og hurtigt).

Helt konkret anses en opgavebesvarelse for hurtig, hvis den besvares hurtigere end den valgte grænseværdi for svartiden på opgaven. Omvendt anses en besvarelse for langsom, hvis den besvares langsommere end denne svartid.

## 7.4 Raschanalyser – valg af opgaver til digital test

Der foretages to Raschanalyser<sup>25</sup> for hver skala, én for den dikotome og én for den polytome. På grundlag bl.a. heraf vælges det skala for skala, om det er den dikotome eller den polytome udgave, der danner grundlag for det videre arbejde. Udgangspunktet er, at det på et empirisk grundlag kan vurderes, om fordelene ved at inkludere svarhastigheden i scoringsmodellen opvejer ulemperne ved den forøgede kompleksitet, som dette medfører.

Problemet med en scoringsmodel, hvor hastighed inkluderes i scoringen, er, at respondenter, der (uden at gøre et reelt forsøg på at forholde sig til opgaven) blot satser på at besvare så hurtigt som muligt, statistisk set kan opnå et uforholdsmæssigt godt resultat. Dette gælder især ved anvendelse af items med kun to svarmuligheder, hvor scoren i så fald vil svare præcist til en fejlfri men langsom besvarelse. Analyser viser imidlertid, at en sådan besvarelsesstrategi er meget usædvanlig og kun i særlige tilfælde (ved relativt svære opgaver) kan give en overnormal score. Desuden er svar, der er skabt ved systematisk meget hurtig svarafgivelse (jf. afsnit 7.3), blevet markeret som ugyldige og scoringsmæssigt behandlet som hvis eleven ikke havde fået stillet opgaven.

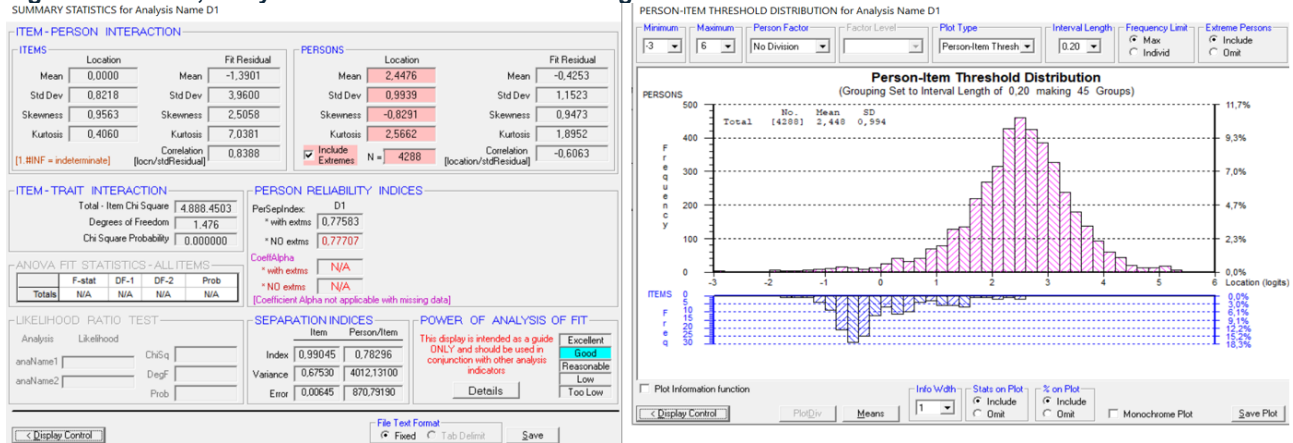
### 7.4.1 Skala 1: Sammenligning af mængder på tid

Gennemgangen af analysen vedr. skala 1 er detaljerig, og i denne sammenhæng forklares de grundlæggende principper og overvejelser. De overordnede resultater af den indledende analyse af den digitale version af skala 1 fremgår af Figur 10.

---

<sup>25</sup> Raschanalyse er en metode, der bygger på en statistisk model udviklet omkring 1960 af den danske matematiker George Rasch. Raschanalysen bruges til at opgøre elevens dygtighed (fx intelligencetest og faglige test) og test af holdninger/vurderinger baseret på kategoriserede data – typisk spørgsmål i en test eller et spørgeskema, der skal besvares ved afkrydsning. Kort sagt beskriver Raschmodellen sammenhængen mellem opgavernes sværhedsgrad og svarpersonernes dygtighed på samme skala. Raschmodellen tilhører familien af Item Respons Modeller, som i dag anvendes over hele verden i alle større undersøgelser af faglig dygtighed (fx PISA, TIMSS og PIRLS) og andre personlige egenskaber/færdigheder (motivation, selvtillid osv.).

**Figur 10: Skala 1, analyse af den dikotomt scorede udgave**



Reliabiliteten er god (PSI= 0,78), men opgaverne (de blåskraverede søjler i figurens højre side) er generelt lettere, end hvad der svarer til elevernes dygtighed (de rødskraverede søjler i figurens højre side). I andre sammenhænge ville dette være et betydeligt problem, da det betyder, at opgaverne gennemsnitligt er for lette til at kunne diskriminere sikkert mellem de dygtigste elever. Imidlertid er formålet med denne test at identificere tegn på talblindhed, hvilket forekommer blandt de lavest præsterende elever.

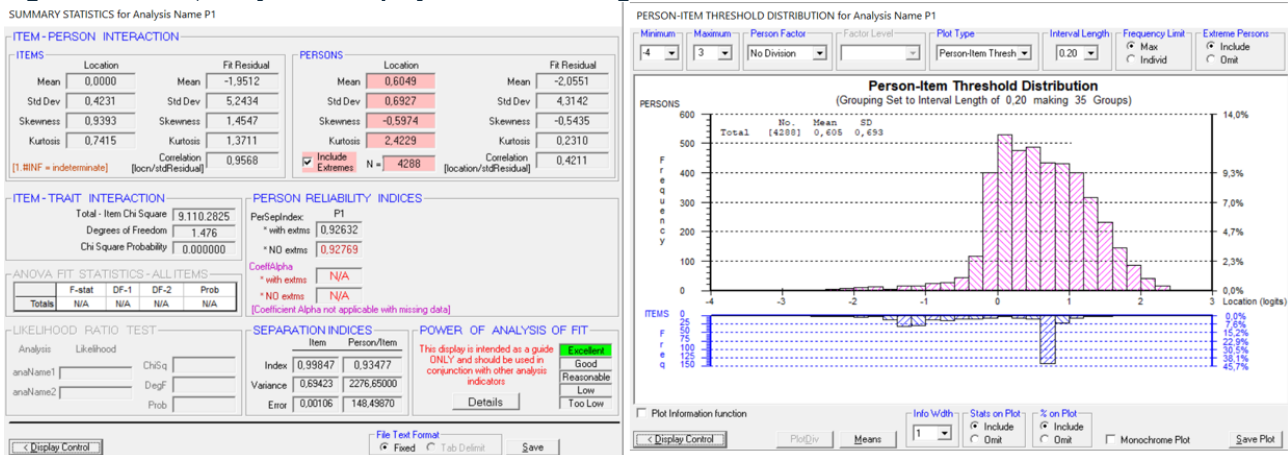
I skala 1 (deltest 3) vises to sæt af prikker i ét sekund, og eleven skal på grundlag af dette glimt-vise indtryk vurdere, hvor der er flest prikker (opgaveformatet ses i Figur 5, øverste højre hjørne). Det vil sige, at for skala 1 er elevens hurtighed i praksis allerede inkluderet i modellen i et vist omfang.

Det fremgår endvidere, at der er en pæn fordeling/spredning af eleverne, og at der er en del items med sværhedsgrader, der passer nogenlunde til den svageste del af eleverne.

Men spredningen i elevernes målte dygtighed (de rødskraverede søjler i de to figurer) er også reduceret i forhold til i Figur 10. Og dygtigheden der skal til at svare rigtigt og hurtigt på en stor del af opgaverne er klumpet noget sammen (den største blå-skraverede søjle).

Af Figur 11 fremgår analysen af den polytome scoring af skala 1. Her er reliabiliteten højere, hvilket for en stor dels vedkommende er en konsekvens af, at opgavernes sværhedsgrader (med inklusionen af tid i opgaven) i gennemsnit passer bedre til eleverne i populationen. Men spredning i elevernes målte dygtighed (de rødskraverede søjler i de to figurer) er også reduceret i forhold til Figur 10, og dygtigheden, der skal til at svare rigtigt og hurtigt på en stor del af opgaverne, er klumpet noget sammen (den største blå-skraverede søjle).

Figur 11: Skala 1, analyse af den polytomt scorede udgave



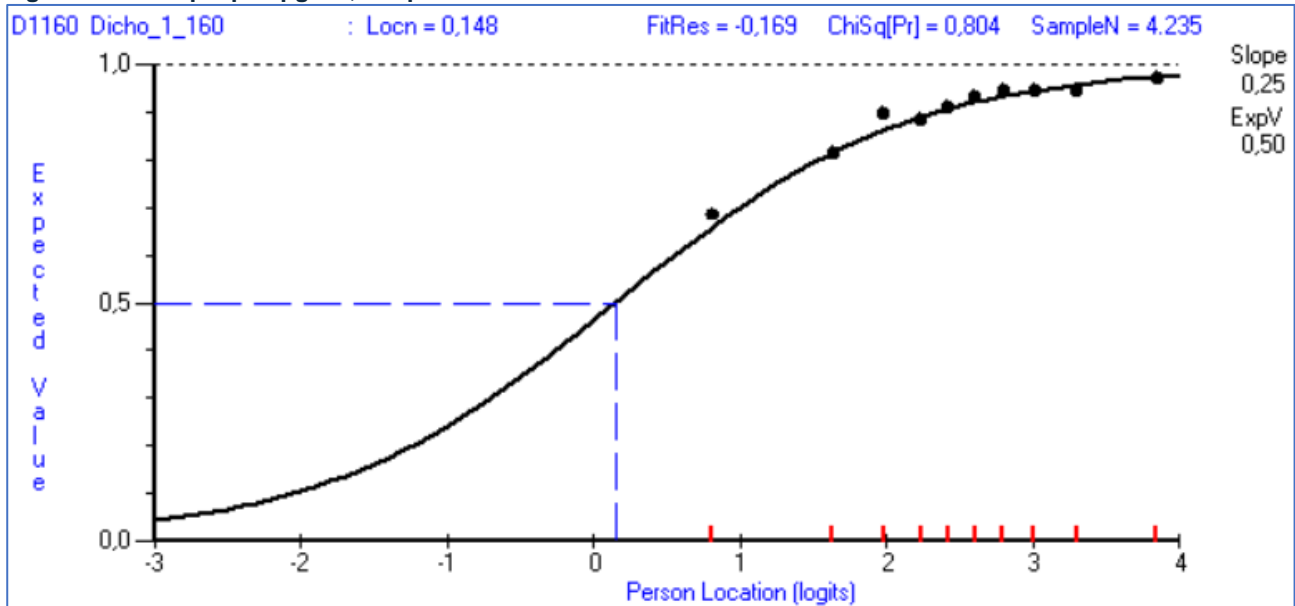
Da elevens hurtighed i praksis allerede er inkluderet i modellen i skala 1 (som den eneste af de 5 skalaer), ville det tilføje et element af redundans også at score opgaverne polytomt, og på den baggrund er det den dikotome scoringsmodel i skala 1, der arbejdes videre med.

Næste skridt er at undersøge, om der er opgaver, der passer dårligt til skalaen. Der fokuseres især på de besvarelser, der stammer fra elever i den lave ende af skalaen. I praksis gøres dette ved at se på fit-residualer og item-karakteristik-kurver.

Fit-residualer er udtryk for, hvor stor del af rest-variationen i besvarelsen af hvert enkelt item, der ikke kan forklares af modellen. Hvis fit-residualet er positivt, betyder det, at eleverne ikke har tilsvarende større sandsynlighed for at svare rigtigt på opgaven ved at være dygtig til det som alle opgaverne tilsammen måler (item underdiskriminerer). Sagt med andre ord: Opgaven passer ikke så godt til skalaen, som resten af opgaverne definerer den. Er fit-residualet negativt, overdiskriminerer item, og er der stærk negativ værdi af fit-residualet, kan man sige, at item fungerer bedre end raschmodellen foreskriver (item er bedre end det burde være til at skelne mellem stærke og svage elever).

Analysen af fit-residualer viser, at der er en række opgaver/items, der ligger uden for den kritiske værdi (konventionen siger, at fit-residualet helst skal ligge mellem +/- 2,5). Af Figur 12 er vist item-karakteristik-kurven for en opgave, der passer godt til modellen (den har et fit-residual meget tæt ved nul: FitRes = -0,169).

Figur 12: Eksempel på opgave, der passer til modellen



Sådan læses figuren:

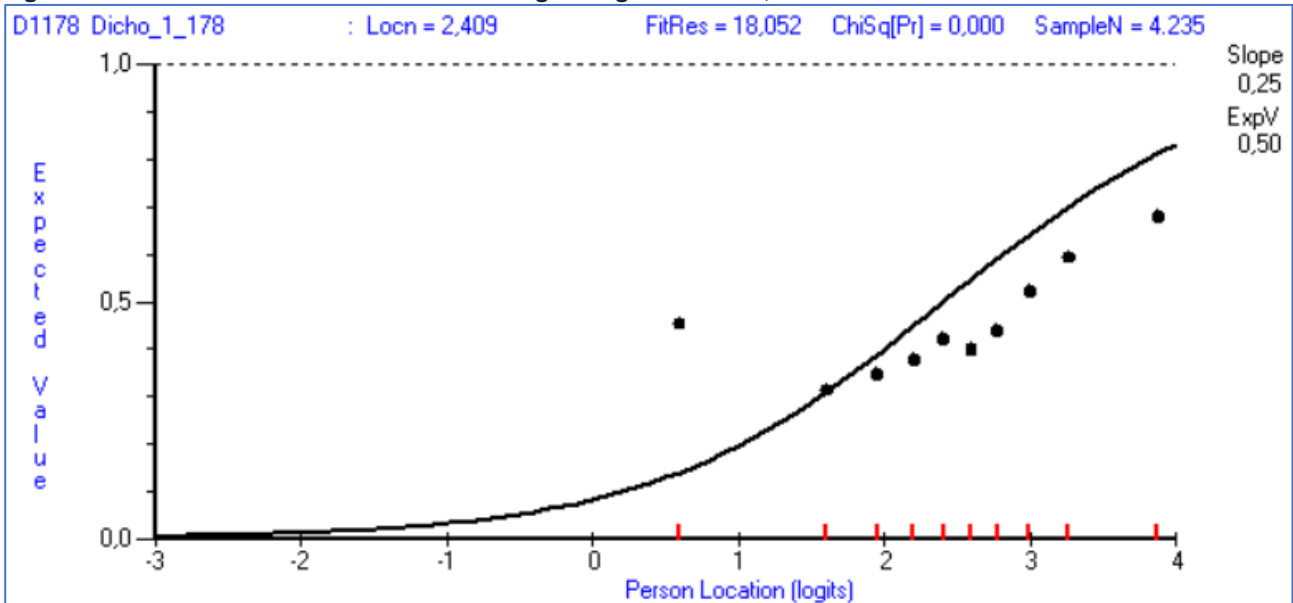
- **Person Location:** Den vandrette akse i figuren angiver elevernes samlede *dygtighed* (Person Location = placering på dygtighedsskalaen). Dygtigheden (Rasch-scoren) måles i logits som også er enheden for opgørelse af items sværhedsgrad. Nulpunktet på akse er defineret ved det midterste/middelsvære af de items, der har været brugt til målingen.
- **Expected value:** Den lodrette akse viser den forventede værdi ved besvarelsen af det konkrete item (dvs. sandsynligheden for korrekt svar). Kan også beskrives som det gennemsnitlige antal rigtige svar (*item-scoren*) for det konkrete item.
- **Klasseintervaller:** Eleverne opdeles i 10 lige store grupper (eller "klasser") efter stigende totalscore (derfor også kaldet "klasseintervaller"). De 10 røde markeringer på den vandrette akse viser altså den gennemsnitlige dygtighed (location) for eleverne i hver af klasseintervallerne.
- **Raschmodellen:** Den fuldt optrukne kurve viser Raschmodellens idealkurve.
- **De ti prikker** viser de 10 klasseintervalleres gennemsnitlige totalscore plottet mod de samme elevs gennemsnitlige item-score (gennemsnitligt antal rigtige).
- **De blå skraverede linjer** markerer dygtigheden for den elev, der efter modellen har 50 pct. sandsynlighed for at svare rigtigt på denne opgave (opgavens location = 0,148 logits).

Der er i alt 2.071 elever, der har besvaret opgaven i Figur 12. Hver af de 10 prikker repræsenterer altså ca. 207 elever. Prikken yderst til højre udgør de 10 pct. af eleverne med den højeste score, og af dem har 97 pct. svaret rigtigt på opgaven. Af de lavest præsterende 10 pct. (prikken til venstre) har 69 pct. svaret rigtigt.

Af Figur 13, nedenfor, er vist item-karakteristik-kurven for det dårligst fungerende item, dvs. den opgave, der har det største fit-residual i skala 1.

Ser man på den bedst præsterende elevgruppe (prikken længst til højre), er det også dem, der har den største sandsynlighed ( $P=0,68$ ) for korrekt svar på opgaven. Princippet i Raschmodellen er, at jo mindre dygtig eleven er, des lavere skal sandsynligheden være for at svare rigtigt på opgaven. Dette passer også nogenlunde frem til det næstsidste klasseinterval ( $P=0,31$ ), mens den svagest præsterende elevgruppe har en markant højere sandsynlighed ( $P=0,45$ ) end de lidt dygtigere elever for at svare rigtigt.

Figur 13: Item-karakteristik-kurven for det dårligst fungerende item, Skala 1



Det kan forekomme ulogisk, at de fagligt dårligst præsterende elever gennemsnitligt klarer sig signifikant bedre i denne opgave end de fleste af dem, der generelt (ud fra den samlede besvarelse af alle opgaverne) er dygtigere til at besvare opgaver. Den mest nærliggende forklaring på dette tilsyneladende paradoks er, at denne opgave (som er den sværeste på skalaen) snyder mange af de elever, der anstrender sig for at svare rigtigt, hvilket langt de fleste gør. Men blandt de 10 pct. af eleverne, der præsterer svagest generelt, er der måske en overvægt, der ikke gør en indsats, men svarer nærmest tilfældigt. Konsekvensen af, at der kun er to svarmuligheder er, at de elever, der ikke gør sig umage, rent statistisk vil ligge tæt på 50 pct. rigtige svar. Denne opgave ser således ud til at være så svær, at de svagest præsterende elever tilsyneladende "står af" på den og gætter. Og det vil netop være blandt de svageste elever, vi vil forvente at finde en relativt høj koncentration af talblinde elever.

En nærmere analyse viser, at der er en helt systematisk sammenhæng mellem items med højt fit-residual og items med høj sværhedsgrad. De opgaver, som raschanalysen kasserer, er altså de svære opgaver, hvilket normalt ville være et problem. Men i denne sammenhæng, hvor fokus er på at kunne skelne mellem de svagest præsterende elever, giver det faktisk god mening at se bort fra de sværeste opgaver. I alt 22 opgaver er kasseret pga. for høje fit-residualer (for ringe diskrimination), især i den lave ende af skalaen.

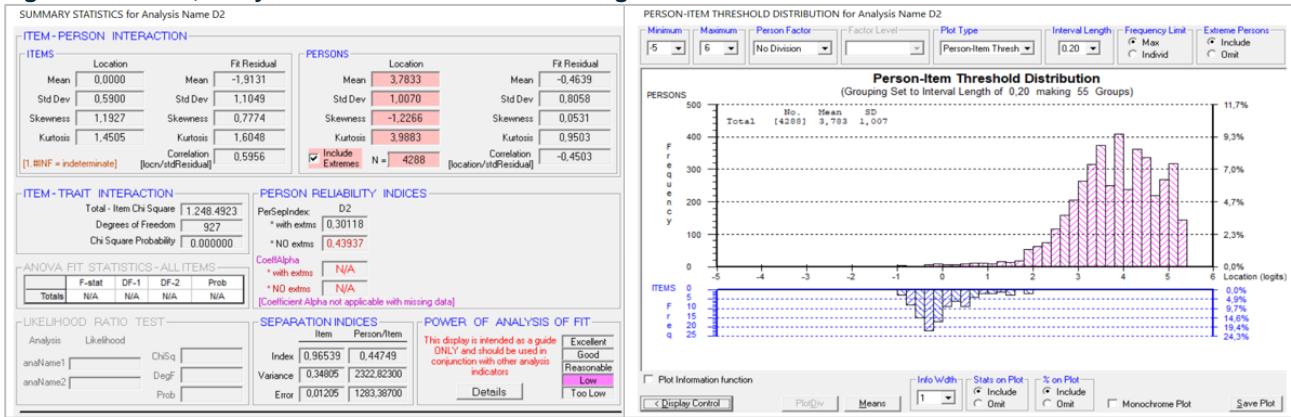
Herefter arbejdes videre med 142 items i skala 1, der i det påkrævede omfang passer til modellen.

### 7.4.2 Skala 2: Sammenligning af talsymboler

Skala 2 består af opgaver, hvor eleven skal klikke på det største af to tal (se Figur 5, øverst i midten). Dette er sammen med skala 3 den letteste af de opgavetyper, der indgår, og det kan være en udfordring i analytisk sammenhæng, at så mange (også mindre dygtige elever) svarer alt rigtigt. Her giver det god mening at inddrage tiden i scoringen for at kunne identificere forskelle mellem dygtige og mindre dygtige elever.

Af Figur 14 fremgår det, at da opgaverne er meget lette, selv i forhold til elevernes dygtighed, resulterer det i en meget lav reliabilitet (PSI=0,30), fordi stort set alle elever (næsten uanset dygtighed) svarer rigtigt på næsten alle opgaver.

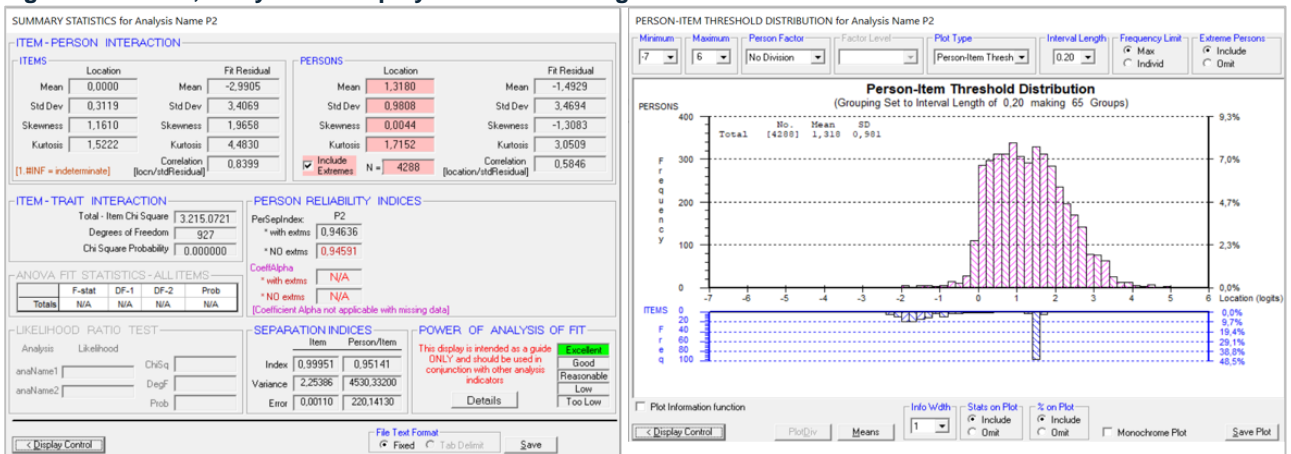
Figur 14: Skala 2, analyse af den dikotomt scorede udgave



Ved at introducere tiden i scoringen opnår man, at kun dem, der både svarer rigtigt og hurtigt, får en høj score.

Af Figur 15 fremgår resultaterne af analysen af den polytomt scorede udgave af skala 2. Her er der store forskelle ift. den dikotome scoring (Figur 14): Der er en meget høj reliabilitet (PSI=0,95), og langt flere elever (de rødskraverede søjler) ligger omkring 0 på location-aksen, hvor testen måler bedst.

Figur 15: Skala 2, analyse af den polytomt scorede udgave

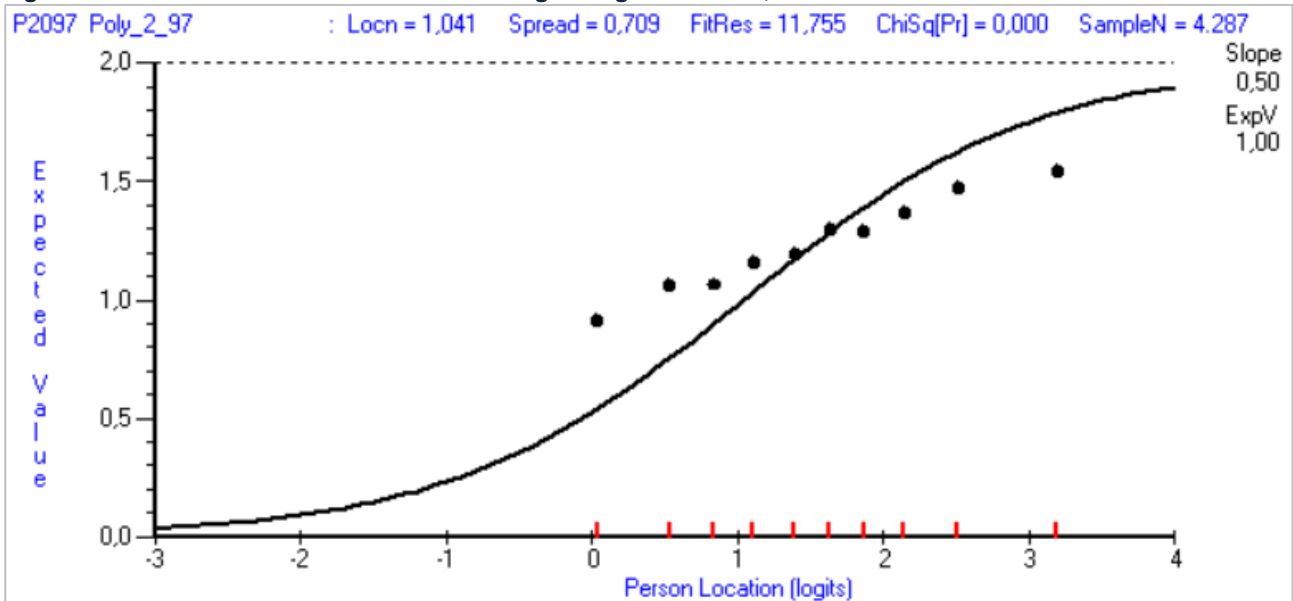


Analysen af data med den polytome scoringsmodel i Figur 15 (hvor tiden er inkluderet) viser altså, at skala 2 giver en mere statistisk pålidelig (reliabel) scoring.

Analysen af fit-residualer viser, at mens alle item holder sig inden for grænseværdien for fit-residualer (under 2,5) i den dikotome scoringsmodel, er der enkelte opgaver, der overstiger denne grænse i betydeligt omfang ved den polytome scoringsmodel.

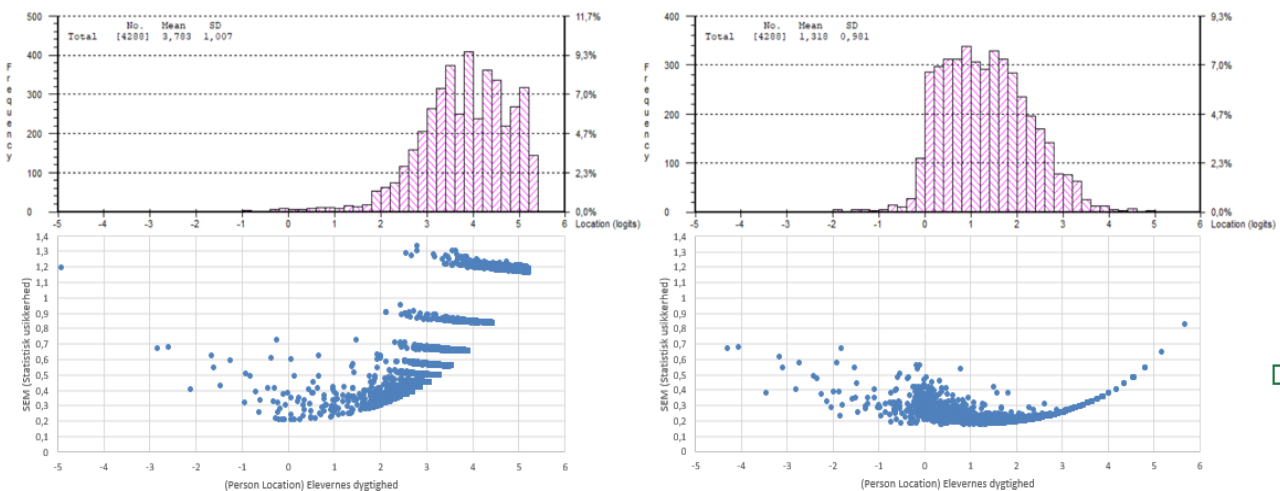
Den opgave, der afviger mest fra modellen, og som dermed er det dårligst fungerende item i skala 2, har en item-karakteristik-kurve, der er gengivet i Figur 16 nedenfor. Det fremgår, at opgaven diskriminerer væsentligt ringere, end modellen foreskriver over hele dygtighedsskalaen.

Figur 16: Item-karakteristik-kurven for det dårligst fungerende item, Skala 2



Når det drejer sig om at identificere talblindhed, er det i særlig grad de lavt præsterende elevers scorer, der skal være pålidelige, og for at illustrere hvordan den statistiske usikkerhed ser ud afhængigt af elevpræstationer, er dette illustreret i Figur 17.

Figur 17: Elevernes dygtighedsfordeling (øverst) og den statistiske usikkerhed (nederst) ved vurderingen af resultater for henholdsvis den dikotome (til venstre) og den polytome (til højre)



Som det fremgår, er SEM (den statistiske usikkerhed) betydeligt lavere ved den polytome scoringsmodel over hele spektret og ikke mindst i den lave ende af spektret.

Ved at sammenholde resultaterne fra analyser af data fra de to scoringsmodeller identificeres i alt 3 opgaver (heraf den der er vist i Figur 16), der generelt passer dårligt til modellen. Disse 3 opgaver, der altså ligger væsentligt uden for den kritiske værdi ved den polytome scoring, slettes. Herefter arbejdes videre med 100 items i skala 2, der i det påkrævede omfang passer til modellen.

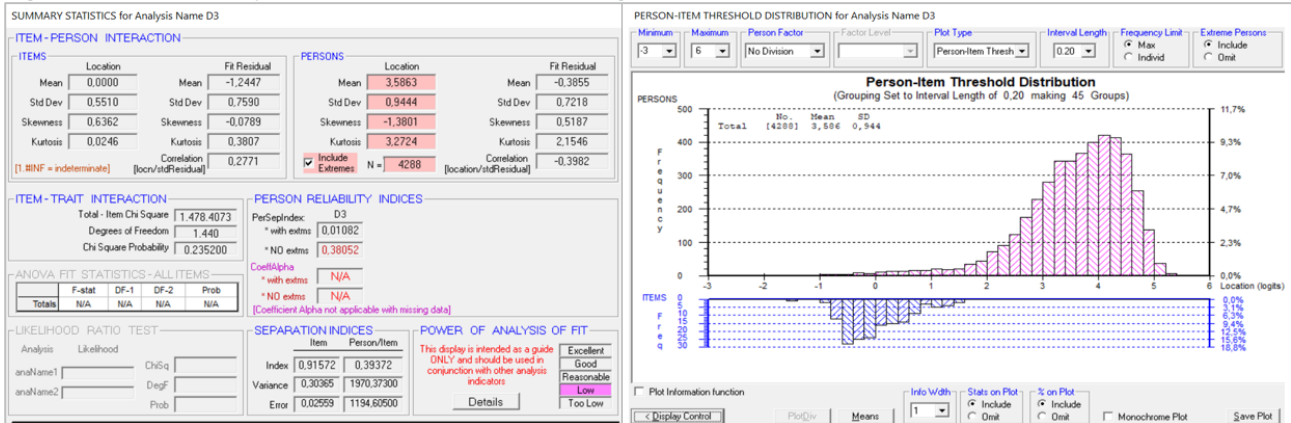
### 7.4.3 Skala 3: Identifikation af små mængder (Subitizing)

Skala 3 består af opgaver, hvor eleven skal svare "ja" eller "nej" til, at der er præcis fire prikker (se Figur 5, øverst til venstre). Denne og skala 2 er de letteste af de opgavetyper, der indgår i den digitale test, og mange (også mindre dygtige elever) svarer alt rigtigt. Her giver det (ligesom ved skala

2) god mening at inddrage tiden i scoringen for bedre-ITEM at kunne identificere forskelle mellem dygtige og mindre dygtige elever.

Af Figur 18 fremgår det, at da opgaverne er meget lette, selv i forhold til elevernes dygtighed, resulterer det i en lav reliabilitet. Dette fordi stort set alle elever svarer rigtigt på næsten alle opgaver. Bortset fra dette ser der ud til at være en overordentlig fin tilpasning til skalaen (Chi2-sandsynlighed=0,23).

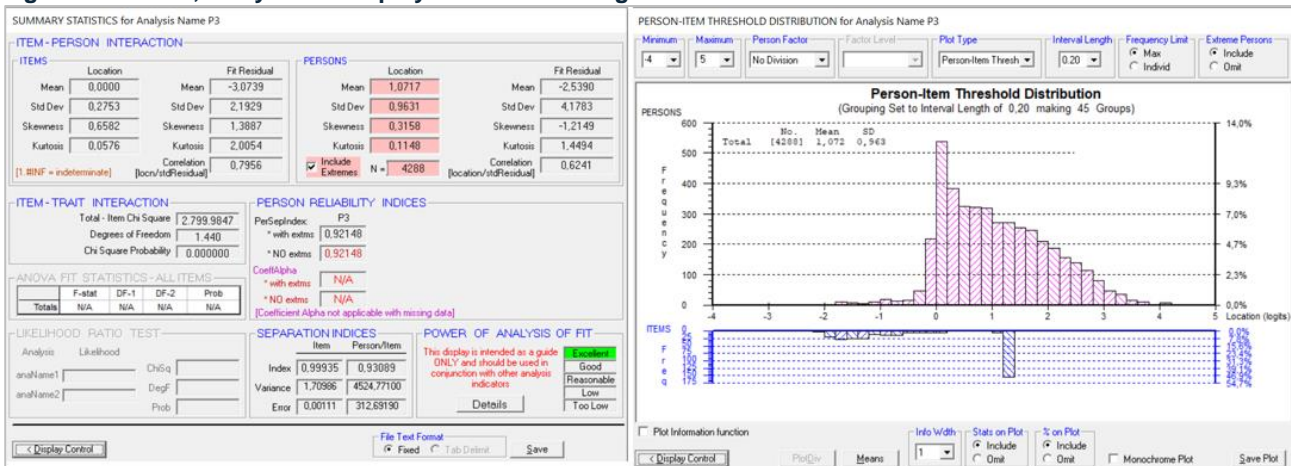
**Figur 18: Skala 3, analyse af den dikotomt scorede udgave**



Ved at introducere tiden i scoringen opnås, at kun dem, der både svarer rigtigt og hurtigt, får en høj score.

Af Figur 19 fremgår resultaterne af analysen af den polytomt scorede udgave af skala 3. Her er de store forskelle ift. den dikotome scoring (Figur 18), at der her er en meget høj reliabilitet (PSI=0,92) og at langt flere elever (de rødskraverede søjler) ligger omkring 0 på location-aksen.

**Figur 19: Skala 3, analyse af den polytomt scorede udgave**



Analysen af data med den polytome scoringsmodel (Figur 19) viser, at skala 3 giver en mere statistisk pålidelig (reliabel) scoring. Der er ingen items, der udviser markante forhøjede fit-residualer, så ingen items kasseres, og der arbejdes videre med de 160 items, vi startede med.

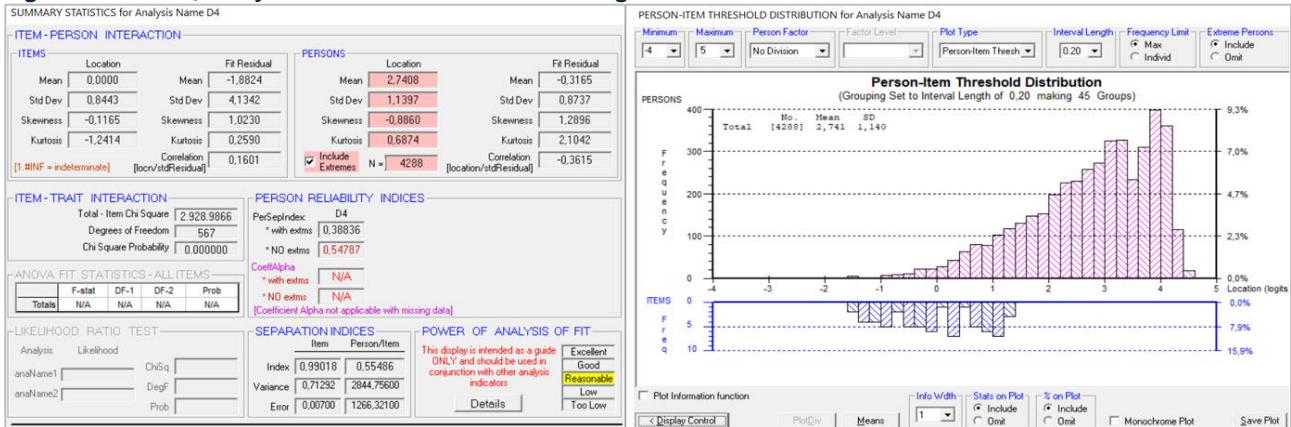


## 7.4.4 Skala 4: Sammenligning af tal og mængder

Skala 4 består af opgaver, hvor eleven skal afgøre, om et antal prikker svarer til et talsymbol (se Figur 5, nederst til venstre). Denne type opgaver tager ca. dobbelt så lang tid som de forrige, og der svares oftere forkert end ved de to foregående skalaer. Her vil det være et empirisk spørgsmål, hvilke af scoringsprincipperne (dikotom eller polytom med tid), der anvendes.

Af Figur 20 fremgår det, at der er en rimelig reliabilitet ( $PSI=0,38$ ) ved den dikotome scoring.

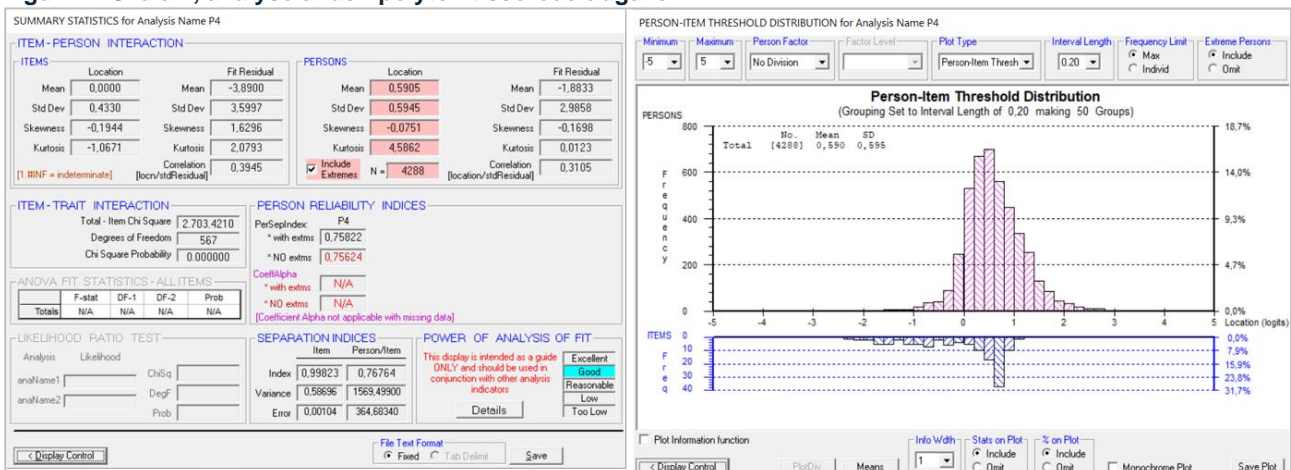
**Figur 20: Skala 4, analyse af den dikotomt scorede udgave**



Af Figur 21 fremgår resultaterne af analysen af den polytomt scorede udgave af skala 4. Analysen af data (hvor tiden er inkluderet) viser, at skala 4 giver en statistisk mere pålidelig (reliabel) scoring ( $PSI=0,76$ ).

Analysen af fit-residualer på de to scoringsmodeller viser, at der er en del flere items med kritiske værdier i den dikotome end ved den polytome, hvilket tyder på, at det eliminerer visse problemer at inkludere tiden i scoren.

**Figur 21: Skala 4, analyse af den polytomt scorede udgave**

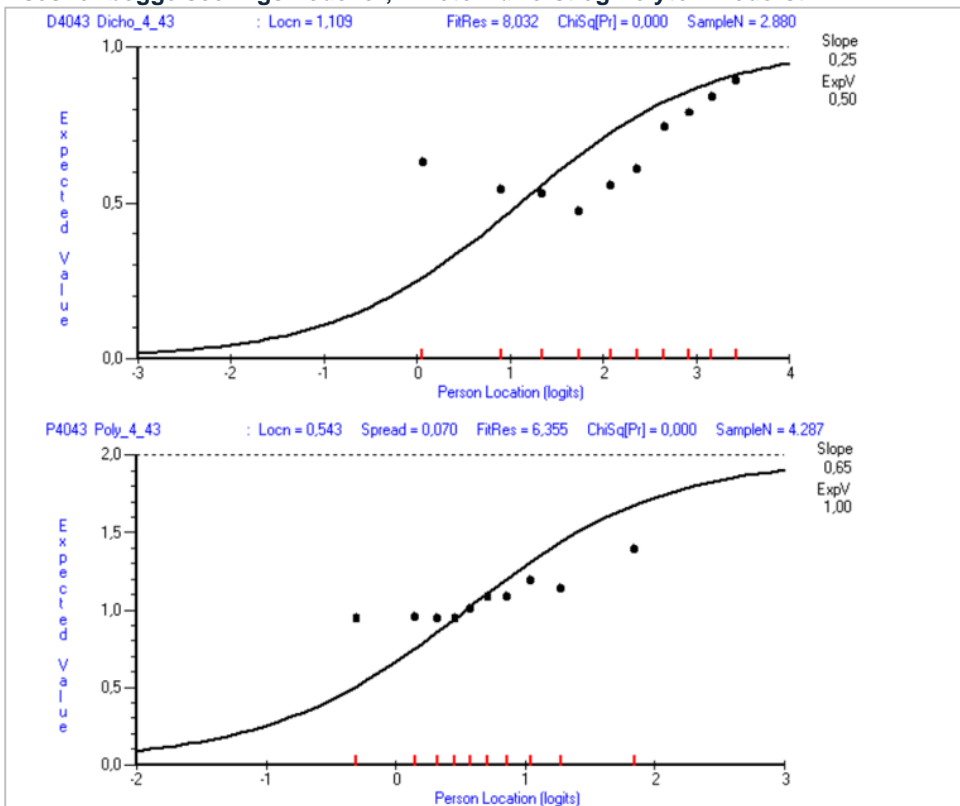


Når man sammenligner fordelingerne i de to figurer ses, at fordelingen i Figur 20 (den dikotome scoring) er højreskæv, hvilket – sammen med det faktum at der besvares ret få opgaver i gennemsnit pr. elev, jf. Tabel 10 – kunne tyde på, at nogle elever tager sig rigeligt med tid og derved kommer til at overpræstere i forhold til intentionerne med opgaveformatet.

Spørgsmålet er, om tidsforbruget fordeler sig jævnt, eller om det er nogle særlige elevgrupper (fx mht. dygtighed), der i særlig grad tager sig god tid til at besvare opgaverne rigtigt. Det kan under-

søges ved at sammenligne item-karakteristik-kurven for henholdsvis den dikotome og den polytome scoring for det enkelte item, hvilket er gjort i Figur 22 for Item 043 (et af de problematiske items i skala 4).

**Figur 22: Item-karakteristik-kurver for et dårligt fungerende item (Skala 4, Item 043). Item-karakteristik-kurven vises for begge scoringsmodeller; Dikotom øverst og Polytom nederst**



Det er tydeligt, at det især er de svageste (måske dem, der er nødt til at tælle på fingrene) og stærkeste elever (fx de meget grundige, der ikke bryder sig om at lave fejl), der klarer sig forholdsvis godt i den øverste del af Figur 22, hvor man ikke tager svartiden med i beregning af opgavens scoring.

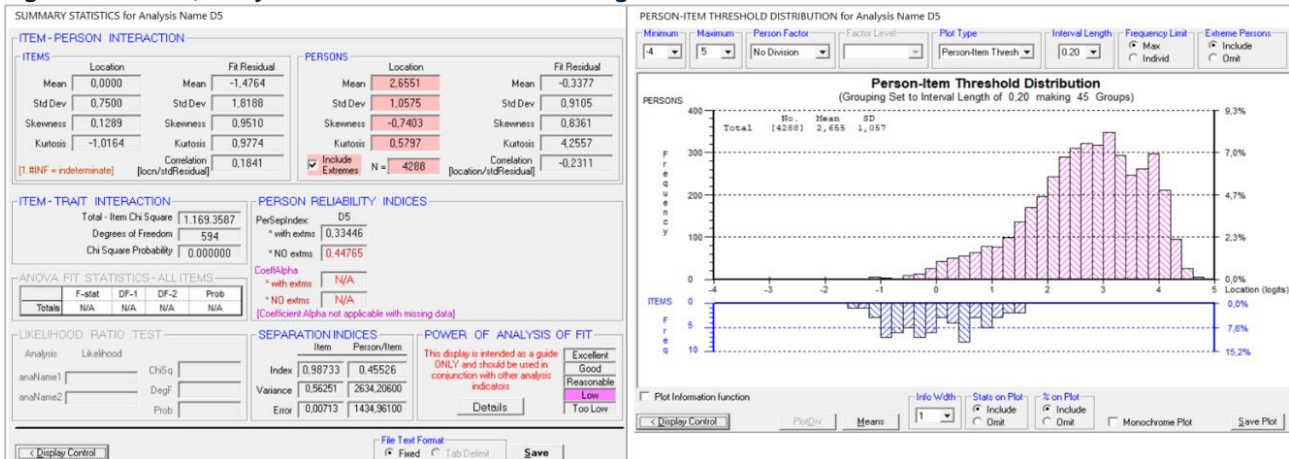
Ved at sammenholde resultaterne fra analyser af data fra de to scoringsmodeller identificeres i alt 7 opgaver (bl.a. Item 043 vist i Figur 22), der generelt passer dårligt til modellen. Disse kasseres, hvorved der resterer 56 opgaver i skala 4.

### 7.4.5 Skala 5: Addition af tal/mængder

I den sidste skala (skala 5) består opgaverne i addition af tal, antal og mængder med forskellige repræsentationer (jf. Figur 5, nederst til højre).

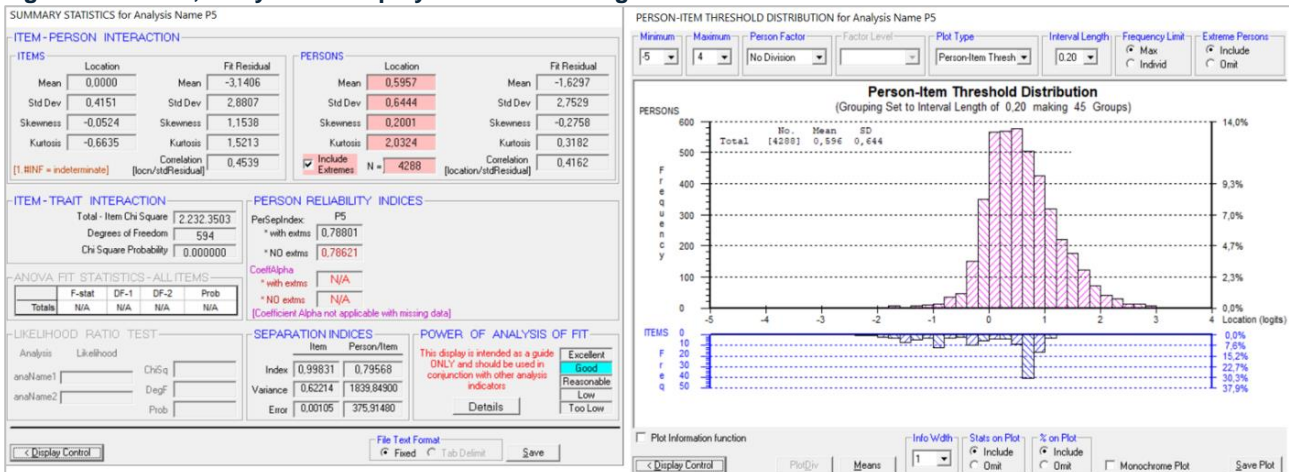
Af Figur 23 fremgår det, at der er en lav reliabilitet ( $PSI=0,33$ ) ved den dikotome scoring.

**Figur 23: Skala 5, analyse af den dikotomt scorede udgave**



Af Figur 24 fremgår resultaterne af analysen af den polytomt scorede udgave af skala 4, hvor der ses en ganske pæn reliabilitet ( $PSI=0,79$ ).

**Figur 24: Skala 5, analyse af den polytomt scorede udgave**



Ved at introducere tiden i scoringen opnår man, at dem, der både svarer rigtigt og hurtigt, får en høj score. Det er i øvrigt ved skala 5 de samme mekanismer som ved skala 4, der gør sig gældende, og det vil være den polytome scoring, der kommer til at danne det bedste afsæt i den egentlige afprøvning.

Der er en enkelt opgave, der kasseres fra skala 5, da den diskriminerer for svagt.

## 7.5 Udvalgelsen af items og scoringsmodel til runde 1 og 2

Den gennemførte analyse har haft til formål at identificere og eliminere de testopgaver, der slet ikke passer til skalaen. Der er nogle opgaver, der afviger noget fra Raschmodellens krav, men som alligevel er bevaret i opgavebanken. Fx visse opgaver i den lette ende, der viste sig at overdiskriminere lidt for meget.

Med denne lempelige tilgang er der sket en begrænset kassering af testopgaver, hvilket fremgår af Tabel 11. Her fremgår det også, hvilken scoringsmodel som vil give det bedste resultat for hver skala.

**Tabel 11: Oversigt over resultatet af kvalitetssikringen af digitale testopgaver**

	Skala 1 (D1)	Skala 2 (P2)	Skala 3 (P3)	Skala 4 (P4)	Skala 5 (P5)	I alt
Antal items fra start	164	103	160	63	66	556
Frasorterede	22	3	0	7	1	33
"Overlevende" items	142	100	160	56	65	523
Bedste egnede scoringsmodel	Dikotom	Polytom	Polytom	Polytom	Polytom	-

Som det ses, virker de polytome skalaer generelt bedst. Det giver en fordel at inkludere tiden eksplicit i scoringen af skala 2-5. For skala 1 er tidsfaktoren allerede indbygget, så det giver redundans at score denne type opgaver med tiden indbygget to steder. Analyserne pegede dog i retning af, at der skulle arbejdes med at finde en anden (lavere) tærskel for hastighed (end mediantiden) med henblik på, at items' sværhedsgrader koncentrerer omkring den dygtighed, som eleverne har (optimal target). Tidsgrænsen, der identificerede de 60 pct. hurtigste elever som udenfor risikozonen, blev vurderet som den mest retvisende.

**Tabel 12: Korrelationsmatrice: Den digitale tests fem delskalaer og i alt**

	D1	P2	P3	P4	P5	D1P25	GnsD1-P5
D1	100%	29%	32%	46%	41%	55%	65%
P2		100%	60%	50%	50%	81%	75%
P3			100%	55%	53%	77%	79%
P4				100%	66%	78%	83%
P5					100%	76%	81%
D1P25						100%	96%

I Tabel 12 er vist korrelationen mellem de fem delskalaer indbyrdes samt to bud på en fælles skala. D1P25 er en skala dannet på grundlag af alle items i de fem delskalaer. GnsD1-P5 er dannet ud fra et råt gennemsnit af de fem delskalaer. Som det fremgår, er korrelationen mellem delskalaerne parvis stærkt varierende og på højst 66 pct. (i tabellen er lavest korrelation markeret med rød og højest korrelation markeret med grøn).

Det har været en udfordring at identificere ét samlet mål for tegn på talblindhed ud fra testbesvarelsenerne, hvilket hænger sammen med, at talblindhed er et komplekst fænomen, der kan komme til udtryk på forskellige måder. Af de to afprøvede modeller (D1P25 og GnsD1-P5) fungerede skalaen baseret på gennemsnit af de fem delscores bedst. Den svage korrelation mellem delskalaerne fortæller imidlertid, at de fem delskalaer måler noget forskelligt. Det er derfor blevet vurderet at give lærerne information om alle fem delscores plus en samlet score. Med henblik på at kunne formidle resultaterne på en relativt enkel og let forståelig måde blev de fem skalaer transformeret til et fælles format (en skala med middelværdi 5 og spredning 2 – afrundet til heltal), og der blev beregnet et simpelt gennemsnit af de 5 skalaer som en samlet score. Præsentationen af resultater til lærere er illustreret i Tabel 14 i afsnit 8.2.1, hvor det også er gengivet, hvordan der er blevet vejledt om fortolkning af resultater.

## 8. Resultater og validering

Analysen af resultater er baseret på en Raschmodel samt ækvivaleringsteknikker, og den tager datamæssigt afsæt i observationsguiden, den digitale test runde 1 og 2 samt samtaletesten. Analysen har til formål at undersøge testenes måleegenskaber, herunder reliabilitet og validitet, samt at estimere grænseværdier (cutscores). Kriterievalideringen af samtaletesten sker ved sammenligning med resultater fra den hollandske RD4-test.

### 8.1 Observationsguide – tjekliste og it-baseret udgave

Observationsguiden er første trin i tragtmødelen, hvor læreren skal forholde sig til en række udsagn (i alt 34 items) om elevens eventuelle vanskeligheder, som beskrevet i afsnit 6.3. Udsagnene blev formuleret i forbindelse med udviklingen af testen.

I projektet arbejdes med to udgaver:

- Dels en simpel liste over udsagnene ([se tjeklisten<sup>26</sup>](#)) som ligger på hjemmesiden,
- dels en egentlig it-baseret udgave, hvor besvarelserne registreres elektronisk

Ved sidstnævnte får læreren feedback umiddelbart efter besvarelsen af items, jf. afsnit 6.3 om opsætning af observationsguiden. Ydermere gemmes besvarelserne, så der gives mulighed for at analysere, om svar på items i observationsguiden, som tilsammen kan beskrive en skala, viser tegn på talblindhed.

#### 8.1.1 Grundlag for analyse af responser

Scoringmodellen for observationsguiden går ud på at tildele point for problemer i forhold til de enkelte udsagn. Hvis eleven "aldrig eller sjældent" har problemer gives 0 point (grøn), hvis eleven har problemer "af og til" gives 1 point (gul), og hvis eleven "ofte eller altid" har problemer gives 2 point (rød). Hvis læreren ikke ved, om eleven har problemer, udgår det pågældende item af scoringen (det behandles som ubesvaret).

Der kræves besvarelse af halvdelen af items (dvs. mindst 17 udsagn, hvor der svares andet end "ved ikke"), før der kan beregnes en score.

Datagrundlaget for analyse af observationsguidens 34 udsagn/items, jf. Tabel 13, består af i alt 864 gyldige vurderinger af individuelle elever (dobbeltsvarelses fraserteret) foretaget af 243 forskellige lærere.

**Tabel 13: De 34 udsagn i observationsguiden/tjeklisten**

Emne	Nr. Observer om eleven...
Tal og tælling	1 . . har svært ved at tælle antal større end 10
	2 . . ikke kan se et antal op til 4 uden at tælle
	3 . . har svært ved at anvende tal til at angive overslag på antal elementer i større mængder
	4 . . har svært ved at læse to- og trecifrede tal højt
	5 . . har svært ved at ordne et- og tocifrede tal, herunder placere dem rimeligt på tallinje

<sup>26</sup> På hjemmesiden <https://epinionglobal.com/talblindhed/talblindhed-materiale/> ligger der link til tjeklisten. Linket er [https://d2kxr106kyasq0.cloudfront.net/wp-content/uploads/2020/09/280920\\_Tjekliste-tegn-pa-talblindhed-for-elever-i-4.-klasse.pdf](https://d2kxr106kyasq0.cloudfront.net/wp-content/uploads/2020/09/280920_Tjekliste-tegn-pa-talblindhed-for-elever-i-4.-klasse.pdf)

	6 . . har svært ved at huske "vigtige" tal, fx eget husnummer
	7 . . har svært ved at anvende titalssystemet, fx at 35 består af 3 tiere og 5 enere
	8 . . har svært ved ikke-kendte talfølger fx 1 - 3 - 5 - 7 eller 14 - 24 - 34 – 44
	9 . . let lader sig forvirre ved læsning og skrivning af tal
	10 . . taster forkert ind på lommeregner
	11 . . har svært ved at tælle baglæns, fx fra 35 til 27
Addition og subtraktion	12 . . har svært ved fortsat tælling
	13 . . har svært ved at fremkalde plustabellen inden for området 0 - 5 uden at tælle
	14 . . mangler regnestrategier ved addition ud over fortsat tælling
	15 . . mangler regnestrategier ved subtraktion ud over at tælle baglæns
	16 . . har svært ved at vurdere om et resultat er næsten rigtigt
	17 . . anvender konsekvent tælling med brug af fingrene ved beregninger
	18 . . ofte anvender streger eller konkrete selv ved enkle additionsstykker
	19 . . vægrer sig ved at give overslag på beregninger
Multiplikation	20 . . har svært ved at automatisere enkle gangestykker som 3·4
	21 . . vægrer sig ved at gætte på resultater
Orientering	22 . . er retningsforvirret og har svært ved at gengive højre og venstre
	23 . . har svært ved at spejle en enkel figur
Tid	24 . . har svært ved at aflæse analoge ure
	25 . . har svært ved at fornemme tidslængde
	26 . . har svært ved at omsætte digitale klokkeslæt til tidspunkter på dagen
Tal og størrelser i hverdagen	27 . . vægrer sig ved at bruge tal ved beskrivelser af hverdagssituationer, fx antal, priser, længde og vægt
	28 . . ofte gætter meget forkert på længdemål og vægt
	29 . . har svært ved at omsætte en pris til "et beløb, der siger dem noget"
	30 . . giver udtryk for, at de har nogen, som hjælper med at håndtere tal i hverdagen
Elevadfærd, oplevelser og holdning	31 . . arbejder meget langsomt med tal og regning med tal
	32 . . viser fysisk ubehag, når der skal arbejdes med tal
	33 . . oplever sig selv som dum og har lavt fagligt selvværd - eleven tror ikke på, at han/hun kan lære at regne
	34 . . har svært ved at huske procedurer

Note om farvemærkning: De ikke-farvede udsagn fungerer som forventet. De gult markerede er lokalt afhængige. De blå markerede skiller sig ud fra de øvrige som adfærdsspørgsmål (se teksten for nærmere forklaring).

Skal observationsguiden anvendes systematisk, er det vigtigt at vide, hvor lang tid det tager at benytte den pr. elev. Den gennemsnitligt registrerede tid fra at respondenter (lærer/vejleder) har åbnet observationsguiden, til den igen lukkes, er urealistisk høj, da det tilsyneladende indeholder data for registreringer, hvor der ikke er afsluttet efter registrering. Ser man bort fra relativt få ekstreme værdier (dvs. de 2 pct. af besvarelsener, der varede under 1,5 eller over 60 minutter) ligger mediansvartiden på 6 minutter pr. elev og middelsvartiden på omkring 10 minutter pr. elev. Så hvis læreren skal gennemgå alle elever i en klasse med 20-25 elever, vil det normalvis tage et sted mellem 2 og 4 timer.

### 8.1.2 Raschanalyse af data fra observationsguiden

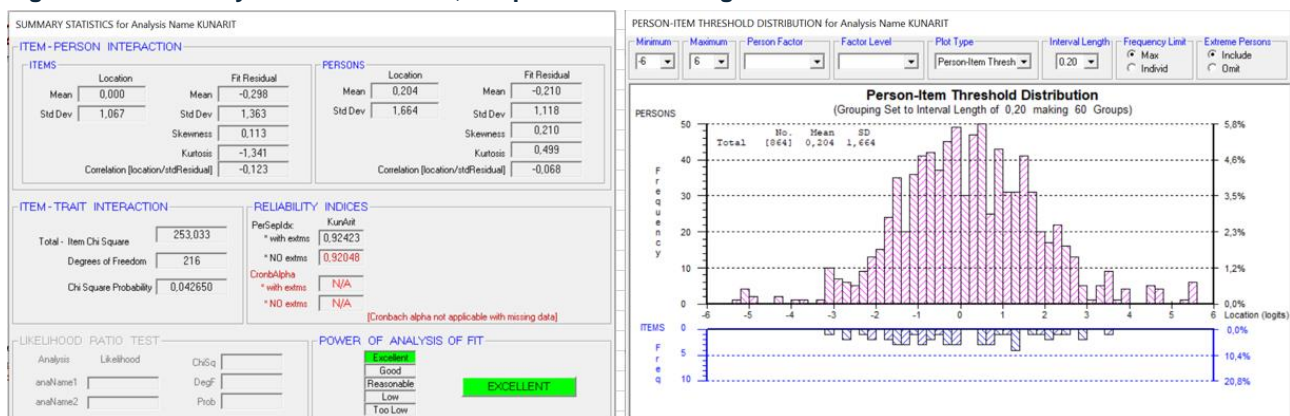
Hovedformålet med denne analyse er at undersøge, om alle udsagnene tilhører samme Raschmodel, om der er lokal afhængighed, eller om data passer bedre til en model, hvor udsagn med dårlig tilpasning udelades.

Analysen viser, at der kan være tale om to skalaer. Der er 10 udsagn, der ikke passer til de øvrige (de har et for stort fit-residual og underdiskriminerer dermed). I Tabel 13 er det item-nummer 17-19, 21-24, 30 og 32-33 (markeret med blå i Tabel 13 ovenfor). En nærmere gennemgang viser, at fælles for disse udsagn er, at de drejer sig om elevadfærd og følelser snarere end aritmetiske færdigheder, som generelt karakteriserer de resterende 24 udsagn. Det kan overvejes at lade de 10 items udgøre en særskilt adfærdsskala.

For de øvrige 24 items viser der sig yderligere lokal afhængighed mellem to sæt af items, nemlig 14 og 15 samt 25 og 26 (markeret med gul). Det betyder fx, at sandsynligheden for høj score på item 15 *"mangler regnestrategier ved addition ud over fortsat tælling"* i højere grad afhænger af svaret på item 14 *"mangler regnestrategier ved subtraktion ud over at tælle baglæns"*, end om eleven samlet set har tegn på talblindhed. Besvarelserne tyder på, at lærerne ikke skelner mellem de to items – de ses som næsten samme udsagn. Man kunne overveje at udelade det ene eller lægge dem sammen til ét spørgsmål.

Af Figur 25 fremgår de overordnede resultater af raschanalysen af de 24 items fra observationsguiden, der passer til modellen – dvs. de items, der har hovedfokus på udsagn om elevens aritmetiske færdigheder.

**Figur 25: Raschanalysen af de 24 items, der passer til en færdighedsorienteret model**



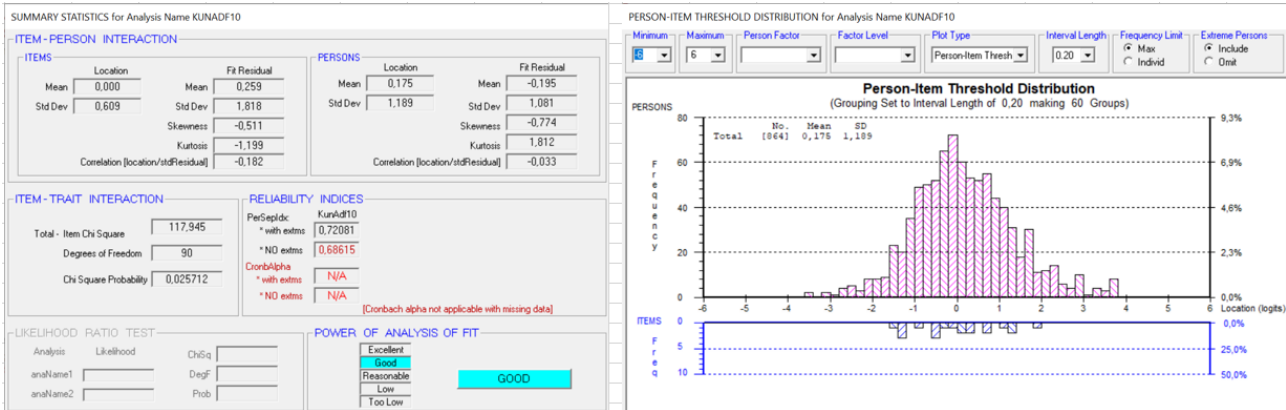
Der synes at være en fin tilpasning til modellen. Af figurens venstre side fremgår, at besvarelserne har en høj reliabilitet (PSI=0,92), og meget tyder på en valid måling (Chi<sup>2</sup> prob.=0,043) med en nærmest perfekt targetting – dvs. en meget fin overensstemmelse mellem fordelingerne af person- og item-locations (de blå og røde søjler i figurens højre side).

### 8.1.3 To dimensioner i observationsguiden

Spørgsmålet er, om de resterende 10 udsagn kan anvendes til noget. Som nævnt havde de 10 items primært fokus på elevernes følelser og adfærd i forhold til arbejdet med tal og mængder. Det skal derfor undersøges, om disse 10 items kan danne en separat adfærdorienteret skala.

Af Figur 26 fremgår de overordnede resultater fra analysen af, om de 10 items kan danne en uni-dimensionel skala.

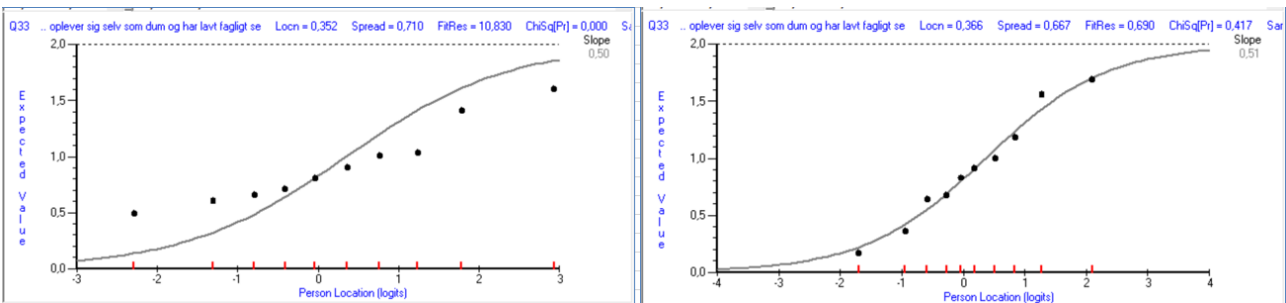
Figur 26. Raschanalyse af, om de 10 items kan passe sammen og danne en adfærdorienteret skala



Det viser sig, at ingen items falder ud af analysen, og set i lyset af at der kun er 10 items, er reliabiliteten acceptabel (PSI=0,72), og det er tæt ved, at modellen passer til de generelle krav ( $\chi^2$  prob.=0,026).

Der er altså udsagn, der ikke passer med en samlet model, hvor læreren vurderer elevernes aritmetiske færdigheder, men som godt kan hænge sammen med de samme læreres vurdering af elevernes holdninger og adfærd. For at illustrere dette kan vi se på item 33 i observationsguiden. Her skal læreren observere om eleven *”oplever sig selv som dum og har lavt fagligt selvværd – eleven tror ikke på, at han/hun kan lære at regne”*

Figur 27: Item-karakteristik-kurver for Item 33 i en samlet færdighedsskala (tv) og i en affærdsskala (th)



Af Figur 27 (venstre side) ser vi, at dette udsagn kun i meget begrænset omfang hænger sammen med lærervurderede tegn på talblindhed beskrevet i observationsguiden. Af figures højre side ses, at hvis vi kun ser på de 10 udsagn, der handler om elevens adfærd, følelser og holdninger, så kommer dette udsagn til at passe nærmest perfekt til modellen.

De 10 adfærdsitems måler – som de fortolkes af de lærere, der har brugt observationsguiden – altså noget andet end resten af observationsguiden og ikke nødvendigvis noget, der hænger tæt sammen med de tegn på talblindhed, som det er intentionen at vurdere med observationsguiden.

Hovedskalaen i observationsguiden består altså af 24 items, når de 10 adfærdsitems er frasorteret. Det kan overvejes at reducere antallet af items i observationsguiden til disse 24. En sådan tilpasning vil endvidere imødekomme tilbagemeldinger fra en del lærere, der synes, at den er (for) ordrig og detaljeret (jf. afsnit 11).

## 8.2 Den digitale test

Den digitale test er afprøvet i to runder med gennemsnitligt fire måneders mellemrum. Første runde af afprøvningen foregik i perioden uge 3-10 (2021), og anden runde strakte sig over uge 19-



26 (2021). Ved begge runder blev der efter testen stillet en række spørgsmål med henblik på at måle elevernes motivation og selvtillid i forhold til tal og regning som beskrevet i afsnit 6.4.5.

### 8.2.1 Præsentation af resultater og læsevejledning til lærerne

Analyserne har vist, at de enkelte delskalaer fungerer bedst hver for sig, hvilket er et tegn på, at de ikke måler helt de samme konstrukter. Det viser sig fx, at nogle af skalaerne parvist korrelerer forholdsvist svagt indbyrdes. Det er derfor vurderet, at den enkleste og lettest forståelige måde at skabe et samlet resultat på, vil være at transformere de fem Raschskalaer til et fælles format (en skala på mellem 0 og 10 point med middelværdi 5 og spredning 2 – afrundet til heltal). Dette giver sammenlignelige skalaværdier, og når der skal beregnes en samlet score, foregår det som et simpelt gennemsnit af de 5 skalaer.

I Tabel 14 vises et eksempel på, hvordan resultaterne er blevet præsenteret for lærere på skolerne (der er lagt en læsevejledning på [projektets hjemmeside](#)).

Tabel 14: Opgørelse af elever i klassen, der har deltaget i afprøvningen (fiktive tal)

Deltest	1		2		3		4		5		Samlet resultat af den digitale talblindhedstest		
	Er der 4 prikker?		Klik på det største tal		Klik på fir-kanten med flest prikker (på tid)		Passer antal prikker til tallet?		Hvor mange giver brikkerne tilsammen?				
Opgave													
Elev	Antal svar		Antal svar		Antal svar		Antal svar		Antal svar		Anbefaling om videre undersøgelse		
	Score (0-10)		Score (0-10)		Score (0-10)		Score (0-10)		Score (0-10)				
Elev1	70	5	67	6	101	5	31	6	25	9	294	6	Samtaletest vurderes ikke nødvendig
Elev2	60	4	73	6	41	3	21	2	31	4	226	4	Samtaletest anbefales
Elev3	50	7	93	5	84	8	30	6	65	1	322	5	Samtaletest vurderes ikke nødvendig
Elev4	88	4	99	2	142	8	55	3	60	3	444	4	Samtaletest anbefales
Elev5	64	6	70	9	93	8	22	7	38	8	287	8	Samtaletest vurderes ikke nødvendig
Reference*	64	5	85	5	94	5	35	5	35	5	311	5	–

I tilbagemeldingen til lærerne fulgte en vejledende tekst til, hvordan resultaterne skal læses. Det blev understreget, at antallet af svar er et udtryk for, hvor hurtigt eleven har besvaret opgaverne (der er afsat netto 2 minutter pr. deltest), og at det først og fremmest gælder om at svare rigtigt, men det forøger scoren, hvis det rigtige svar er givet hurtigt.

Endvidere blev det beskrevet, at de fem deltest (skalaer) er beregnet, så scoren 5,00 er gennemsnittet af resultaterne fra de deltagende elever, og at scoren går fra 0-10 point (afrundet til heltal, hvor værdier under nul sættes til 0 og værdier over 10 sættes til 10). Det samlede resultat er et gennemsnit af resultaterne for de fem deltest.

Lærerne blev informeret om, at der vurderes risiko for talblindhed, hvis elevens samlede score er under cutscoren, dvs. blandt de lavest præsterende 40 pct. af de deltagende elever, og at der i disse tilfælde anbefales en samtaletest. Når en elev er blandt de 40 pct. lavest præsterende deltagere, svarer det til, at eleven er blandt de 10 pct. lavest præsterende på landsplan. Cutscoren er for den digitale test som helhed beregnet til 4,68 (se afsnit 8.4.3 om estimation af cutscores).

Samtidig blev det kommunikeret, at eleven normalvis ikke vil være i risikogruppen for talblindhed, hvis elevens samlede score ligger over cutscoren, hvorved en samtaletest ikke vil være nødvendig. Af vejledningen fremgik det dog også, at resultaterne ikke nødvendigvis giver et sikkert billede af, om eleven er talblind, og at læreren derfor altid bør holde resultaterne op mod sine andre erfaringer med og evalueringer af eleven.

I forhold til vurderingen af om et testresultat kan betragtes som retvisende, blev det påpeget, at man skal være særligt opmærksom på, hvis eleven har svaret meget langsomt, meget hurtigt eller ikke har fuldført testen. Det kan være et tegn på misvisende testresultater, hvis eleven har svaret meget langsomt (evt. hvor eleven har været nødt til at tælle på fingrene), eller hvis eleven har svaret meget hurtigt (hvor eleven har klikket sig igennem opgaverne uden af forsøge at svare rigtigt).

Afslutningsvist blev det angivet, at der som tommelfingerregel bør gennemføres en samtaletest, hvis læreren er i tvivl.

## 8.2.2 Resultater fra de to afprøvningsrunder

Det samlede antal elever og testforløb i de to testrunder fremgår af Tabel 15.

**Tabel 15: Datagrundlag for hhv. digital test og samtaletest**

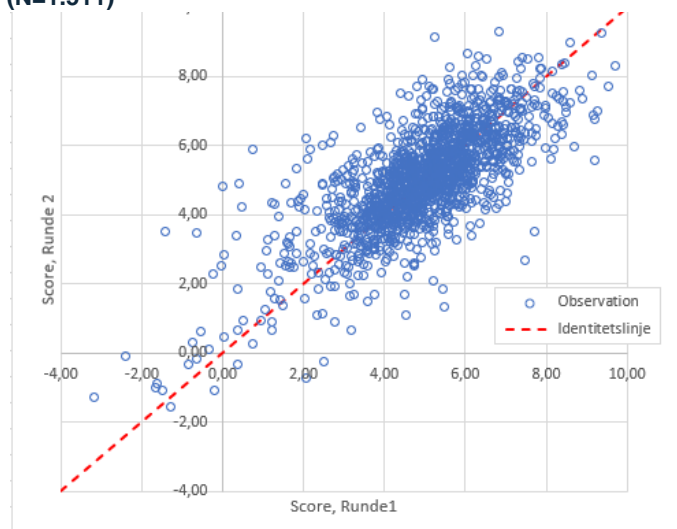
	Digital test	Samtaletest	Begge testtyper
Elever i Runde 1	3.292	1.599	-
Elever i Runde 2	1.730	326	-
Elever der deltog i begge runder	1.511	108	-
Antal deltagende elever (renset for dubletter)	3.511	1.817	1659

Af afsnit 8.3.3 fremgår det, hvor stærk test-retest-sammenhængen er (opgøres ved at se på korrelationen mellem de to testresultater) i forhold til samtaletesten. I Tabel 16 og Figur 28 illustreres tilsvarende for den digitale test. Der arbejdes her med ikke-afrundede scorer.

**Tabel 16: T-test, digital test  
(Parvis dobbelt stikprøve for middelværdi)**

	ScoreR1	ScoreR2
Middelværdi	4,96	5,04
Varians	2,51	2,18
Observationer	1511	1511
Pearson-korrelation	0,737	
Fg	1510	
t-stat	-3,104	
P(T<=t) en-halet	0,00097	
t-kritisk en-halet	1,64586	
P(T<=t) to-halet	0,00194	
t-kritisk to-halet	1,96154	

**Figur 28: Digitale test-scorer, runde 1 og 2 for samme elever (N=1.511)**

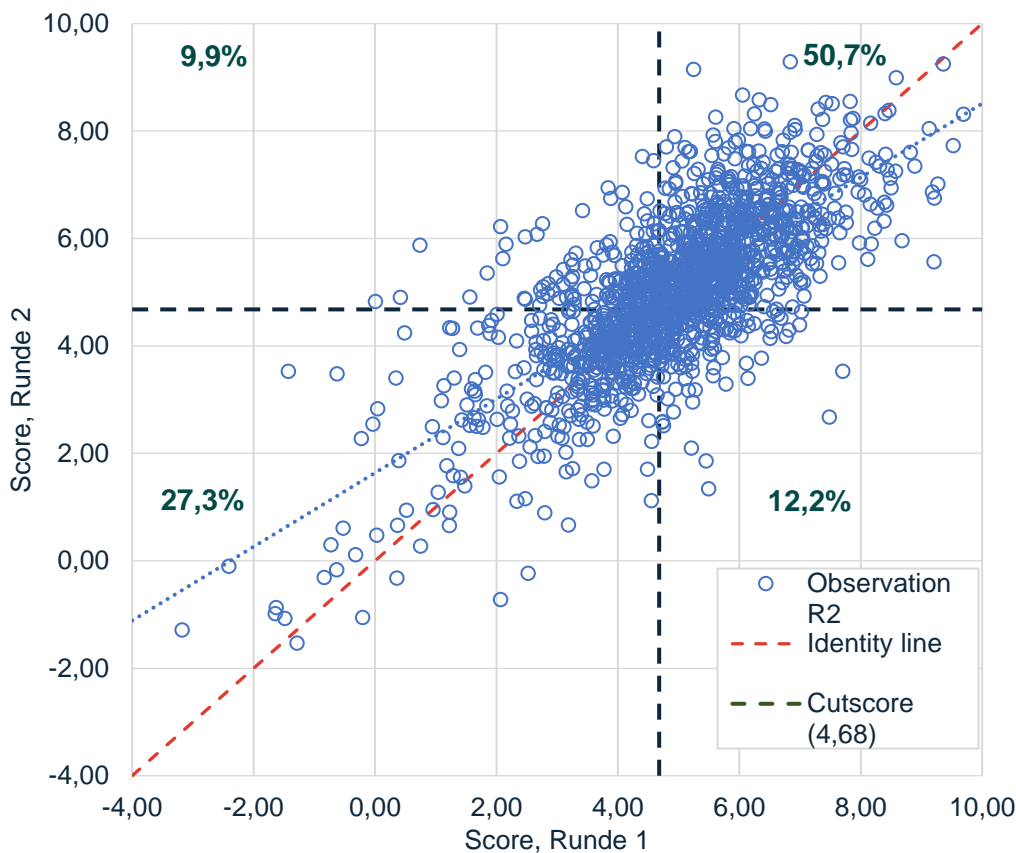


Som det fremgår, er der i den digitale test en tydelig, men ikke overvældende, korrelation mellem præstationerne i første og anden testrunde ( $r=0,73$ ). Endvidere er forskellen mellem resultaterne i første og anden testrunde klart signifikant ( $\alpha<0,01$ ). Forskellen i den gennemsnitlige score mellem

første og anden runde er imidlertid beskeden (0,08 point på skalaen med middelværdi 5 og SD=2), så det kan således ikke ses med det blotte øje, at punktsværmens i Figur 28 ligger signifikant over identitetslinjen. Test-retest-korrelationen for den samlede skala, der er vist ovenfor, er højere end test-retest-korrelationen for de enkelte skalaer ( $r=0,54$  til  $0,65$ ).

I Figur 29 er vist samme data som i Figur 28 med tilføjelse af cutscorer for de to test. Metoden for estimation af cutscores er uddybet i afsnit 8.4.3. De to cutscorer inddeler figuren i fire felter, der illustrerer andelen af elever, der har scoret det samme (over eller under cutscoren), højere eller lavere i første og anden testrunde. Således har 50,7 pct. af eleverne scoret over cutscoren i både runde 1 og 2, og omvendt har 27,3 pct. scoret under cutscoren i begge testrunder. Til gengæld har over en femtedel af eleverne (22,1 pct.) scoret på den anden side af cutscoren i anden testrunde end i den første. Hvor 9,9 pct. har scoret over cutscoren i runde 2, har 12,2 pct. scoret lavere.

**Figur 29: Digitale test-scoring, runde 1 og 2 for samme elever med markering af cutscoren (N=1.511)**



Hvis eleverne har svaret præcist i overensstemmelse med deres dygtighed ved begge testrunder, vil en forskel i udfald af de to test betyde, at deres dygtighed har ændret sig – enten i opad- eller nedadgående retning. Dog kan en række andre faktorer end dygtighed påvirke testforløbet og dermed resultatet. Der kan fx være lærerens instruktion (har læreren gennemført en demotest med eleverne? Har læreren motiveret eleverne og forklaret dem, hvad testresultaterne skal bruges til?) eller tekniske forhold (en mus, et tastatur eller skolens internetforbindelse). Den væsentligste årsag til forskelle mellem de to testresultater skal dog formentlig findes i elevernes aktuelle dagsform (koncentration og engagement) og personlige forhold. Givet testopgavernes format, hvor svarhastighed spiller en væsentlig rolle, betyder koncentration meget for testens udfald. Hvis eleverne har andre ting, der optager dem, har sovet dårligt om natten eller er kommet afsted uden morgenmad,

kan det medvirke til et dårligere testresultat på dagen. Ikke desto mindre har langt størstedelen af eleverne (knap 80 pct.) svaret konsistent i forhold til cutscoren.

På trods af følsomhed over for udsving i elevernes indsats, har den digitale test vist sig som et stærkt måleredskab ved at levere en høj målingsreliabilitet og validitet. Den digitale test kan dermed udgøre et værdifuldt redskab, der skaber et hurtigt og enkelt overblik. Dette har flere lærere givet udtryk for. Eksempelvis siger en lærer i forbindelse med spørgeskemaundersøgelsen (jf. afsnit 10):



*"Den digitale test vil være et fantastisk screeningsværktøj, fordi den er let og hurtig at gennemføre med mange" – (Lærersurvey)*

Men det er samtidigt vigtigt at være opmærksom på, at testresultater er data, som skal fortolkes.

En høj score vil som udgangspunkt være udtryk for, at eleven har velfungerende aritmetiske færdigheder og dermed ikke har tegn på talblindhed (en elev svarer ikke rigtigt på 90 pct. af 300 opgaver på ti minutter ved et tilfælde). En lav score er tegn på talblindhed, men en lav score kan også have en række andre årsager – fx manglende engagement og/eller fokus på at svare hurtigst muligt, utilstrækkelig instruktion, tekniske problemer med udstyret eller internetforbindelsen eller simpelthen en dårlig dag for eleven (ikke sovet godt, ikke spist morgenmad, bekymringer som forstyrrer). Når elever har en lav score i en digital test, er der derfor behov for videre undersøgelser af elevens tegn på talblindhed. Man kan lade eleven gentage den digitale test, hvis der er grund til at tro, at eleven ikke præsterede på et retvisende niveau første gang. Men det vil som regel give mere viden at gennemføre en samtaletest.

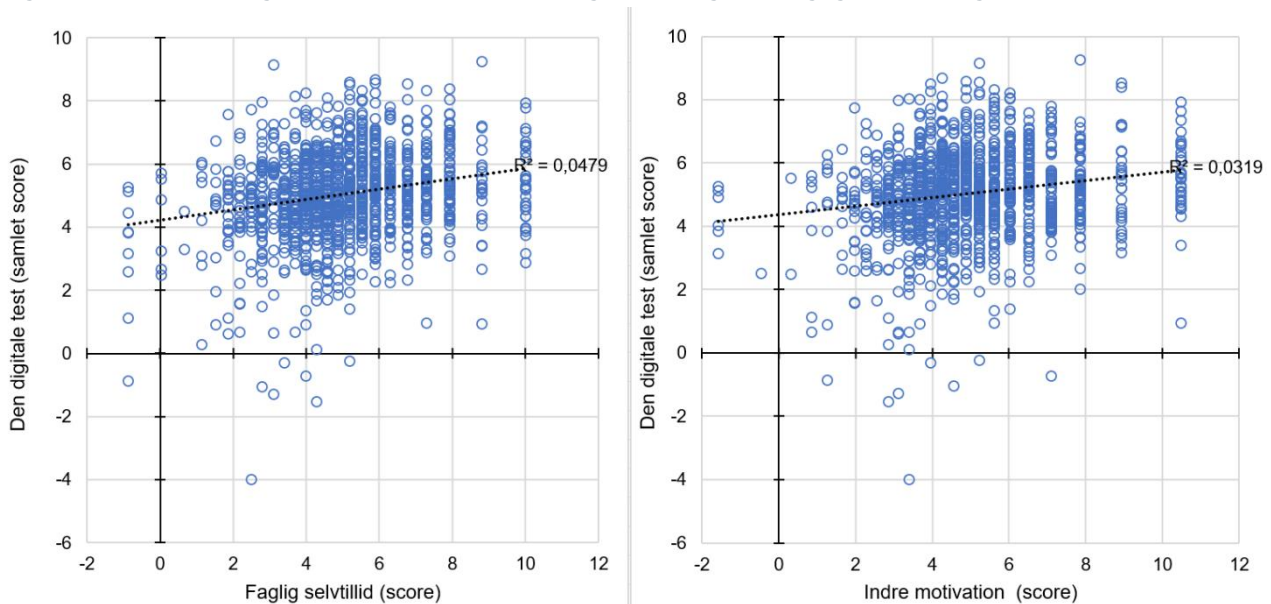
### 8.2.3 Motivation og selvtillid

I forlængelse af den digitale test (de fem delskalaer) blev eleverne stillet en række spørgsmål med det formål at afdække deres indre motivation og faglige selvtillid i forhold til tal og regning (jf. afsnit 6.4.5).

Det er veldokumenteret, at elever, som har indre motivation og faglig selvtillid, i højere grad er parate til at yde en indsats, når de møder faglige udfordringer<sup>27</sup>. Det er ofte antaget, at elever, der scorer højt i faglige test og prøver, automatisk også er dem, som er mest motiverede og har størst selvtillid. Sådan hænger det imidlertid ikke altid sammen. Tværtimod ses der ofte kun en svag sammenhæng mellem motivation, selvtillid og faglige resultater. Dette viser sig også at være tilfældet for eleverne, der har besvaret den digitale test inklusive motivations- og selvtillidsspørgsmålene, som det fremgår af Figur 30. Her ses ganske vist en klart signifikant ( $\alpha < 0,01$ ), men kun svagt positiv, sammenhæng mellem den digitale test og hhv. selvtillid ( $R^2 = 0,048$ ) og motivation ( $R^2 = 0,032$ ). Ligesom de fem delskalaer er motivation og selvtillid scoret ved Raschanalyse og transformeret til skalaer med middelværdien 5 og en spredning på 2.

<sup>27</sup> Makransky, G., Wandall, J., Madsen, S. R., Hood, M., & Creed, P. (2020). Development and Validation of the UiL-Scales for Measurement of Development in Life Skills—A Test Battery of Non-Cognitive Skills for Danish School Children. *Scandinavian Journal of Educational Research*, 64(4), 612-627.

Figur 30: Sammenhængen mellem samlet score i digital test og hhv. faglig selvtillid og indre motivation



Undersøgelsen viser altså, at de elever, der scorer lavt i den digitale test, statistisk set har en lavere motivation og selvtillid end dem, der scorer højt – omend forskellen ikke er ret markant. Sammenlignet med andre elever vil elever med talblindhed have et større behov for motivation og selvtillid, hvis de skal yde den ekstra indsats, som er nødvendig for at klare sig i skolen og i hverdagslivet. Forskning i non-kognitive færdigheder viser, at det er væsentligt nemmere at påvirke elevers motivation og selvtillid, end det er at påvirke deres kognitive færdigheder (såsom matematik- og regnefærdigheder) direkte<sup>28</sup>. En indsats målrettet elever med tegn på talblindhed kan derfor med fordel have motivation og selvtillid som løftestang til at styrke elevernes relation til matematikken og lyst til at arbejde med det.

Et væsentligt skridt på vejen til faglig udvikling hos udfordrede elever kan altså bestå i at arbejde bevidst og systematisk med at styrke deres interesse for faget (indre motivation) og deres tro på, at indsatsen gør en forskel (faglig selvtillid). Dette kan fx gøres gennem anerkendelse, måling og synliggørelse af fremskridt kombineret med systematisk konstruktiv feedback.

## 8.3 Samtaletesten

Samtaletestens resultater bliver analyseret særskilt, og de bliver yderligere sammenholdt med den hollandske RD4-test med henblik på at undersøge kriterievaliditeten. RD4-testen er en hollandsk samtaletest, der i forvejen er valideret i forhold til en generel standard (CITO's matematiktest).

### 8.3.1 Analyse af data

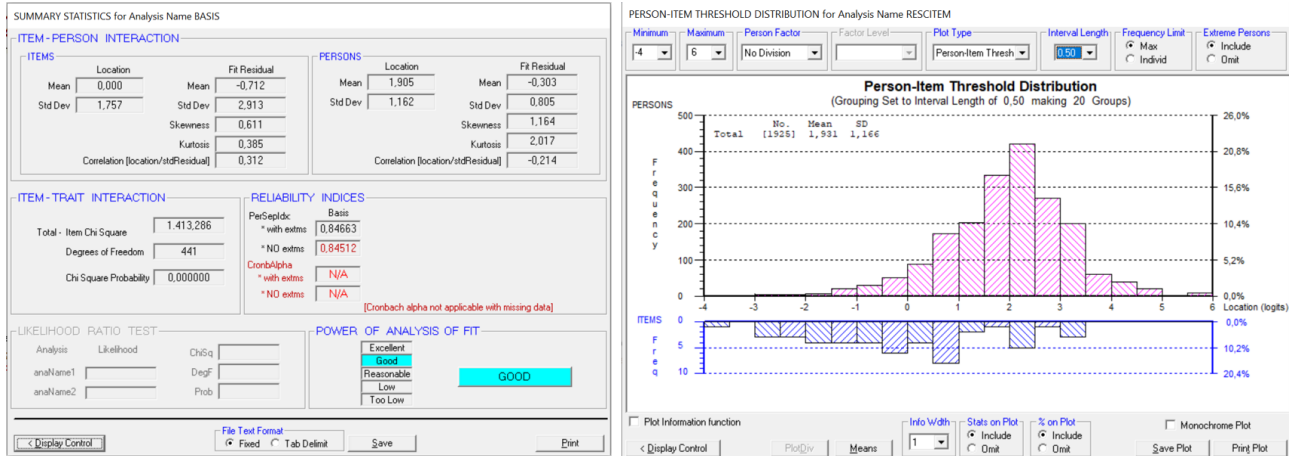
I de to afprøvningsrunder kom der i alt 1.925 gyldige besvarelser af samtaletesten. Ved den første runde var der 1.598 besvarelser frem til d. 25. april 2021. Frem til sommerferien var der yderligere 327 afprøvningsrunder af samtaletesten. Af disse var de 108 gengangere fra runde 1 (jf. afsnit 8.3.3).

<sup>28</sup> Heckman, J. J. (2008). Schools, skills, and synapses. *Economic inquiry*, 46(3), 289-324.

### 8.3.2 Raschanalyser af data fra samtaletesten

På grundlag af besvarelserne af samtaletesten gennemføres en Raschanalyse, og resultaterne herfra, som fremgår af Figur 31, ser overordnet fornuftigt ud med en reliabilitet på 0,85 (PSI), men med en noget højreskæv fordeling.

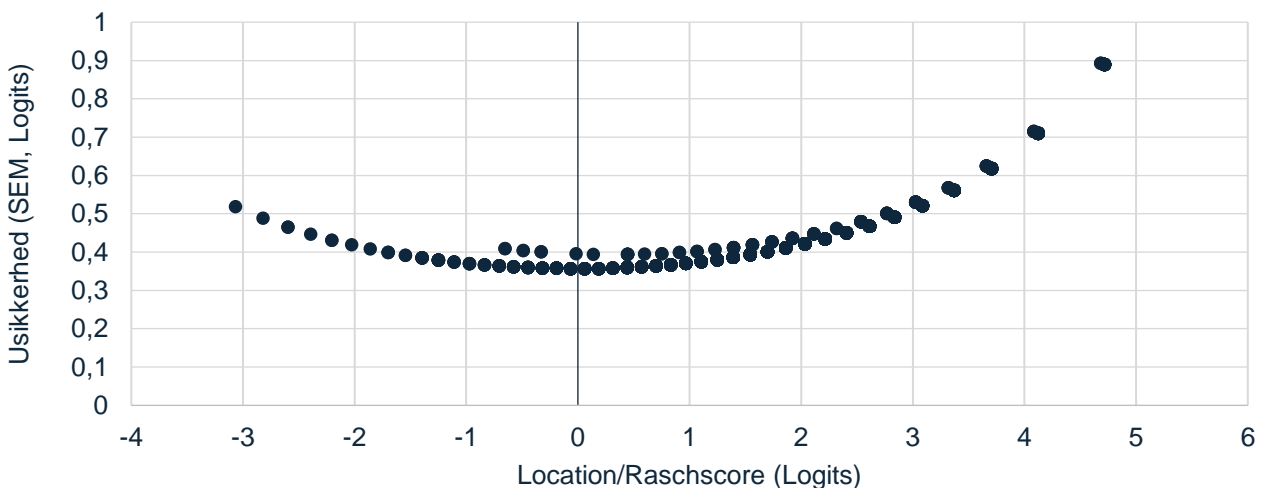
Figur 31: Samtaletesten, analyse af alle 1.925 besvarelser



Opgaverne er altså gennemsnitligt meget lette for den population, der gennemførte afprøvningen. Dette udgør ikke nødvendigvis et problem, da det især er eleverne i den lave ende af skalaen, som det i denne sammenhæng er mest relevant at få en god måling på.

I Figur 32 er vist sammenhængen mellem den statistiske usikkerhed for elever med forskellige scorere i samtaletesten.

Figur 32: Statistisk usikkerhed ved måling med samtaletesten (N=1.925)



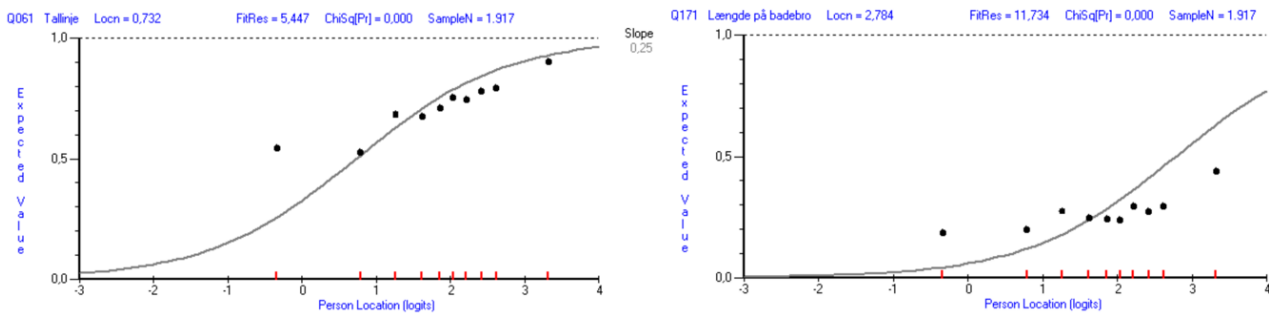
Det ses, at testen, der hvor den måler bedst, leverer målinger med en usikkerhed på SEM=0,36. Ved at sammenligne højre side af Figur 31 med

Figur 33 kan man komme frem til, at testen faktisk måler mest præcist der, hvor man vil forvente, at de fleste elever med talblindhed befinder sig.

Der er yderligere foretaget en nærmere analyse af de enkelte items i samtaletesten. Denne viser, at enkelte opgaver passer mindre godt til det samlede opgavesæt som helhed. Især skal fremhæves opgave 6 (tallinje, som skal aflæses) og opgave 17 (billede af en badebro, hvis længde skal anslås). Item-karakteristik-kurverne (ICC) for disse to opgaver er vist i

Figur 33.

Figur 33: ICC for opgave 6 og 17 i samtaletesten



Som det fremgår, er fit-residualerne for disse opgaver for store. Det vil sige, at disse opgaver underdiskriminerer i betydeligt omfang. Af den årsag har det været overvejet at lade dem udgå af beregningerne af elevernes scorer. Fjernes disse items fra modellen vil det imidlertid betyde en lille reduktion i reliabilitet, og da samtaletesten yderligere er sammensat af komponenter, der skal passe sammen rent fagligt, er det besluttet ikke at kassere disse to items i forbindelse med denne afprøvning. Det vil dog være hensigtsmæssigt at tilpasse eller erstatte dem, hvis samtaletesten skal bruges igen på et senere tidspunkt – fx i en fase 2.

### 8.3.3 Gengangere i samtaletesten

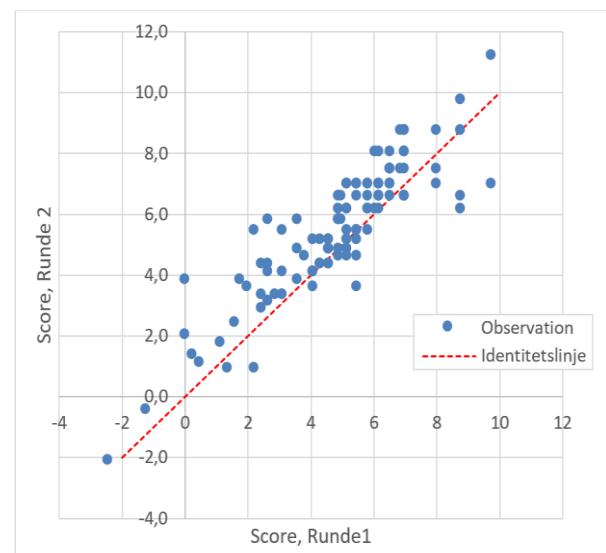
Som nævnt var der 108 elever, der havde besvaret samtaletesten i både runde 1 og 2. En sammenligning af deres resultater viser, at disse 108 elever præsterede signifikant bedre i runde 2 end runde 1, hvilket er illustreret i Tabel 17 og Figur 34.

Tabel 17: T-test (samtaletest)

(Parvis dobbelt stikprøve for middelværdi)

	Runde 1	Runde 2
Middelværdi	36,27	38,80
Varians	64,29	51,09
Observationer	108	108
Pearson-korrelation	0,864	
Fg	107	
t-stat	-6,5084	
P(T<=t) en-halet	1,2585E-09	
t-kritisk en-halet	1,659	
P(T<=t) to-halet	2,517E-09	
t-kritisk to-halet	1,982	

Figur 34: Sammenhæng mellem scorer i de 2 runder



Dette kan skyldes to forskellige forhold eller en kombination heraf. Enten 1) at eleverne er blevet dygtigere til den type matematik, der afprøves i samtaletesten, eller 2) at der er et test-retest-problem. Dvs. at eleven – pga. dialogformen som den ofte praktiseres i samtaletesten – lærer opgaverne og de korrekte svar at kende og derfor præsterer bedre anden gang end første. Men uden at være blevet bedre til faget generelt.

Det ses af Figur 34, at der er en meget solid korrelation mellem præstationerne i første og anden testrunde ( $r=0,86$ ) og endvidere, at forskellen mellem resultatet i første og anden testrunde er

stærkt signifikant ( $t=6,5$ ). Eleverne får gennemsnitligt 2,5 flere opgaver rigtigt i runde 2 (juni) end i runde 1 (april). Af Figur 34 fremgår det også, at alle eleverne har forbedret deres resultater næsten lige meget mellem runde 1 og 2 (langt de fleste observationer ligger lige over identitetslinjen).

Læring foregår sjældent så ensartet fra elev til elev, så det er sandsynligt, at der her er et test-re-test-problem: At eleverne kun klarer sig så meget bedre anden gang, fordi de er blevet trænet i de konkrete opgaver og nok også har fået det rigtige svar at vide i løbet af første runde. Det er således tvivlsomt, om samtaletesten kan bruges flere gange på de samme elever og med de samme opgaver med et validt resultat til følge.

### 8.3.4 Kriterievalidering med RD4

For at sikre at samtaletestens måde at måle på lever op til internationale standarder, er det valgt at foretage en kriterievalidering af samtaletesten op mod den hollandske RD4-test. RD4-testen er forudgående blevet valideret med den hollandske matematiktest fra CITO, som udgør en generelt accepteret standard inden for området. Det skal understreges, at RD4-testen udelukkende bruges til validering af samtaletesten og ikke skal bruges fremadrettet som en del af testbatteriet.

### 8.3.5 RD4 – en hollandsk samtaletest

RD4 er en matematikfaglig samtaletest, som er opbygget efter det hollandske curriculum. Testen er opdelt i 23 faglige dimensioner indenfor matematik og opgjort på halvårsmål, hhv. 'mid' og 'end' resultat for hvert klassetrin, som vist i Figur 35, hvor M3 betyder halvvejs igennem 3. klassetrin og E3 er ved udgangen af 3. klassetrin.

Figur 35: Elevprofil i RD4-testen

Domain	Subject	EK1	MK2	EK2	M1	E1	M2	E2	M3	E3	M4
C	1 Counting in order	T1-EK1	T1-MK2	T1-EK2							
	2 Subitize	T2-EK1	T2-MK2	T2-EK2							
	3 Determine an amount	T3-EK1	T3-MK2	T3-EK2	T3-M1	T3-E1					
	4 Counting forwards	T4-EK1	T4-MK2	T4-EK2	T4-M1	T4-E1	T4-M2	T4-E2	T4-M3	T4-E3	T4-M4
	5 Counting backwards	T5-EK1	T5-MK2	T5-EK2	T5-M1	T5-E1	T5-M2	T5-E2	T5-M3	T5-E3	T5-M4
	6 Skip counting				T6-M1	T6-E1	T6-M2	T6-E2	T6-M3	T6-E3	T6-M4
	7 Positioning				T7-M1	T7-E1	T7-M2	T7-E2	T7-M3	T7-E3	T7-M4
N	8 Many/ few/ order	G8-EK1	G8-MK2	G8-EK2	G8-M1	G8-E1	G8-M2	G8-E2	G8-M3	G8-E3	G8-M4
	9 Composition	G9-EK1	G9-MK2	G9-EK2	G9-M1	G9-E1	G9-M2	G9-E2	G9-M3	G9-E3	G9-M4
	10 Partitioning	G10-EK1	G10-MK2	G10-EK2	G10-M1	G10-E1	G10-M2	G10-E2	G10-M3	G10-E3	G10-M4
AS	11 Writing numbers	G11-EK1	G11-MK2	G11-EK2	G11-M1	G11-E1	G11-M2	G11-E2	G11-M3	G11-E3	G11-M4
	12 Naming numbers	G12-EK1	G12-MK2	G12-EK2	G12-M1	G12-E1	G12-M2	G12-E2	G12-M3	G12-E3	G12-M4
	13 Addition		OA13-MK2	OA13-EK2	OA13-M1	OA13-E1	OA13-M2	OA13-E2	OA13-M3	OA13-E3	OA13-M4
MD	14 Subtraction		OA14-MK2	OA14-EK2	OA14-M1	OA14-E1	OA14-M2	OA14-E2	OA14-M3	OA14-E3	OA14-M4
	15 Adding/ Estimating		OA15-MK2	OA15-EK2	OA15-M1	OA15-E1	OA15-M2	OA15-E2	OA15-M3	OA15-E3	OA15-M4
	16 Subtraction/ mental math		OA16-MK2	OA16-EK2	OA16-M1	OA16-E1	OA16-M2	OA16-E2	OA16-M3	OA16-E3	OA16-M4
M	17 Multiplication		VD17-MK2	VD17-EK2	VD17-M1	VD17-E1	VD17-M2	VD17-E2	VD17-M3	VD17-E3	VD17-M4
	18 Division		VD18-MK2	VD18-EK2	VD18-M1	VD18-E1	VD18-M2	VD18-E2	VD18-M3	VD18-E3	VD18-M4
Ra	19 Measuring length						M19-M2	M19-E2	M19-M3	M19-E3	M19-M4
	20 Measuring weight						M20-M2	M20-E2	M20-M3	M20-E3	M20-M4
	21 Measuring volume						M21-M2	M21-E2	M21-M3	M21-E3	M21-M4
Re	22 Ratio						VH22-M2	VH22-E2	VH22-M3	VH22-E3	VH22-M4
	23 Relations							VB23-E2	VB23-M3	VB23-E3	VB23-M4

På mange områder ligner RD4 den danske samtaletest, idet de begge 1) gennemføres 1:1 med en testperson og en elev, 2) er baseret på faglige, matematiske opgaver og 3) er baseret på en testpersons vurdering af elevens besvarelse på baggrund af velbeskrevne vurderingskriterier.

Dog er der punkter, hvorpå de to tests ikke ligner hinanden, eksempelvis er RD4 adaptiv i modsætning til den danske samtaletest. I RD4 lokaliseres elevens niveau på to dimensioner: 1) halvårsniveau og 2) niveauet inden for det hollandske halvårs curriculum. Den danske samtaletest kan i højere grad betragtes som éndimensionel og lineær og dækker en mindre del af matematikken (fokus på talforståelse, aritmetiske færdigheder og anvendelse i hverdagssammenhænge). Denne forskel har imidlertid også betydning for, hvor lang tid det tager at teste eleverne. Hvor RD4 kan forventes

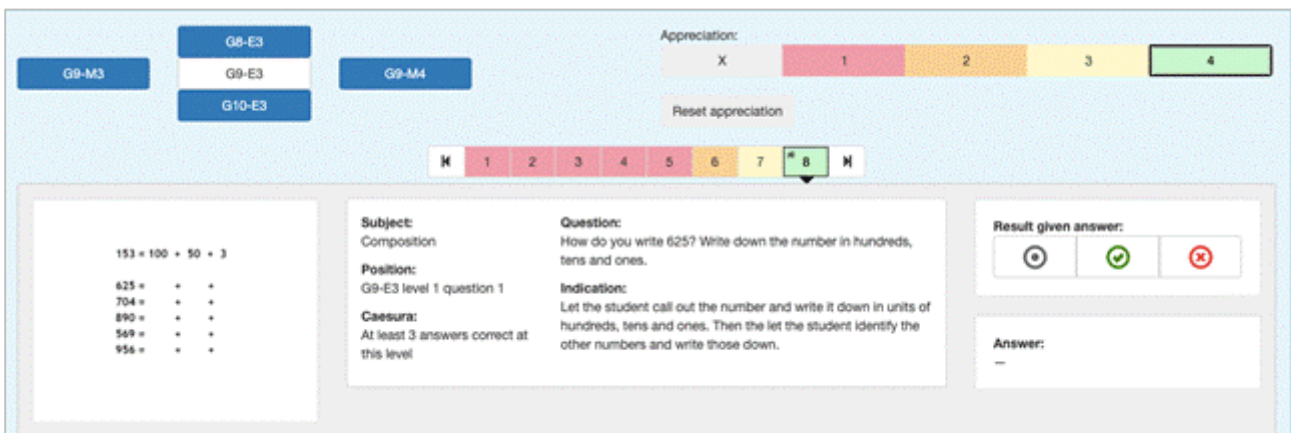


at tage omkring 1½ time (45 min. pr gennemløb), er den danske samtaletest estimeret til at tage lidt under ½ time. Hvor adaptiviteten hos RD4 i samspil med de 23 afgrænsede dimensioner gør testen velegnet til afdækning af elevens konkrete udfordringer over hele curriculum, er det en meget ressourcekrævende øvelse at lave en fuld elevprofil. Den digitale test og samtaletesten er derimod mindre tidskrævende og kan samlet set gennemføres på 45 minutter, men giver ikke samme indblik i elevens konkrete udfordringer med samme nuanceringsgrad som RD4.

### 8.3.6 RD4-testens opbygning

Bag hvert felt i Figur 35 ligger en række opgaver, der er kategoriseret efter forskellige niveauer (hhv. rød, orange, gul, grøn) og dertilhørende håndfaste vurderingskriterier som vist i Figur 36. Elevens præstation vurderes ved, at testpersonen markerer den farve, der er passende for elevens besvarelser med afsæt i vurderingskriterierne. Hvis eleven ikke kan præstere på nogle af de forskellige niveauer inden for et givent felt, markeres dette som gråt ('eleven kan ikke svare').

Figur 36: Eksempler på opgaver fra RD4



The screenshot shows a digital test interface. At the top, there are buttons for different levels: G9-M3, G8-E3, G9-E3, G9-M4, and G10-E3. To the right, there is an 'Appreciation' bar with four colored segments: 1 (red), 2 (orange), 3 (yellow), and 4 (green). Below this is a 'Reset appreciation' button. In the center, there is a row of buttons labeled 'M', 1, 2, 3, 4, 5, 6, 7, 8, and 'M'. Below this is a task card for 'Composition'. The task card includes the following information:

- Subject:** Composition
- Position:** G9-E3 level 1 question 1
- Caesura:** At least 3 answers correct at this level
- Question:** How do you write 625? Write down the number in hundreds, tens and ones.
- Indication:** Let the student call out the number and write it down in units of hundreds, tens and ones. Then the let the student identify the other numbers and write those down.

On the right side of the task card, there is a 'Result given answer' section with three buttons: a grey circle with a white dot, a green circle with a white checkmark, and a red circle with a white 'X'. Below this is an 'Answer' field with a dashed line.

Testen udføres 1:1 – dvs. med en enkelt elev, der skal løse faglige opgaver, og en testperson, der faciliterer testen ved at stille eleven opgaverne, vurdere elevens præstation på baggrund af vurderingskriterierne og observere elevens måde at løse opgaverne på.

Som udgangspunkt afvikles testen i to gennemløb: Først gennemgås tabellen lodret på det niveau, eleven forud for testen er vurderet til at ligge på (i eksemplet ovenfor er en 4. klasseelev vurderet til at være et halvt år bagud og således ligge på niveau E3), og dernæst testes der vandret i tabellen, alt efter hvordan eleven er vurderet inden for den faglige dimension. Hvis eleven ikke lever op til det niveau, der er forventet på halvårsniveauet (farven i højre side af hvert felt i tabellen), testes eleven på et lavere halvårsniveau, og omvendt testes der på et højere halvårsniveau, hvis eleven præsterer op til et forventet niveau. Således er testens anden gennemgang adaptiv, idet den tilpasses elevens besvarelser og tilstræber at lokalisere niveauet, hvorpå eleven akkurat kan besvare opgaverne korrekt. Slutmålet er en profil som den, der fremgår af Figur 35.

RD4-testen blev gennemgået med udvalgte elever fra kerneskolerne (N=71), hvorefter elevernes profiler blev fremsendt til Mieke van Groenestijn med henblik på at få analyseret resultaterne. Dette vil blive yderligere udfoldet i afsnit 9.1.2.

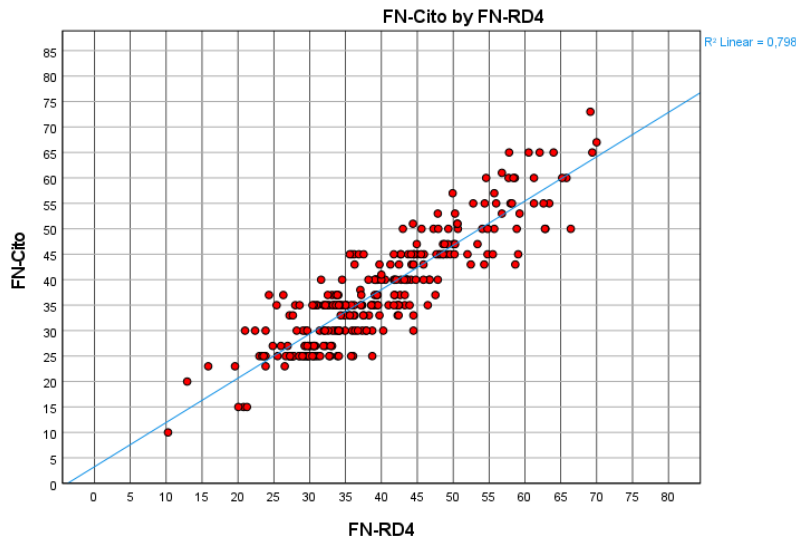
### 8.3.7 Valideringen af RD4-testen med CITO's matematiktest

Som tidligere nævnt er RD4-testen valideret gennem sammenligning med CITO's matematiktest, der udgør en generelt accepteret standard. Begge hollandske test dækker over hele den holland-

ske grundskole fra indskoling til de afsluttende klasser. Denne validering gentages med jævne mellemrum for at sikre, at sammenhængen fortsat er stærkt, senest i 2021, hvor 284 hollandske elever, spredt over skolens klassetrin, tog både RD4-testen og CITO's matematiktest.

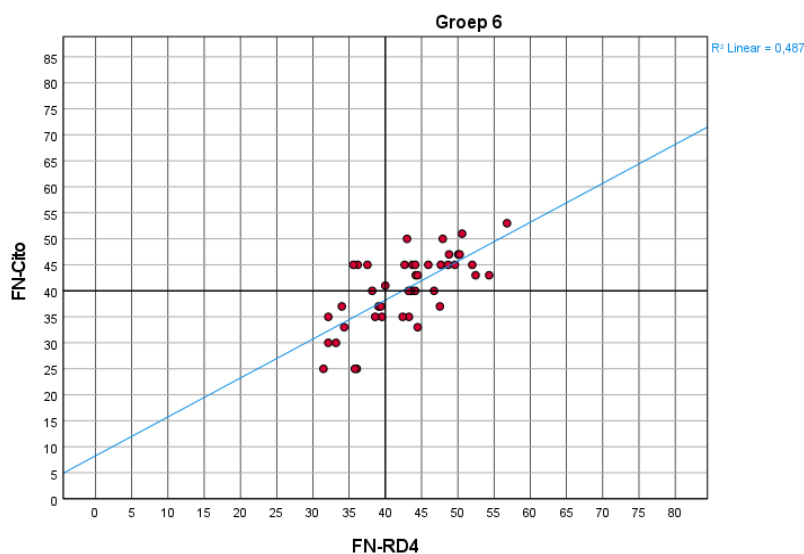
Som vist i Figur 37 ses en overordentlig pæn sammenhæng mellem resultater fra RD4 og CITO ( $R^2=0,798$ ). Én af årsagerne til, at sammenhængen er så markant, er, at når valideringen bygger på elever fra 1.-9. klassetrin, så vil materialet indeholde en stor spredning i præstationerne.

**Figur 37: Korrelation mellem resultater fra CITO- og RD4-testen, alle klassetrin**



Afprøvningen af dette projekts samtaletest omfatter kun elever fra 4. klassetrin, hvorfor der må forventes en betydeligt mindre spredning i resultaterne. En stærk sammenhæng som den, der ses i Figur 37, kan vi ikke forvente at finde, når vi kun ser på elever fra et enkelt klassetrin alene af den årsag, at spredningen i elevernes færdigheder vil være langt mindre. For at skabe et reelt sammenligningsgrundlag er resultaterne fra 4. klasseeleverne trukket ud af figuren ovenfor og vist i Figur 38. Som forventet er korrelationen mellem CITO-testens resultater og RD4-testen betydeligt svagere ( $R^2=0,487$ ) end den, der fremgik i Figur 37.

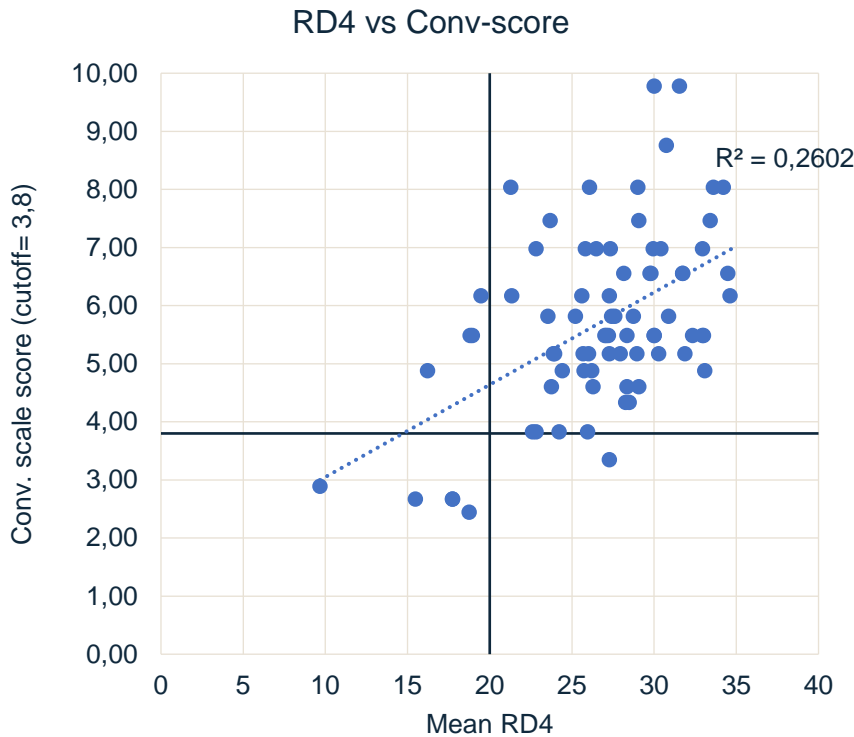
**Figur 38: Korrelation mellem resultater fra CITO- og RD4-testen, 4. klassetrin**



### 8.3.8 Validering af samtaletesten med RD4

Da RD4 i forvejen er valideret<sup>29</sup>, er det muligt at anvende denne som grundlag for at kriterievalidere den danske samtaletest på baggrund af data fra samme population (N=76) ved begge test. Valideringen er baseret på elever fra de otte kerneskolers resultater i hhv. RD4-testen og samtaletesten. Sammenhængen mellem resultaterne fra RD4-testen og projektets samtaletest (på engelsk, conversational test/conv. score) fremgår af Figur 39.

Figur 39: Korrelation mellem resultater fra RD4- og samtaletesten, 4. klassetrin



Som følge af, at det alene er 4. klasseelever, og det, der måles i de to test, ikke er helt det samme, vil man forvente en relativt svagere sammenhæng mellem RD4-testens resultater og samtaletestens. Hvor samtaletesten i højere grad er fokuseret på enkelte dele af matematikken med særlig fokus på talforståelse og aritmetiske færdigheder, dækker RD4-testen det fulde curriculum<sup>30</sup> indenfor matematikfaget.

Set i lyset heraf er en korrelation på  $R=0,51$  ( $R^2=0,26$ ) et rimeligt resultat, som peger på, at projektets samtaletest måler det, som er hensigten.

## 8.4 Samspillet mellem den digitale test og samtaletesten

I tragmodellen gennemføres samtaletesten, hvis den digitale tests resultater peger på, at der er risiko for, at eleven er talblind. Det rejser et spørgsmål om, hvorvidt den form for talblindhed, som de to test måler, faktisk repræsenterer det samme konstrukt. Dette kan dels undersøges ved at sammenligne de to testtypers respektive konstruktioner, fordele og ulemper. Derudover kan det belyses og undersøges ved at se på, hvordan elever, der har gennemført begge test, faktisk scorer

<sup>29</sup> Bunck, M. J. A., Terlien, E., van Groenestijn, M., Toll, S. W. M., & Van Luit, J. E. H. (2017). Observing and analyzing children's mathematical development, based on action theory. *Educational Studies in Mathematics*, 96(3), 289-304.

<sup>30</sup> Det hollandske og det danske curriculum for matematik har stærke lighedspunkter.

i hver af disse. Med udgangspunkt heri vil en konkret anbefaling til den praktiske udmøntning af tragtmodellen kunne undersøges.

## 8.4.1 Fordele og ulemper ved hver type af test

Rent testfagligt er den digitale test og samtaletesten grundlæggende forskellige konstruktioner, som hver især har meget forskellige fordele og ulemper – disse er uddybet i oversigten nedenfor:

	Fordele	Ulemper
<b>Digital test</b>	Den digitale test er hurtig, let og administrativt enkel at gennemføre som gruppeprøve for en lærer/vejleder/skole. Den giver på ca. 15 min. et statistisk meget stærkt resultat baseret på gennemsnitligt 300 opgaver pr. elev. Testens skalaer udgør redskaber til at afdække noget, der ikke lader sig belyse på anden vis. Kan gennemføres mange gange uden test-retest-problemer (egnet til at vurdere progression).	Testopgaver og -resultater er abstrakte indikatorer på problemstillinger, der ikke ligner dagligdags undervisning, og læreren har typisk intet udbytte af adgang til responser. Det kan være vanskeligt for den enkelte lærer at bruge og fortolke testresultaterne, da scoringsmodellen vil forekomme ugenomsagelig. Ved ethvert måleredskab er der fejlkilder, og det kræver en betydelig indsigt i testteori og processer ved afviklingen at afsløre en dårlig måling.
<b>Samtale-test</b>	Samtaletesten gennemføres i dialog mellem eleven og læreren, der har en meget direkte tilgang til proces og resultater. I modsætning til den digitale test ligner opgaverne materiale, der kunne indgå i matematikundervisningen. Det medfører, at prøvesituationen bliver mindre markant, og at lærerens observationer i forbindelse med gennemførelsen ofte anvendes direkte som input til undervisningen.	Samtaletesten er individuel, tager normalt 20-50 min. med efterfølgende bearbejdning (dvs. forholdsvis ressourcekrævende at gennemføre). Resultatet er stærkt påvirkeligt af, hvilken gennemførelsesstil og bedømmelseskriterier den enkelte lærer/vejleder anlægger. Da der er relativt få opgaver, giver det en relativt lav reliabilitet (stor statistisk usikkerhed) – især for meget stærke elever, men også for de lavest præsterende. Der er tydelige test-retest-problemer ved gentagen brug af samme opgavesæt.

Kort sagt er den digitale test kvantitativ, statistisk stærk med fokus på reliable og valide scorerer (del- og totalscorer), mens samtaletestens forcer ligger på det relationelle og kvalitative i svarprocessen samt responser på de enkelte opgavers anvendelighed i undervisningsøjemed.

## 8.4.2 Sammenhæng mellem resultater i digital test og samtaletest

Jf. Tabel 15 er der 1.659 elever, der har deltaget i både en digital test og en samtaletest. I forlængelse af tragtmodellen, der beskriver en udskilningsproces, er der grundlag for to forskellige måder at beskrive samspillet mellem de to testtyper på: 1) Hvis der er en meget stærk korrelation mellem testresultatet i de to test, kan det lede til en antagelse om, at de måler den samme form for talblindhed (samme konstrukt, samme dimension). Det vil i så fald kunne diskuteres, hvorvidt det er meningsfuldt at have to separate målinger. 2) Hvis der på den anden side slet ikke er nogen korrelation, kan det lede til at overveje, om det er noget så grundlæggende forskelligt, der måles, at det ikke giver mening at lade det omfatte af samme batteri.

Som det fremgår af

Figur 40, er der en signifikant, men relativt svag korrelation mellem resultaterne af de to test ( $r=0,36$ ). Det tyder altså på, at de to test måler nogle grundlæggende forskellige færdigheder.

Figur 40: Resultater fra elever, der har taget både digital test og samtaletest (N=1.659)



### 8.4.3 Estimation af Cutscores

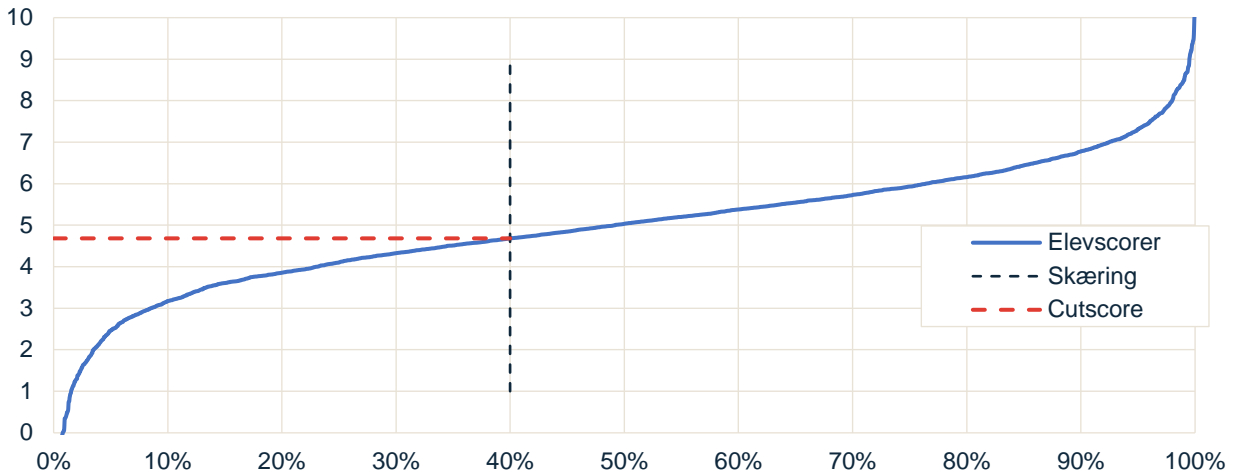
Tragtmodellen bygger på en metode, der går ud på at rangordne eleverne og frasortere dem i trin efter faldende score. Grænsen for hvornår en elev i tragtmodellen på et givet trin anses for at være i risiko bestemmes altså ved at beregne en score. Denne score er skæringspunkt for, hvornår elever anses for at være uden for risiko for talblindhed. Denne grænseværdi betegnes efterfølgende som cutscoren.

Elever i 4. klasse, der ligger over de 10 pct. lavest præsterende i den digitale test på landsplan, er ifølge modellen sandsynligvis uden for risiko for talblindhed. Samtaletesten har derefter til formål at frasortere de 70 pct. højest præsterende blandt disse 10 pct. Derved vil der efter samtaletesten være 3 pct. af eleverne, der ligger under begge cutscores og dermed er i risiko for talblindhed.

Indledningsvist i projektet har vi bedt de tilmeldte skoler om at identificere de 25 pct. (4-6 elever pr. klasse) lavest præsterende i matematik (bl.a. ved hjælp af resultater fra nationale test og andre prøver) baseret på en antagelse om, at denne gruppe vil omfatte alle de elever, der er i risiko for talblindhed. Populationen udgjorde 3.511 elever, der har besvaret den digitale test. Vores efterfølgende beregninger bygger på, at skolerne har været i stand til at identificere de lavest præsterende elever, og at den stikprøve, som udgør forsøgets population, udgør et tilnærmelsesvist repræsentativt udsnit af de 25 pct. lavest præsterende danske 4. klasseelever på de deltagende skoler. Identifikation af de lavest præsterende 10 pct. på landsplan udgør dermed de 40 pct. lavest præsterende i den digitale test, da stikprøven i projektet omfatter 25 pct. af eleverne.

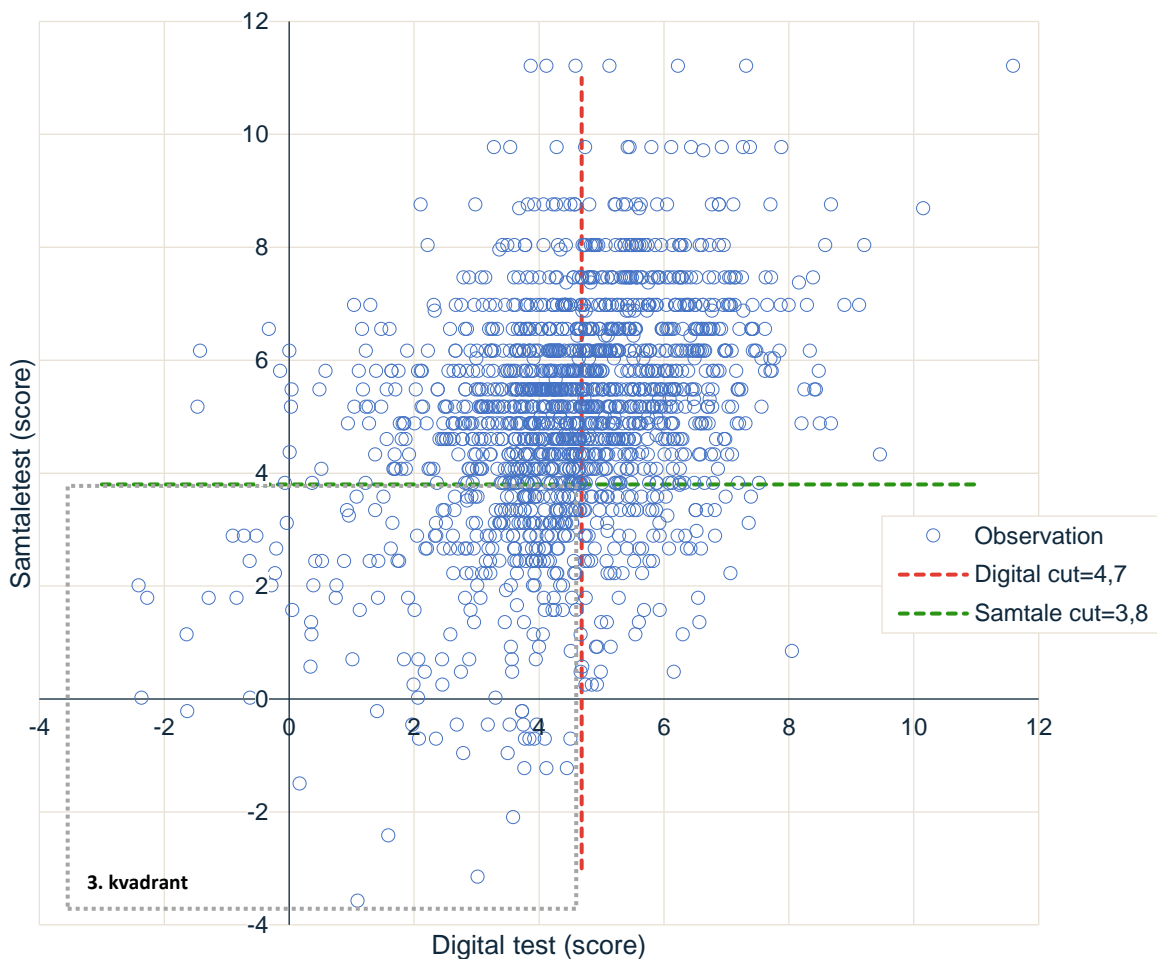
Metoden til at estimere cutscoren for den digitale test er, at elevernes resultater listes i rangorden, hvorefter elevscoren for den elev, der ligger på 40-percentilen (svarende til  $3.511 \cdot 0,4 = 1.404$  elever), identificeres. I den digitale test er denne score blevet bestemt til 4,7 som vist i Figur 41.

Figur 41: Fordeling af elevscorer (N=3.511) og bestemmelse af cutscore=4,7 (ved percentil=40)



Metoden for beregning af cutscoren ved samtaletesten: Af de 1.659 elever, der har taget begge, identificeres de elever, der ligger under cutscoren til den digitale test (i alt 914 elever). Disse elever rangordnes efter deres samtaletest-score, hvorefter cutscoren for elever under 30-percentilen (i alt 274 elever) identificeres. Samtaletestens cutscore er således estimeret til 3,8. De to cutscorer er indtegnet på Figur 42, og eleverne i risiko for talblindhed er således dem i 3. kvadrant (markeret med en smal stiplede linje).

Figur 42: Resultater fra elever, der har taget både digital test og samtaletest (N=1.659)



## 9. Samarbejde med kerneskolere

Beskrivelse af rekruttering og udvælgelsesproces er beskrevet ovenfor i afsnit 5.2 om 'Rekruttering af kerneskolere og reserveskolere'.

Kerneskolerne havde overordnet tre centrale roller i projektet:

1. At udvælge/justere/udvikle undervisningsforløb målrettet egne elever i risiko for talblindhed i samarbejde med projektet, hinanden og ekspertgruppen (som de indgik i) samt at afprøve de udviklede forløb.
2. At indsamle og formidle erfaringer med processen, test og de pædagogiske forløb.
3. At afprøve alle dele af den samlede test mindst en gang. Den digitale test skulle afprøves to gange – både før og efter undervisningsforløbene – blandt andet med henblik på måling af effekten af forløbene.

### 9.1 Udvikling og afprøvning af pædagogiske indsatser

For at facilitere udviklingsprocessen og dialogen med og mellem skolerne blev der indledningsvist aftalt et set-up, som skulle kunne holdes inden for en relativt snæver ressourcemæssig ramme og samtidig levere de ønskede resultater.

I Tabel 18 fremgår aktivitets- og ressourceplanen for samarbejdet i forbindelse med kortlægning, udvikling og afprøvning af lovende pædagogiske indsatser. Oversigten markerer de dele, der vedrører samarbejdet mellem projektet og kerneskolerne. Selve udviklingsopgaven blev dimensioneret efter skolernes råderum til opgaven samt kommunale retningslinjer for undervisning i forbindelse med COVID-19. Afprøvningen har således været præget af et tilpasningsbehov til den omskiftelige kontekst af restriktioner i skolepraksis, fx hvad angår at lave mindre specialundervisningshold med elever fra forskellige klasser.

**Tabel 18: Opdateret aktivitets- og ressourceplan**

	Uge	Indhold	Varighed	Forberedelse
Kortlægning	Uge 50	Fælles intromøde	1 time	Ingen
	Uge 51	Møde med kerneskolerne parvis	1 time	Orientering i dokumenter fra emu (½ time)
	Uge 2	Opsamling mellem skolerne parvis	1 time	Ingen
	Uge 3	Igangsættelse af udviklingsprocessen	1 time	Ingen
	Uge 6	Deltagelse i ekspertgruppemøde	2 timer	½ time
Udvikling	Uge 7/8	Møde med kerneskolerne parvis	1 time	Afhænger af skolernes indsats/ressourcer
	Uge 10	Afrunding på udviklingsprocessen	1 time	Afhænger af skolernes indsats/ressourcer
	Uge 11	Afholdelse af to seminarer med eksperter	2 timer	½ - 1 time
Afprøvning	Uge 12-23	Afprøvning		

Grundet praktiske udfordringer med afprøvning relateret til COVID-19-restriktioner, blev afprøvningsperioden for indsatserne forlænget i takt med, at fristen for gennemførelse af testrunde 2 blev udskudt. Bestræbelsen var at give skolerne maksimal fleksibilitet til at komme længst muligt, idet der dog blev lagt vægt på, at anden testrunde skulle gennemføres, efter indsatserne var gennemført (i størst muligt omfang).

Tidsplanen for afprøvning af observationsguide, digital test og samtaletest, runde 1 og 2, har fulgt planen for de øvrige deltagende skoler.

### 9.1.1 Kortlægning

Som grundlag for kerneskolernes udviklingsarbejde blev forud afsøgt, hvilken eksisterende viden der forelå på feltet. I regi af arbejdet med talblindhed har først SFI og siden Epinion udarbejdet kortlægninger over pædagogiske indsatser i forbindelse med talblindhed. Begge disse er baseret på en gennemgang og review af forskningsartikler. I den offentliggjorte forskningsoversigt ([Se Forskningsoversigten på EMU](#)) fremgår i afsnit 12.3 en oversigt over "Virkningsfulde pædagogiske indsatser/tiltag/værktøjer og redskaber i forskningen". Denne oversigt har udviklingsarbejdet taget afsæt i, og det er også tydeligvis inspireret heraf. Herfra har der dog kun kunnet trækkes begrænset viden om indsatsernes præcise udformning, så grundlæggende er denne kortlægning funderet i en praktisk tilgang med udgangspunkt i erfaringer fra forskningsoversigten, projektets ekspertgruppe og koordinatorene på kerneskolerne.

De fleste materialer, der blev nævnt af koordinatorene, er en kombination af kortlægningsmateriale og undervisnings-/indsatsmateriale. Kortlægningsmaterialer skal forstås som materialer, der kan illustrere, hvorvidt, og på hvilke områder, eleven lader til at have udfordringer, men som ikke nødvendigvis er målrettet håndtering af disse udfordringer (såsom Hogrefes MG-prøver og MAT-test). Kombinationsmaterialerne består ofte af en test eller anden form for materiale til screening af elevernes matematiske færdigheder og dernæst inspiration til, hvad der kan gøres ved dette, men sjældent af et færdigt forløb, der kan tages direkte i brug. Materialerne, der blev nævnt, anvendes overvejende til kortlægning, og det er få materialer, der anvendes til egentlig undervisning af elever med matematikudfordringer. Flere af koordinatorene udtrykte behov for yderligere materiale til undervisning – særligt målrettet de ældre elever (mellem- og udskolingsniveau).

Materialer, der hovedsageligt anvendes til kortlægning:

- "Matematikvanskeligheder - Tidlig intervention" af Lena Lindenskov og Peter Weng
- "M-OBS, Matematisk observation i børnehaveklasse og indskoling" af Jesper Hviid Andersen, Michael Wahl Andersen og Nancy Fuglsang
- "Rummelighed i matematik" af Michael Wahl Andersen
- "TIM (Tidlig Indsats i Matematik)"
- Björn Adlers Matematikscreening I, II og III
- MAT-test og MG-prøver (Begge fra Hogrefe)
  - Bruges bl.a. som samtaletest (ligesom RD4) – brugt "adaptivt" ved at vælge test fra højere/lavere klassetrin.

Materialer, der anvendes til indsats/undervisning:

- "RoS - Regning, Observation og Strategi" af Pernille Pind
- "Matematikvanskeligheder - Tidlig intervention" af Lena Lindenskov og Peter Weng
  - Herunder "TMTM - Tidlig Indsats Til Marginalgrupper i matematik"

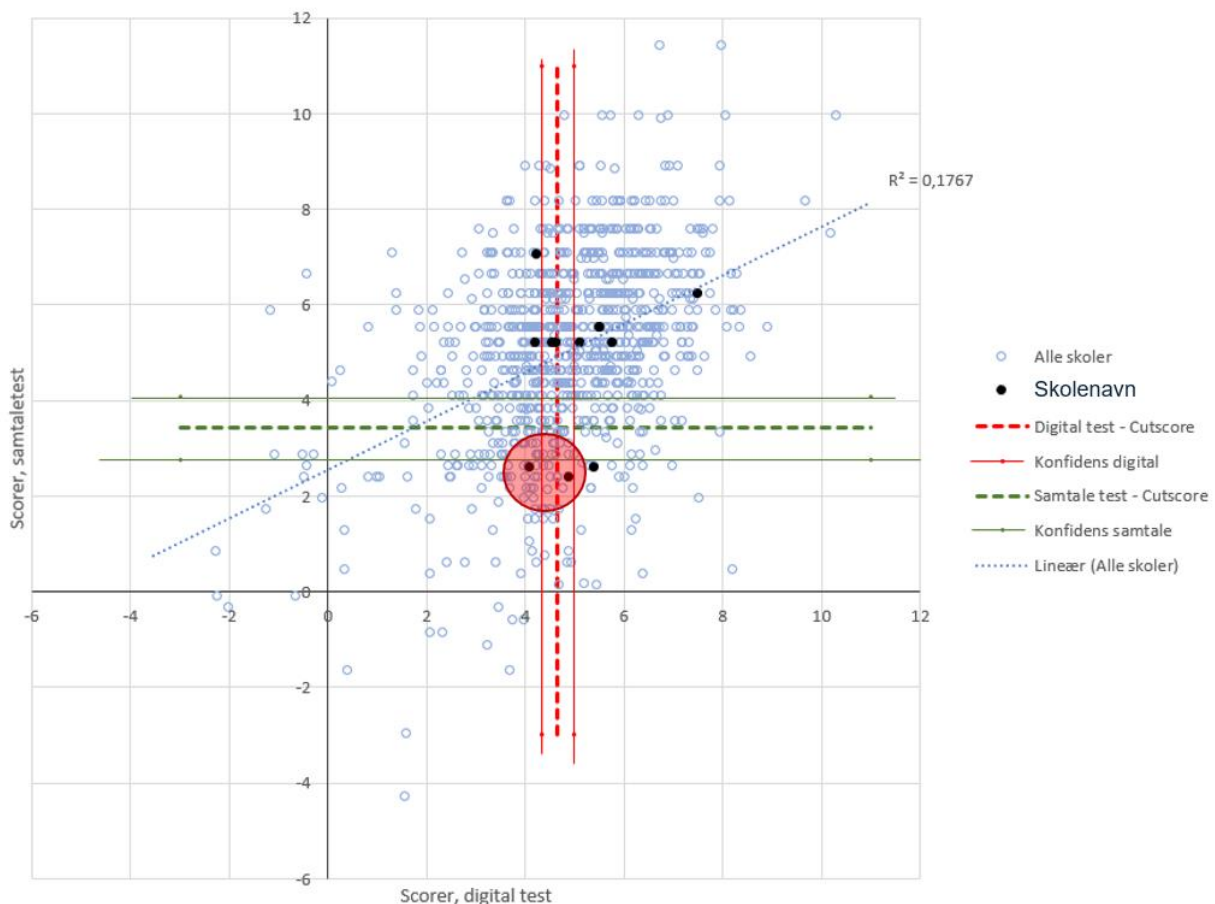


### 9.1.2 Tilbage melding af resultater og udvælgelse af elever

Udvælgelsen af elever til indsatserne er hovedsageligt foregået med afsæt i resultater fra runde 1 af den digitale test og samtaletesten samt resultater fra RD4. Som tilbage melding på testene fik koordinatorene tilsendt et dokument bestående af 1) et diagram med visualisering af deres elevers resultater fra den digitale test og samtaletesten som vist i figuren nedenfor og 2) Mieke van Groenestijns<sup>31</sup> analyser og vurderinger af elevernes risiko for talblindhed på baggrund af resultater fra RD4-testen (se RD4-testen kort beskrevet på engelsk: [RD4-in-brief](#)) samt anbefalinger til fokusområder til undervisning af eleverne (Bilag 1). Således indebærer både analyserne på baggrund af den digitale test, samtaletesten og RD4 en risikovurdering af, om eleven udviste tegn på talblindhed.

I diagrammet over resultaterne fra den digitale test og samtaletesten (jf. Figur 42) er markeret cutscorer for begge test. I Figur 43 markerer en tynd linje den statistiske usikkerhed, der gælder for cutscoren. Disse to linjer markerer konfidensintervallet, der er det område hvori vi ikke med rimelig sikkerhed kan afgøre, om eleven ligger over eller under cutscoren. De elever, der lå i konfidensintervallet, opfordrede vi koordinatorene til at inddrage i de pædagogiske indsatser sammen med eleverne under cutscoren.

Figur 43: Eksempel på oversigt over elevresultater på en kernesko



<sup>31</sup> Medlem af projektets ekspertgruppe og medforfatter til RD4-testen.

I Figur 43 fremgår eleverne fra en af kerneskolerne (de sorte prikker). Det ses, at der er to elever (markeret ved den røde cirkel), der opfylder kriterierne for at blive udvalgt til indsatserne i dette projekt; én, der ligger under begge cutscorer, og én, der ligger under cutscoren for samtaletesten og over cutscoren, men indenfor konfidensintervallet, for den digitale test.

Resultaterne fra RD4 blev formidlet i form af en RD4-profil (som koordinatorene fik adgang til i umiddelbar forlængelse af testene) samt en analyse af elevens risiko for talblindhed med afsæt i denne profil (vedhæftet i tilbagemeldingsdokumentet). Der var generelt god overensstemmelse mellem Mieke van Groenestijns analyser og vurderingerne på baggrund af den digitale test og samtaletesten.

Eleven fra figuren, som i talblindhedstesten ligger under begge cutscorer, fik følgende tilbagemelding fra Mieke:

**At risk for dyscalculia.** Severe difficulties on all subdomains.

**Advice:** further diagnostic testing, compare with results on other domains. Start with composition of numbers, order, and positioning, writing and naming numbers. Addition and subtraction with real hands-on materials in relation with writing the procedures on paper. Give meaning to numbers. Talk about numbers in real life situations.

Eleven fra figuren, der i talblindhedstesten lå under den ene cutscore og i konfidensintervallet for den anden, fik følgende tilbagemelding fra Mieke:

**At risk for severe difficulties, might be dyscalculia.** Difficulties on number, counting, order and partitioning. Understands composition (Hundreds, Tenths, Ones). Writing and naming numbers, composition and addition is good. Difficulties with subtraction, multiplication, and division. Shows some basic knowledge on relations.

**Advice:** Good and careful intervention at number (counting forwards, order, positioning, partitioning in relation with composition) and basic operations. Start with doing computations with nice numbers (tenths, hundreds, and fifths). Use calculator for doing basic and further complex operations.

Som en yderligere faktor til udvælgelsen af elever har COVID-19-restriktioner resulteret i, at nogle af koordinatorene har måttet prioritere ift., hvilke elever der udvælges. Opfordringen har generelt været, at de elever, hvis resultater ligger længst under begge cutscorer, bør prioriteres.

I Tabel 19 nedenfor er givet et overblik over de forskellige aktiviteter og elevgrupper på de 8 kerneskolers. Tabellen viser således forskellige oplysninger om kerneskolerne og deres elevgrundlag for deltagelse i projektet, opdelt i 1) basisoplysninger, 2) oplysninger om gennemførte tests og 3) oplysninger om elever udvalgt til indsats. Som det ses, var der oprindeligt udvalgt 97 kerneskolerelev til at deltage i projektets aktiviteter, hvor der var forventet knapt det dobbelte antal. Det beror sandsynligvis på, at kerneskolernes elevgrundlag (pga. udvælgelseskriterier for kerneskolerelev) er en del bedre præsterende, end hvad der er repræsentativt for elever i 4. klasse generelt.

Af de 97 elever er 24 udvalgt til at deltage i afprøvningen af de udviklede pædagogiske indsats med afsæt i ovenstående udvælgelsesmetode.

**Tabel 19: Oversigt over antal kerneskolerelev, der har været udvalgt til/deltaget i projektets aktiviteter**

Antal kerneskolerelev	Kernescole, nummer								I alt
	1	2	3	4	5	6	7	8	
<b>1) Basisoplysninger om de deltagende elever</b>									
Antal elever på 4. årgang på skolen	68	69	94	124	47	99	110	118	729
Oprindeligt forventet antal elever udvalgt til projektet (25 pct. af antal elever på årgangen)	17	17	24	31	12	25	28	30	182
Antal elever udvalgt af skolerne til projekt	16	11	18	13	5	7	21	6	97
<b>2) Oplysninger om gennemførte tests</b>									
Antal elever udvalgt til RD4	16	11	18	13	5	7	8	6	84
Antal elever gennemført RD4*	11	10	17	11	5	6	7	4	71
Antal elever gennemført digital test	16	11	20	14	8	8	6	7	90
Antal elever gennemført samtaletest	12	10	20	14	8	7	6	7	84
Antal elever gennemført både digital og samtaletest	12	10	20	14	8	7	6	7	84
<b>3) Oplysninger om elever udvalgt til indsatser</b>									
Oprindeligt forventet antal elever med tegn på talblindhed (3 pct. af antal elever på årgangen)	2	2	3	4	1	3	3	4	22
Antal elever med tegn på talblindhed baseret på cut-scores for både digital- og samtaletest	1	0	1	0	1	0	0	0	3
Antal elever med tegn på talblindhed inkl. elever i konfidensintervallerne**	2	1	2	1	1	1	0	1	9
Antal elever udvalgt til indsats	2	3	7	2	1	0	8	1	24

\*Nogle af koordinatorene valgte at gennemføre RD4-testen med elever, der ellers ikke var inddraget i projektet (fx fra andre årgange). Disse elever er ikke inddraget i tabellen.

\*\*Dvs. elever, der ligger under cutscoren eller inden for den statistiske usikkerhed.

Af Tabel 19's nederste del (Oplysninger om elever udvalgt til indsatser) fremgår hhv. forventet antal elever med tegn på talblindhed og faktisk antal elever med tegn på talblindhed. Som det ses, er der 3 elever, der ligger under begge cutscores – og yderligere 6 elever, der ligger over en eller begge cutscores, men inden for konfidensintervallerne, dvs. 9 i alt. Det forventede antal af elever med tegn på talblindhed (22) var godt dobbelt så højt som det faktiske (9), hvilket stemmer godt overens med, at kerneskolerne er udvalgt, så de gennemgående har fagligt stærkere elever end gennemsnittet af skolerne.

I tabellens sidste række fremgår antallet af elever, der har deltaget i de forskellige indsatser, som er udviklet og afviklet kommunevis på hhv. skole 1 og 2 (indsats 1), skole 3 og 4 (indsats 2), skole 5 og 6 (indsats 3) og skole 7 og 8 (indsats 4).

Indsatserne gennemgås nedenfor i afsnittene 9.1.4 – 9.1.7

## 9.1.3 Udvikling og afprøvning af indsatser

Udviklingen af pædagogiske indsatser er sket parvis i samarbejde mellem skolerne i hver kommune, bl.a. med afsæt i materialer fra ovenstående kortlægning, lærernes erfaringer og vurderinger af elevens faglige niveau, resultater fra RD4-testene samt enkelte udvalgte artikler og rapporter til inspiration<sup>32</sup>.

Ved et introducerende møde, som havde til formål at afklare eventuelle overvejelser, bekymringer, udfordringer og muligheder, der skulle tages til efterretning eller overvejelse, blev det bl.a. påpeget, at der på flere af skolerne var ingen eller meget få af de testede elever, der i henhold til testresultaterne var i risiko for talblindhed, jf. Tabel 19.

Denne udfordring blev bragt op på ekspertgruppemøde 2, hvor koordinatorene fra hver kerneskole deltog og fremlagde deres ideer og forslag til indsatser med henblik på sparring og feedback til inspiration. Eksperterne foreslog, at koordinatorene udvalgte de elever, der umiddelbart havde størst behov for ekstraundervisning og hvor udfordringerne i et vist omfang kunne overlappes med talblindhed (fx hvis de havde svært ved de grundlæggende regningsarter, plus, minus, gange og dividere), hvilket understøttede den generelle opfordring om at udvælge de elever med de laveste resultater i den digitale test og samtaletesten.

Eksperterne gav derudover både konkret feedback til de fremlagte forslag til indsatser og generelle råd og anbefalinger til undervisning af elever med talblindhed. I forlængelse af denne drøftelse blev der efter ønske fra de deltagende skoler arrangeret to seminarer (et med deltagelse af dansksprogede eksperter og et med de engelsksprogede), hvor eksperterne fik lejlighed til at fortælle om deres erfaringer, og hvor koordinatorene fik mulighed for at få uddybende feedback og inspiration til indsatserne.

Der var en klar opfattelse blandt koordinatorene af, at projektet i særlig grad var målrettet afprøvning af test. Men det blev eksplicit fremhævet, at de var glade for, at projektet rummer et undervisningsaspekt ift., hvordan man kan hjælpe elever med tegn på talblindhed. En af koordinatorene udtalte således:

*”Jeg synes, det giver super meget mening, at vi har den her anden side [udvikling af pædagogiske indsatser] med i hele projektet for os, der varetager elevens og lærerens perspektiv. Nu har vi trukket dem igennem en hollandsk test og to runder dansk test, så vil vi også gerne have lov til at hjælpe eleverne, og det kan vi så gøre her”.*

Afprøvning af indsatserne blev sat i gang i uge 12 og afviklet med inspiration i princippet fra bl.a. aktionslæring om, at forløbet evalueres og tilpasses mellem hver undervisningsgang. Nogle af koordinatorene gennemførte indsatsen næsten som planlagt, men grundet praktiske udfordringer i relation til COVID-19-restriktioner måtte størstedelen af forløbene gennemføres delvist (med færre elever og/eller færre lektioner end planlagt), og en enkelt koordinator fik slet ikke mulighed for at begynde indsatsen på grund af lokale restriktioner på skolen.

---

<sup>32</sup> Mahmoud, M. S., Zainal, M. S., Rosli, R. & Maat, S. M. (2020) Dyscalculia: What We Must Know about Students' Learning Disability in Mathematics? Universal Journal of Educational Research 8(12B), Faculty of Education, Universiti Kebangsaan Malaysia, Malaysia  
 Undervisningsministeriet (2019a) Pædagogisk vejledning: Arbejde med de rationale tal i praksis.  
 Undervisningsministeriet (2019b) Pædagogisk vejledning: Arbejde med de naturlige tal i praksis.

Da indsatserne er udviklet parvis af koordinatorene fra hver kommune, er der udviklet fire indsatser, som er gennemført parallelt. Afviklingen er sket med små forskelle mellem skolerne med hensyn til de enkelte elevers behov og niveau, men med fælles grundlæggende principper for, hvad der blev afprøvet.

Som fællesnævner var alle indsatserne struktureret som særundervisning uden for klassens almindelige rum og skema i små hold eller én-til-én med den/de udvalgte elever. Indholdet af indsatserne varierede derimod mellem kommunerne, men som et grundlæggende princip er der lagt vægt på at visualisere og konkretisere tal, mængder og regnestykker på forskellige måder, fx vha. terninger, penge, ternet papir, centicubes, penge mv. Dette var både baseret på koordinatorenes erfaringer som matematikvejledere og pba. eksperternes råd til håndtering af talblindhed.

Nedenfor beskrives de afprøvede materialer, forløb og indsatser på de fire kommuners kernesko-ler. Koordinatorernes beskrivelser og materiale tilhørende indsatserne er vedlagt i bilag 2-5.

### 9.1.4 Indsats 1: Addition og multiplikation

I kommune 1 blev et eksisterende materiale med dertilhørende forløb (Pernille Pinds RoS-kuffert) anvendt som grundlag for indsatsen. I denne kommune var der grundlæggende forskellige udgangspunkter for de to skolars elever, idet der på den ene skole var to elever, der blev vurderet i risiko for talblindhed, og på den anden var der ingen – dette med afsæt i resultater fra både den digitale test og samtaletesten samt analyserne af resultater fra RD4-testen.

De to skoler har hver især arbejdet med kontinuerlige fokusområder, hhv. multiplikation og addition. Additionsforløbet er oprindeligt tiltænkt elever i 1-2. klasse, og multiplikationsforløbet er tiltænkt elever på 3-6. klassetrin. Fokusområderne er udvalgt på baggrund af kendskab til eleverne, deres faglige niveau og konkrete udfordringer.

Indsatsen er særligt fokuseret på at anvende konkrete materialer til at omsætte regnestykker, eksempelvis perlekæder og tallinjer. Anvendelse af konkrete materialer var en gennemgående god erfaring fra koordinatorene samt anbefaling fra eksperterne:

*“My advice is to use real-life materials, like money, because you can feel them. Development of number sense and quantity is unlike numbers related to telephone numbers, time, etc. The basis of number sense is to use real-life objects, and in particular important for children with an indication of dyscalculia”* (Mieke van Groenestijn, Ekspertgruppemøde 2)

På den ene skole, hvor der blev anvendt et materiale, der som udgangspunkt var tiltænkt lavere klassetrin, var erfaringen, at når eleverne er 'massivt udfordret' i matematik, er det en fordel at planlægge indsatsen på en måde, der giver eleverne nogle succesoplevelser. Koordinatoren beskriver, at additionsstykkernes lave niveau betød, at der var mere fokus på, hvordan eleverne regnede, end hvad de regnede. Beskrivelsen af indsatsen på denne skole er yderligere udfoldet i Bilag 2A.

På den anden skole, hvor indsatsens niveau i højere grad var alderssvarende, oplevede koordinatorene, at *“eleven får god træning i at omsætte abstrakte gangestykker til forskellige konkrete repræsentationer”*. Beskrivelsen af indsatsen på denne skole er yderligere udfoldet i Bilag 2B. Dog understreges vigtigheden af, at indsatsen tilpasses den enkelte elevs faglige niveau og læringshastighed.

### 9.1.5 Indsats 2: 10'er venner og farverepræsentationer

I kommune 2 blev der arbejdet med varierende fokus fra gang til gang, hhv. positionssystemet, systemer og talmønstre/tabeller, tal i hverdagen, partition og regnestrategier, addition og subtraktion. Der var i alt udvalgt ni elever til indsats fordelt på de to skoler (hhv. to og syv elever). Ud af disse ni elever var der to elever, som på baggrund af den digitale test og samtaletesten lå i risikozonen for talblindhed, én fra hver skole.

På de to skoler blev der bl.a. arbejdet med 10'er venner og 100'er venner i forskellige farverepræsentationer, overslagsregning og regning med penge. Farverepræsentationerne af 10'er og 100'er venner er illustreret i Figur 44. Beskrivelsen af indsatsen samt tilhørende materialer afprøvet på disse skoler er yderligere udfoldet i Bilag 3A og 3B.

Til et ekspertgruppemøde påpegede Steve Chinn, at farverepræsentationerne kan fungere som en god måde at skabe opmærksomhed om de forskellige cifre, men at det er vigtigt at understøtte dette med en mængderepræsentation gennem visualiseringer eller konkrete materialer.

På de to skoler er der arbejdet med mængderepræsentationer i form af penge, talbrikker, centicubes, 10'erstænger og hundredeplader, bl.a. gennem et webbaseret program ([Number Pieces by The Math Learning Center](#)), hvor eleverne kan flytte rundt på figurer af hhv. én, ti og hundrede klodser. I forbindelse med regnestrategier er der anvendt en metode, hvor eleverne selv visualiserer regnestykker<sup>33</sup> ved at tegne forskellige former, der repræsenterer tallenes værdi (firkanter er 100'er, streger er 10'ere og prikker er 1'ere). Desuden er der gennemgående arbejdet med at anvende spil, hvori der indgår tal og regnestykker i forskellige afarter, herunder terningspil, kortspil, vendespil og pilekast.

Figur 44: Farverepræsentationer af 10'er og 100'er venner

1	9	10	90
2	8	20	80
3	7	30	70
4	6	40	60
5	5	50	50

### 9.1.6 Indsats 3: Talforståelse: talsymboler og tallenes orden

I kommune 3 tog indsatserne indledningsvist udgangspunkt i de tre tematikker: subitizing<sup>34</sup>, talforståelse og at tælle, hhv. baglæns og 'skip counting'/tabeller.

Udfordringer inden for delområdet 'subitizing' er i forbindelse med gennemgang af RD4-testens delområder blevet fremhævet som en af indikatorerne på talblindhed og er ligeledes inkorporeret som et parameter i den digitale test. Dog blev det til et ekspertgruppemøde pointeret, at udfordringer i forbindelse med subitizing i højere grad skal betragtes som en indikator for talblindhed end et fokusområde for undervisning:

*"It's important not to think that if you train subitizing, you will get better math skills. It's not the way it works. If you make exercises in subitizing, you will get better at subitizing, but there is not a strong relationship to addition or subtraction, for example, but it's a good indicator"* (Bent Lindhardt, Ekspertgruppemøde 2).

<sup>33</sup> Skolen har givet lov til at denne video kan deles til formål i tilknytning til projektet.

<sup>34</sup> Evnen til automatisk at genkende mængder op til 5 uden at tælle.

[...] and subitizing is only up to five, so that's not developmental, that's an indication of early very poor number sense. (Steve Chinn, Ekspertgruppemøde 2)

Den endelige indsats var centreret omkring talforståelse, herunder talsymboler og tallenes orden. Dette med særlig fokus på at anvende gentagelser. Hvor undervisningsgangene havde fokus på forskellige talsymboler, gik rammen for undervisningen igen for hver lektion. Hver lektion blev endvidere afrundet med matematikspillet Multi Math. Beskrivelsen af indsatsen på denne skole er yderligere udfoldet i Bilag 4.

Grundet COVID-19-restriktioner var det kun muligt at gennemføre indsatsen på én af de to skoler med én elev. Denne elev var det tydeligste eksempel på, hvad der inden for projektets forståelsesramme kan betragtes som talblindhed. Koordinatoren beskriver:

*"Interventionerne til talblinde elever må i høj udstrækning tilpasses den enkelte elev. Eksempelvis har en af mine elever sværest ved netop 3, 6 og 7, har et meget anstrengt forhold til centicubes, er umiddelbart dygtig til at læse tal og tælle baglæns og er en dygtig læser og har en meget åben tilgang til det, at han sandsynligvis er talblind"* (Bilag 4)

Tre undervisningsgange i indsatsen har været dedikeret til at arbejde med hhv. tallene 3, 6 og 7 ved at finde mange forskellige former for repræsentationer, visualiseringer og konkretiseringer af tallene og de mængder, de repræsenterer. Denne tilgang er understøttet af eksperternes anbefaling, der bl.a. retter sig mod at fokusere på sammenhængen mellem talsymboler og mængder; frem for blot at knytte tal og regnestykker an til talsymboler, bør de knyttes til konkrete materialer, der kan give eleven en fornemmelse for tallenes størrelser.

### 9.1.7 Indsats 4: Positionssystemet og gangespil

I kommune 4 blev der arbejdet med to forskellige overordnede formål, først positionssystemet og dernæst gange, med afsæt i varierende spil og materialer fra gang til gang. Her blev der bl.a. arbejdet med forskellige former for terningspil til at understøtte gange-øvelser og øvelser i positionssystemet. Materialer og beskrivelse af indsatsen på disse skoler er yderligere udfoldet i Bilag 5A, 5B, 5C, 5D og 5E.

På den ene skole var der ingen elever, der lå i risikozonen for talblindhed – hverken pba. den digitale test eller samtaletesten. På den anden skole var der én elev i risikozonen; eleven lå i konfidensintervallet for samtaletesten og under cutscoren for den digitale test.

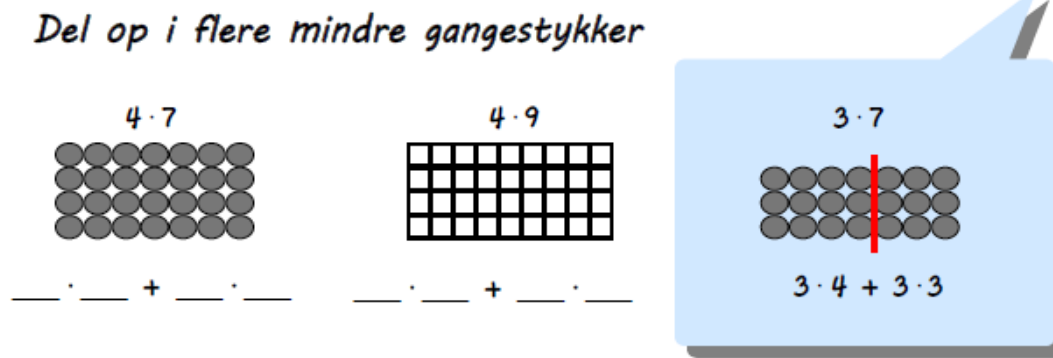
De første to lektioner var centreret om positionssystemet, hvortil der bl.a. blev anvendt et webbaseret spil ([talmat](#)), terningspil, spil med talkæder og centicubes. De resterende tre lektioner havde gangestrategier som fokuspunkt. Her blev der gennemgående arbejdet med to øvelser: 1) At eleven skulle visualisere forskellige gangestykker på ternet papir, og 2) eleven skulle opdele gangestykker, visualiseret i tern, på forskellige måder som vist i



Figur 45.



Figur 45: Opdeling af gangestykker



## 9.2 Erfaringer fra samarbejdet med kerneskolerne

Erfaringerne fra samarbejdet med og mellem skoler har været præget af, at de enkelte skolars vilkår har vist sig at være ret forskellige. Fx har elevgrundlagene, på trods af homogene udvælgelseskriterier, vist en vis variation mellem skolerne. Det har fx givet udslag i, at man i én af kommunerne ikke har kunnet finde elever med tegn på talblindhed (enkelte elever lå under cutscoren i den digitale test, men ingen af eleverne lå under samtaletestens cutscore), mens vi i de tre andre kommuner har fundet et meget begrænset antal elever med tegn på talblindhed både i den digitale test og samtaletesten.

### 9.2.1 Kan man undervise sig ud af talblindhed?

Erfaringer fra både teori og praksis peger på, at elever, som har udfordringer med de grundlæggende aritmetiske færdigheder, kan få reduceret omfanget af udfordringerne gennem træning, undervisning og passende støtte. At det er muligt at hjælpe elevernes tegn på talblindhed gennem pædagogiske indsatser er understøttet af forskningen generelt, og dette danner grundlag for projektets delopgave 3.

Som beskrevet i afsnit 8.4.2 'Sammenhæng mellem resultater i digital test og samtaletest' peger projektets data på, at den digitale test og samtaletesten måler nogle grundlæggende forskellige færdigheder. Koordinatorerne fra kerneskolerne har også observeret denne forskel. De vurderer, at samtaletestens opgaver kunne være en del af det, der normalt arbejdes med i matematikundervisningen. Opgaverne i den digitale test afviger derimod fra almindeligt undervisningsmateriale både i forhold til form og indhold.

Mens den digitale test i højere grad måler et intuitivt, neurologisk aspekt af tal- og mængdeforståelse samt grundlæggende aritmetiske færdigheder, giver samtaletesten snarere et indblik i elevens evne til at håndtere matematiske udfordringer både i matematikundervisningen og i dagligdagen. Samtaletesten giver dermed et billede af, i hvilket omfang talblinde elever har tilegnet sig metoder til at kompensere for deres neurologiske udfordringer med tal og mængder. Denne skelnen mellem, hvad de to test måler, svarer til en generel forskningsmæssig distinktion mellem neurologiske og didaktiske perspektiver på talblindhed, der bl.a. blev diskuteret og understøttet af projektets eksperter på ekspertgruppemøde 2.

Når den digitale test i højere grad er fokuseret på de neurologiske aspekter af talblindhed, er der dele af testen, hvor elevens resultater kun i begrænset omfang vil kunne forbedres gennem matematikundervisning (eksempelvis delskala 3; subitizing og delskala 1; ANS). Samtaletesten, der er

mere fokuseret på regning og anvendelse af tal, overlapper med matematikundervisningens formål, hvorved resultatet herfra i højere grad vil forbedres gennem pædagogiske indsatser i matematikundervisningen.

Denne mekanisme understøttes af fordelingen af elevernes resultater i de to test holdt overfor hinanden som ses i Figur 42 på s. 58. I figuren ses et forholdsvist lille antal elever, der ligger i 4. kvadrant af diagrammet (som scorer lavt i samtaletesten, men højt i den digitale test). Det vil typisk være sådan, at hvis eleven har en god intuitiv tal- og mængdeforståelse, målt ved den digitale test, så vil eleven også have relativt bedre forudsætninger for at klare sig godt i samtaletesten. Det gælder ikke nødvendigvis omvendt: En elev, der scorer meget lavt i den digitale test, vil – med god undervisning og grundig træning – stadig have pæn sandsynlighed for at få en høj score i samtaletesten, fordi det, der måles i samtaletesten, lettere kan påvirkes igennem pædagogiske indsatser. Ved gennemgang og analyse af kerneskolernes elevresultater i runde 1 så vi et eksempel herpå. En elev scorede under cutscoren i den digitale test, men meget højt i samtaletesten. Koordinatoren begrundede dette med, at eleven havde fået den rette hjælp og støtte fra både skolen og hans forældre under Corona-nedlukningen, og dermed havde han trænet kompenserende strategier intensivt og med god effekt.

Som vist i Tabel 19 (s. 63) fandt vi betydeligt færre elever med tegn på talblindhed end forventet på de otte kerneskolers. Dette kan måske forklares ved ovenstående mekanisme. Udvælgelsen af kerneskolers med et relativt stærkt elevgrundlag kan have betydet, at elever, som tidligere har ligget i risiko for talblindhed, er blevet hjulpet og støttet i et sådant omfang, at deres talblindhed ikke længere kunne registreres med samtaletesten.

## 9.2.2 Koordinatorernes erfaringer

Ved afslutningen af forløbet med de pædagogiske indsatser blev koordinatorenne bedt om at udveksle og dele deres erfaringer med at deltage i projektet. De gjorde det tydeligt, at deltagelse i projektet havde været en lærerig proces, som de satte pris på at have været en del af. De understregede dernæst, at der er et presserende behov for en talblindhedstest i folkeskolen, og at de håber, den snart bliver stillet til rådighed. Dernæst påpegede de, at det tidsmæssige omfang af indsatserne (1-2 lektioner i 5 uger) har været utilstrækkeligt, hvis der skal gøres en synlig forskel, og flere af koordinatorenne havde intentioner om at fortsætte indsatserne efter projektets afslutning.

De påpegede i øvrigt, at det, som eksperterne fremhævede som gavnligt for talblinde elever, også fungerer med elever, der har andre udfordringer med aritmetik, tal og mængder af anden karakter end talblindhed. Det er dog vigtigt at understrege, at koordinatorennes erfaringer tager afsæt i et meget begrænset datamateriale, og at generelle konklusioner på dette grundlag derfor er usikre.

Vi bad dem yderligere dele, hvad de så som de vigtigste og mest markante erfaringer til inspiration for andre lærere og vejledere, som står over for opgaven at undervise elever med tegn på talblindhed.

Koordinatorernes vigtigste erfaringer kan opsummeres i fem punkter:

1. Mulighed for succesoplevelser
2. Positive relationer og rammer for matematikken
3. Samarbejde med elevens øvrige lærere
4. Brug af konkrete repræsentationer og gentagelser
5. Afdækning af udfordringer

### 9.2.3 Mulighed for succesoplevelser

Både ekspertgruppen og resultater fra forskning på området påpeger, at elever med talblindhed ofte vil have et forhold til matematik, der er præget af nederlagsoplevelser. Flere af koordinato-  
rerne fremhævede nederlagsfølelser som en hindring for elevernes udvikling, da det påvirkede ele-  
vernes tro på, at de kan blive bedre. Nogle af eleverne lod til at tænke lavere om deres egne fær-  
digheder, end hvad koordinato-  
rerne vurderede egentlig passede. En af koordinato-  
rerne forklarede:



*"Det var ikke nogen af mine elever, der bonede ud i risikozonen, så for dem handlede det nok lige så meget om at have nogle positive øjeblikke med matematik og blive anerkendt i, at de sagtens kan. De opfattede sig selv i større vanskeligheder, end de egentlig var."*

Det blev fremhævet som en motivationsskabende faktor at sætte et niveau for undervisningen, hvor eleven havde mulighed for at besvare nogle opgaver korrekt og derved skabe plads til succesoplevelser med matematikken. Således udtrykker en koordinator det:



*"Havde det været en 'ægte situation', så tror jeg det havde været nemmere at motivere dem og danne en relation, for vi ville aldrig starte lige på og hårdt, vi ville starte blødere ud [ift. RD4-testen]. Man ville give dem noget, de kunne mestre først. Med RoS-kufferten har vi nok skudt lidt højt. Skulle vi gøre det igen, ville vi bruge opgaverne, men vi ville nok tilpasse det mere til den enkelte elev."*

### 9.2.4 Positive relationer og rammer for matematikken

Som beskrevet i afsnit 9.1.3 'Udvikling og afprøvning af indsatser', var alle indsatserne struktureret som særundervisning uden for klassens almindelige rum og skema i små hold eller én-til-én med den/de udvalgte elever. Dette fremhævede koordinato-  
rerne som relationsfremmende, både hvad angik deres egne relationer til eleverne samt elevernes relationer til hinanden. Som beskrevet i forrige afsnit, har mange af de pågældende elever et anstrengt forhold til matematikken. Hertil kan positive relationer og rammer for matematikundervisningen bidrage til at skabe et rum, hvor det er trygt at udfordre sig selv med risiko for at lave fejl. En koordinator udtrykker:



*"Bare det at komme ud og lave noget andet, og at der er et positivt fokus på dem ved at man sidder i mindre grupper, hvor der er masser af plads og tid til at blive hjulpet og blive kigget på undervejs."*

En anden siger:



*"Relationen til børnene, ved at være i de der små grupper, det har rykket enormt meget, ved at tegne tingene og tage et par skridt tilbage. Mit indtryk er, at de ligesom har fået et selvtilids-boost, og det har løftet dem."*

Ved at fokusere på lærer-elev og elev-elev-relationerne – i samspil med at eleven får nogle succesoplevelser – kan elevens relation til matematikken forbedres. For elever med talblindhed eller tegn derpå kan det være en lettelse i sig selv, at elevens udfordringer anerkendes. Én af koordinato-  
rerne, der havde en elev med klare tegn på talblindhed, beskriver:



*"Jeg tror, den største ændring er sket idet, der er nogen, der har sagt: "du er talblind" [...] Det virkede som om, det var en kæmpe lettelse, da der var nogen, der sagde: "det er nok det her, det handler om". Den har han hurtigt grebet, og nu er det dét, han siger nede i klassen, når der er nogen, der spørger. Han har fuldstændig taget det til sig, det er sådan, det er, det er noget, jeg har svært ved, og det vil jeg gerne arbejde med – og I må gerne vide det alle sammen."*

## 9.2.5 Samarbejde med elevens øvrige lærere

Samarbejdet mellem matematikvejlederen og elevens øvrige lærere blev også fremhævet som et særskilt opmærksomhedspunkt. Særligt elevens matematiklærere, men i nogle tilfælde også lærerne i de øvrige fag, afhængigt af omfanget af elevens udfordringer. En elev med talblindhed kan fx være udfordret i naturfag, madkundskab eller idræt, hvor tal ofte indgår i fagenes indhold og opgaver. Også i de kulturelle fag som dansk eller historie indgår tal og mængder, fx i beskeden om, at eleverne skal læse nogle bestemte sider eller kapitler – eller i form af årstal i forbindelse med, at der arbejdes med historiske eller litterære perioder. Tal og mængder er en naturlig del af vores hverdag, og dette gælder ikke mindst fagene i skolen. Det kan derfor være en fordel at etablere et samarbejde med elevens øvrige lærere i forbindelse med afdækning af matematikudfordringernes omfang og/eller i forbindelse med pædagogiske indsatser. En koordinator understreger dette således:



*"En af de vigtigste pointer er, at talblindhed i den grad rækker ud over faget matematik. Det ses ikke kun i matematiktimerne. Det er over hele paletten. Det er også dansklæreren, der skal være opmærksom på, at her sidder et barn med nogle vanskeligheder, der også rækker ind i danskfaget og helt ud i livet."*

## 9.2.6 Brug af konkrete repræsentationer og gentagelser

I projektets løb er det stået frem, både fra eksperternes side, fra forskningen og fra koordinatorernes erfaringer, at der særligt er to didaktiske virkemidler, som er gavnlige i undervisningen af elever med talblindhed. Det første virkemiddel er at bruge konkrete materialer til at illustrere mængder. For elever, der har svært ved tal og mængder, heriblandt elever med talblindhed, kan det være en fordel at inddrage de taktile sanser ved brug af konkrete materialer til at repræsentere mængder. Konkrete materialer knyttes til tallenes symboler og giver eleven en fornemmelse af de mængder, symbolerne repræsenterer. I beskrivelsen af den pædagogiske indsats, understreger en af koordinatorerne vigtigheden heraf:



*"Det er vigtigt, at man har en bred vifte af repræsentationer i sin undervisning, men jeg oplever, at det er meget vigtigt med mange repræsentationer gennem konkrete materialer, når vi arbejder med elever, der har massive udfordringer i matematik." (Bilag 2a)*

Det andet virkemiddel er gentagelser. I forlængelse af det forrige punkt kan et samarbejde med elevens matematiklærer understøtte, at de emner, der arbejdes med, gentages i flere sammenhænge og ofte. En koordinator forklarer:



*"Jeg har en gang om ugen fået skrevet til matematiklæreren, hvad det er, vi har arbejdet med, så de har kunne flette det ind i deres undervisning. Noget af det, jeg tager med, er, at det er rigtig vigtigt, at det, vi laver med eleverne, kommer med i den almene undervisning – fordi det er gentagelse, gentagelse, gentagelse og at bruge det, bruge det, bruge det."*

I gentagelser ligger også at give tid og plads til de emner, der kan virke basale. En af koordinato-  
rerne beskriver:



*"Noget af det, der blev meget tydeligt er, hvor stort fokus, vi må have på talforståelse, hvor vi måske tidligere er gået lidt for hurtigt frem."*

## 9.2.7 Afdækning af udfordringer

Som et sidste punkt fremhæves betydningen af et grundigt afdækningsarbejde som en særskilt, vigtig erfaring. Talblindhed er et meget forskelligartet fænomen, som kan komme til udtryk i forskellige former. Meget tyder på, at der ikke findes generelle og universelle forløb, der kan hjælpe elever med talblindhed, men at der må ligge en afdækning til grund for at vurdere, hvad der vil være meningsfuldt for den enkelte elev. En af koordinato-  
rerne beskriver dette således:



*"Talblindhed er mere komplekst end som så. Jeg arbejder også meget med ordblindhed, og jeg synes dette [talblindhed] er vanskeligere. Det er sværere at sætte en finger på, fordi det er så individuelt fra barn til barn."*

En grundig afdækning er nødvendig for at kunne tilrettelægge den bedst mulige indsats målrettet elevens konkrete udfordringer. Indsatser rettet mod elever med talblindhed uden andre vanskeligheder kan således bygges op ud fra andre forudsætninger, hvor der kan trækkes på elevens styrkesider. En koordinator formulerede det således:



*"Jeg synes, der er en meget stor forskel på at have lavet en indsats med et barn, som har generelle vanskeligheder og så de her par børn, jeg har siddet overfor i dette projekt, som har så specifikke vanskeligheder. Med disse børn har jeg lettere kunne trække på deres styrkesider, fordi dem er der så mange af. Det er sværere, når jeg sidder over for et barn, der har generelle vanskeligheder."*

## 9.3 Kerneskolernes afprøvning af talblindhedstesten

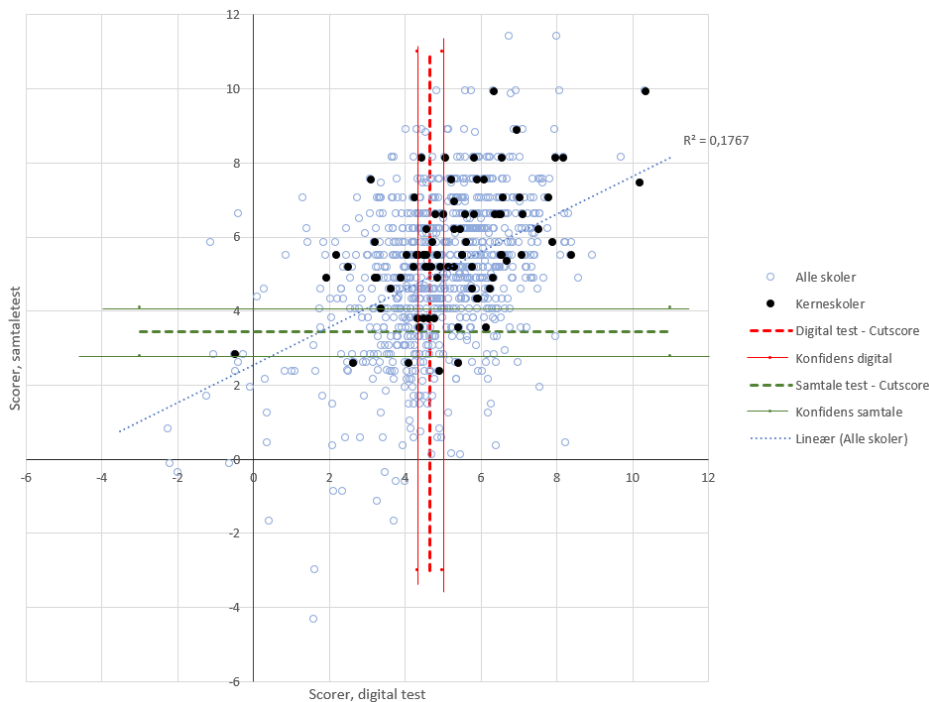
Som nævnt har kerneskolerne med deres deltagelse indvilget i at bidrage til afprøvningen af alle fire dele af testbatteriet, dvs. observationsguiden (enten it-udgaven eller tjeklisten), den digitale test, samtaletesten og PPR. Resultaterne af denne afprøvning vil blive udfoldet i de følgende afsnit.

### 9.3.1 De første tre trin i tragtmodellen

Afprøvningen af observationsguiden og resultaterne herfra er beskrevet i afsnit 6.3 og 8.1 – hertil har kerneskolerne bidraget. Yderligere har kerneskolerne som beskrevet tidligere bidraget til afprøvning af både den digitale test og samtaletesten.

Af Figur 46 fremgår alle elever, der har taget både en digital test og en samtaletest (de blå prikker) og de elever der kommer fra kerneskolerne (de sorte prikker).

**Figur 46: Kerneskolernes resultater i den digitale test og samtaletesten (N=1.659)**



Det udestår herefter at se på de af kerneskolernes elever, der ligger i risikozonen for talblindhed, og det udestående spørgsmål er, om disse elever falder ind under projektets målgruppe (dvs. elever med "ren talblindhed" – uden andre personlige, social, sproglige, kognitive mv. problemer) eller om en flerhed af problemer (komorbiditet) kan være den egentlige årsag til elevernes lave scorer i testene.

### 9.3.2 PPR i tragtmodellen

Set fra et overordnet perspektiv var PPR's rolle defineret i den oprindelige testmodel: Når skolerne ved hjælp af den digitale test og samtaletesten havde lokaliseret de lavest præsterende 3 pct. af eleverne, skulle PPR identificere den halvdel, der forventedes at have talblindhed. Dette viste sig imidlertid at være forbundet med komplikationer i praksis, fordi der ikke eksisterer konsensus om, hvad begrebet talblindhed dækker over og at diagnosticering af talblindhed ligger udenfor, hvad PPR forudsættes at have ekspertviden om (fx matematikvanskeligheder og effekter af utilstrækkelig undervisning).

### 9.3.3 Møder med PPR i de 4 kommuner

Der har været holdt et fælles møde med de fire kerneskolokommuner – herunder PPR-medarbejdere fra alle kommunerne. Formålet var at informere om projektet og fremdriften samt at initiere en dialog med PPR om afklaring af PPR's rolle i projektet.<sup>35</sup>

<sup>35</sup> I mødet deltog også Steen Polk (Pædagogisk Psykologisk konsulent hos Svendborg Kommune) som bistår Epinion og DPU i denne del af opgaven som en del af ekspertgruppen og som i sidste fase af udviklingsprojektet var rundt og besøge de skoler, hvor samtaletesten havde udpeget 10 elever som talblinde. Steen Polk havde testet eleverne med redskaber, som mange skoler har til rådighed mhp. at afklare kognitiv kapacitet og havde haft samtaler med eleverne selv og deres lærere. På baggrund af disse vurderinger af de 10 børn, kunne 5-7 kategoriseres som reelt talblinde med en kognitiv begavelse i normalområdet, mens de øvrige af de 10 kunne anses for at

Konkret initierede vi en proces, som involverede PPR i kerneskoledommunerne. PPR i hver af disse kommuner blev inviteret til et dialogmøde i starten af og midt i processen med henblik på erfaringsudveksling. Formålet var at identificere, hvilke ideer, viden og behov, de fire PPR har i forhold til at kunne vurdere eleverne, samt hvordan samarbejdet med skolerne bedst kunne organiseres.

Møderne tog afsæt i spørgsmålene:

- Hvad er talblindhed, og hvad ved vi/PPR om årsagerne hertil?
- Hvilke testmaterialer kan PPR med fordel bruge, og hvordan kan de anvendes bedst muligt?
- Hvad er elever med talblindheds primære problemer, og hvad kan PPR/skolen gøre for at hjælpe dem?

Fra alle fire kommuners PPR blev der signaleret nysgerrighed, men også uafklarethed i forhold til hvad PPR's rolle kan være i processen, og der blev efterlyst retningslinjer og klarere beskrivelser af den funktion, som PPR forventedes at udfylde i projektet. Disse møder og drøftelserne bragte os således ikke væsentligt nærmere på, hvordan den efterspurgte afklaring kunne tilvejebringes.

På baggrund af møderne besluttede vi at inddrage projektgruppens ekspert Steen Polk med henblik på at få afprøvet en model for samarbejde mellem skolerne og PPR. Der blev taget afsæt i en metode, som Steen Polk tidligere selv har analyseret på, og modellen kan beskrives således:

Hvis den digitale test og samtaletesten begge peger i retning af, at eleven har tegn på talblindhed, men hvor skolen er i tvivl om, hvorvidt der er andre årsager, så kan der foretages følgende undersøgelser fx vha. generelle kognitive redskaber, som mange skoler råder over<sup>36</sup>: Undersøgelse om 1) eleven har normal intelligens (kan undersøges med Ravens Progressive Matricer, sets A, Ab, B, eller CHIPS), 2) om hukommelsen er intakt, og 3) om eleven i øvrigt har godt fagligt overblik og kan løse andre opgaver på aldersvarende niveau (fx læsning). Det skal understreges, at disse undersøgelser i givet fald vil supplere talblindhedstestene og ikke være en del heraf.

Hvis skolen herefter fortsat er i tvivl, om der fx kan være tale om generelle indlæringsvanskeligheder, eller skolen er i tvivl om, hvad resultaterne viser, vil det være nærliggende, at skolen tager PPR med på råd. PPR vil her kunne supplere med andet testmateriale hvis nødvendigt.

To af kerneskoledommunerne tog imod et tilbud om at deltage i en demonstration af, hvordan generelle kognitive redskaber kunne anvendes ift. de deltagende elever. Den samlede vurdering var, at denne undersøgelse ikke gav anledning til mistanke om generelle indlæringsvanskeligheder eller andre massive vanskeligheder i specialundervisningsregi, som skolen ikke vil kunne magte alene. I dette tilfælde sås derfor ikke anledning til yderligere involvering af PPR.

---

have generelle vanskeligheder. Det var bl.a. med afsæt i disse erfaringer, at det vurderedes som hensigtsmæssigt i udviklingsprojektet at involvere PPR som sidste led i tragmodellen.

<sup>36</sup> Materiale til undersøgelse af 1) Pædagogisk profil, Ravens Progressive Matricer, sets A, Ab, B, 2) Visuel/motorisk kodeprøve, 3) Kompleks figurtegning og 4) Auditiv hukommelsesspændvidde, serielt og reversibelt. Materiale til de foreslåede undersøgelser kan hentes på hjemmesiden <https://talblind.com/>.

## 10. Spørgeskemaundersøgelse til lærere

Som opfølgning på skolernes deltagelse er alle koordinatore fra de 799 tilmeldte skoler blevet inviteret til en spørgeskemaundersøgelse. Formålet har været at indsamle feedback på processen, testredskaber mv. fra de involverede pædagogiske medarbejdere.

Undersøgelsen er gennemført som et webbaseret spørgeskema blandt koordinatore på de 799 skoler tilmeldt projektet. Koordinatorerne er blevet bedt om at videresende linket til spørgeskemaet til øvrige lærere og vejledere på skolen, der har været engageret i projektets aktiviteter. Foruden faktuelle baggrundsspørgsmål har de svaret på en række vurderingsspørgsmål med afsæt i egne erfaringer med talblindhedsprojektet og dets elementer.

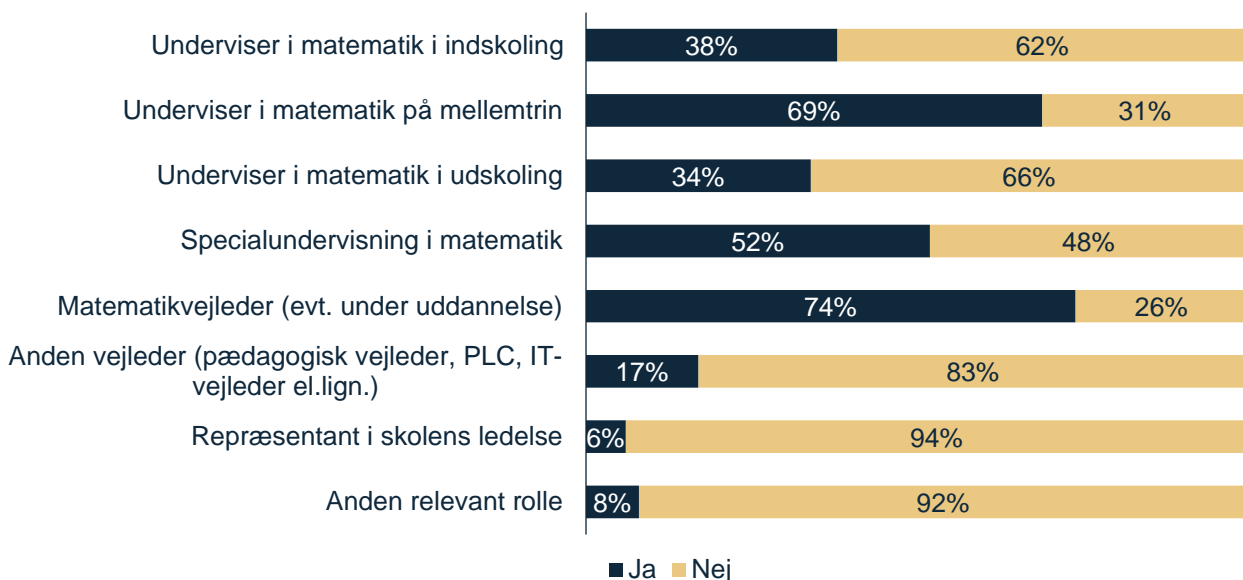
Data er indsamlet i perioden 6. juni til 6. september 2021, og der er gennemført 376 besvarelser fordelt på 350 skoler. Kun fra 20 skoler er der modtaget mere end ét svar. Opgjort på elevniveau er 64 pct. af de 3.241 deltagende elever repræsenteret via en lærers besvarelse af skemaet. De 64 pct. er således et vægtet tal, hvor lærerbesvarelsen vejes med antallet af elever, der indgår i analysen af de digitale testresultater.

### 10.1 Lærernes baggrund og rolle på skolen

I undersøgelsen af lærernes baggrund er det overordnede billede, at det er erfarne lærere, der har deltaget i afprøvningen af talblindhedstesten. Således har godt og vel tre ud af fire lærere mere end 10 års undervisningserfaring i matematik.

Figur 47 viser lærernes rolle i relation til matematikundervisningen, herunder at lidt over halvdelen udfører specialundervisning i matematik, og at knap tre ud af fire er matematikvejledere på skolen.

Figur 47: Hvad er dine nuværende roller ift. matematikundervisning? N=376



Langt de fleste lærere, der har besvaret undersøgelsen, er således matematiklærere. Ca. halvdelen udfører specialundervisning i matematik og langt de fleste har ydermere en vejlederfunktion.

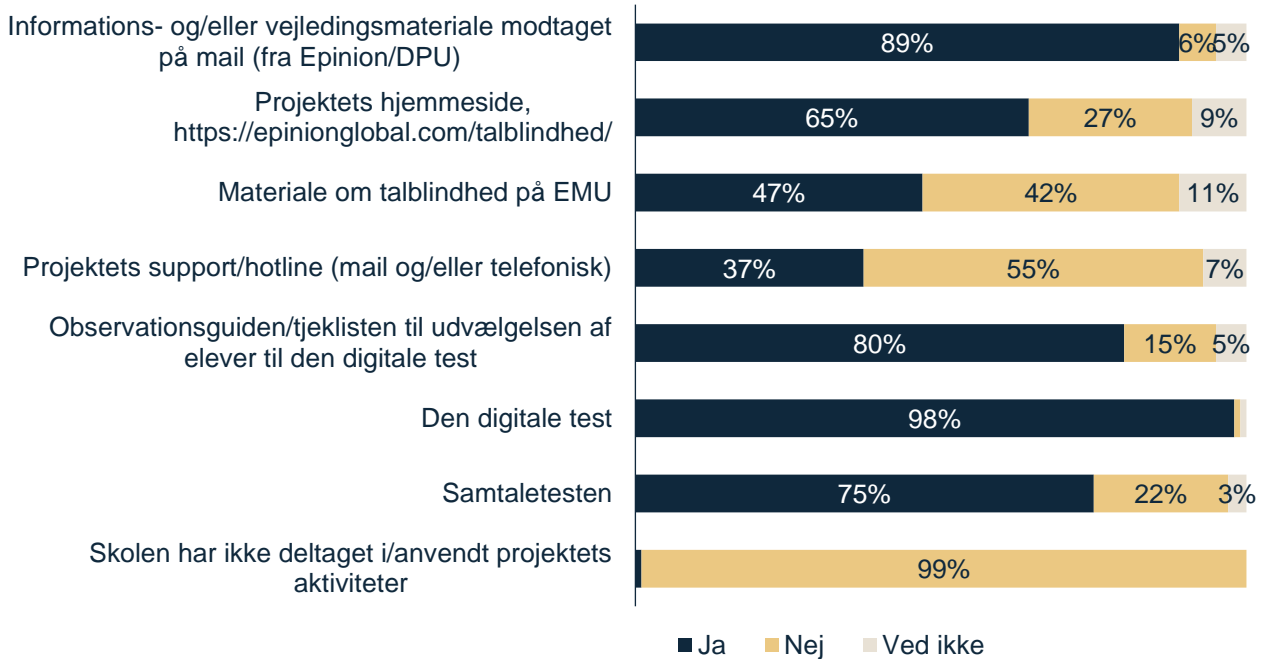
### 10.2 Skolens deltagelse i projektets aktiviteter

I Figur 48 nedenfor har lærerne angivet, hvilke aktiviteter skolen har deltaget i ifm. projektet. 98 pct. har deltaget i den digitale test, som har været minimumskravet til skolen for at kunne deltage i



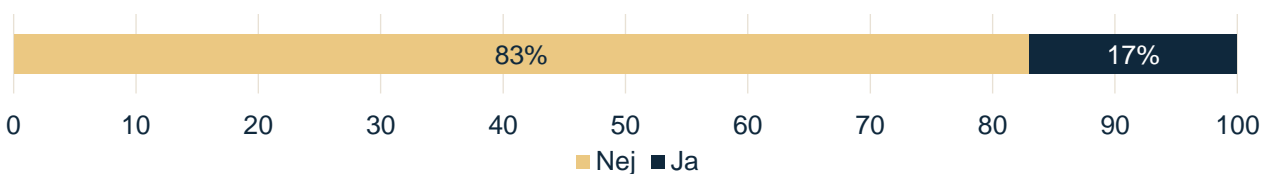
projektet. Derudover har fire ud af fem benyttet observationsguiden/tjeklisten til indledningsvis at udvælge relevante elever til afprøvning af den digitale test. 75 pct. har anvendt samtaletesten, og hertil skal bemærkes, at der er nogle skoler, som ikke har haft elever for hvem det er blevet anbefalet at gennemføre en samtaletest.

**Figur 48: Hvilke af projektets aktiviteter har skolen deltaget i/anvendt? Spørgeskemaundersøgelse blandt lærere**



Endvidere er lærerne blevet spurgt til pædagogiske indsatser mellem de to afprøvningsrunder af den digitale test. Figur 49 viser, hvor stor en andel af de lærere, der har deltaget i undersøgelsen, som har taget særlige initiativer/indsatser.

**Figur 49: Andel af lærere, der har taget særlige initiativer (N=299)**



Således svarer 17 pct., at de har taget særlige initiativer med hensyn til matematik for de pågældende elever. De 17 pct. omfatter også kerneskolere, men undersøgelsen viser, at en væsentlig større del har iværksat indsatser med afsæt i testens resultater. Dette vil blive uddybet i afsnit 11.1. Indsatserne er af varierende karakter, fx beskriver en lærer således:



*”Vi har forsøgt at hjælpe de elever, der har det svært med matematik med en to-lærertime om ugen” – (Lærersurvey)*

Andre indsatser var i højere grad fokuseret på et samarbejde med hjemmet om, at eleven har særligt behov for støtte:



*”Størstedelen af tiden har der været hjemmeundervisning (Covid-19), men der har været ekstra kontakt til disse elevers hjem for at sikre, at de kunne følge med, da der fra skolens side var opmærksomhed omkring, at det var elever der havde behov for støtte” – (Lærersurvey)*

Der er også blevet iværksat særlige indsatser, én-til-én mellem lærer og elev, hvor eleven har fået særlig undervisning:



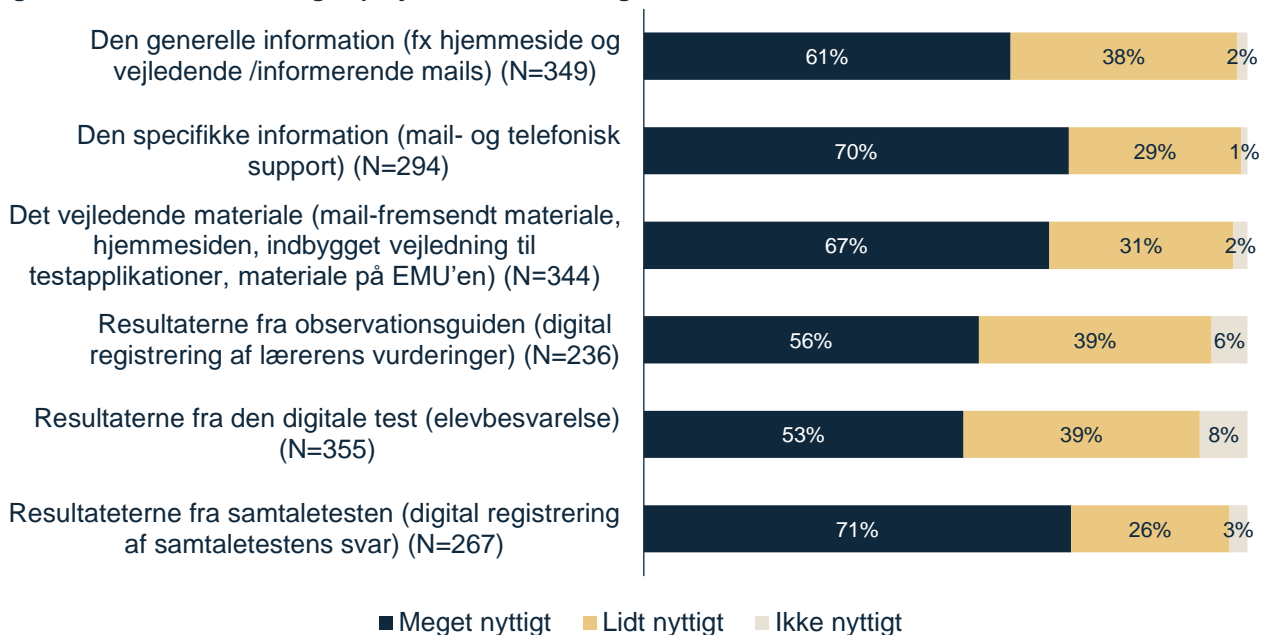
*"Jeg har givet enkelte lektioner til to elever, der efter samtalen er "i risiko for talblindhed" en-til-en undervisning i forhold til at udvikle deres talforståelse" – (Lærersurvey)*

Overstående citater illustrerer et udsnit af den brede variation af indsatser og initiativer, skolerne har iværksat på baggrund af testenes resultater.

### 10.3 Vurdering af projektets aktiviteter og testresultater

Nedenfor, i Figur 50, har lærerne vurderet anvendeligheden af projektaktiviteterne og testresultater fra de forskellige deltest. Her er det overordnede billede, at alle projektaktiviteter og testresultater findes meget nyttige af størstedelen af lærerne. Det er dog værd at bemærke, at 39 pct. af lærerne finder resultaterne fra hhv. observationsguiden og den digitale test for "lidt nyttigt". Det indikerer, at der kan være behov for at udfolde testresultaterne og deres betydning yderligere fremadrettet for at øge anvendeligheden.

**Figur 50: Lærernes vurdering af projektets aktiviteter og information**

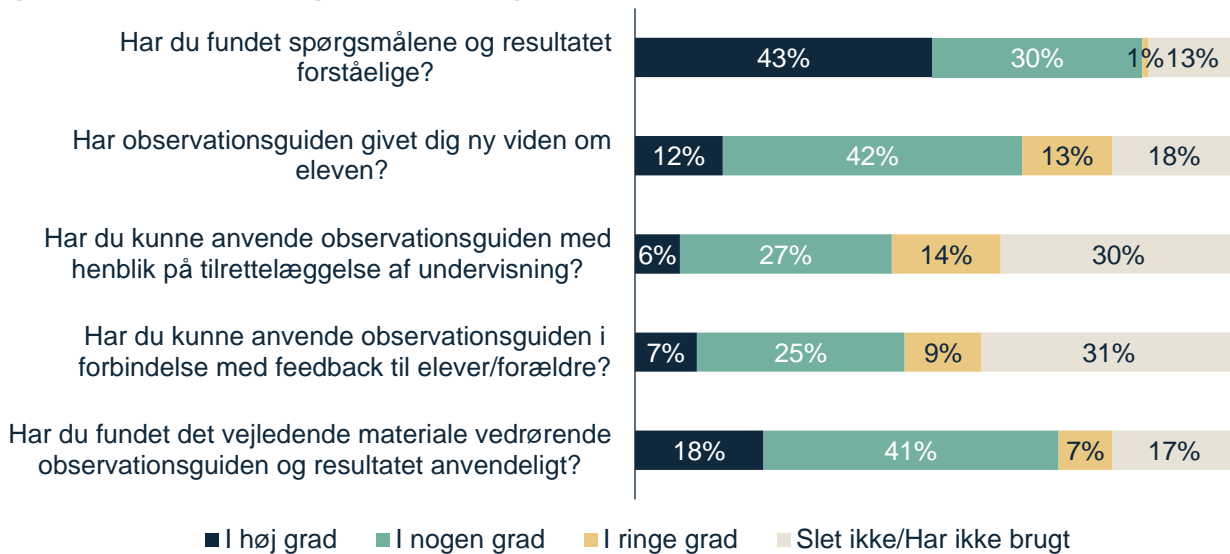


#### 10.3.1 Lærernes erfaringer med observationsguiden

Af Figur 48 fremgår det, at 80 pct. af lærerne angiver at have brugt observationsguiden i større eller mindre omfang. Dog tyder lærernes tilbagemeldinger på, at observationsguiden i mange tilfælde er brugt på en anden måde, end hvad der var hensigten, og flere lærere udtrykker, at udbyttet af observationsguiden ikke står mål med indsatsen.

I Figur 51 har lærerne vurderet specifikke forhold om resultaterne fra observationsguiden.

Figur 51: Lærernes vurdering af observationsguiden



Undersøgelsen tyder på, at lærerne i det store hele har fundet observationsguiden og dens spørgsmål forståelige og anvendelige. Observationsguiden har i nogen grad givet ny viden om eleven og i mindre grad givet input til undervisningen og informationer til dialog med elever og forældre. Flere lærere udtrykker, at udbyttet af observationsguiden ikke står mål med indsatsen. En lærer udtrykker, at observationsguiden kan "[...] virke lidt overvældende, selvom det alle er vigtige opmærksomheder - det er rigtig meget at forholde sig til." (Lærersurvey)

En anden lærer tilslutter sig, at observationsguiden var for omfangsrig, men fandt til gengæld observationsguiden brugbar til formål som ikke var forudset, nemlig til brug ved dialog med kollegaer:



*"Der var for mange ord. Det var ikke tydeligt om modtager var egen lærer, der kender barnet, eller en vejleder, som jeg er (kender ikke børnene). Men jeg kunne selvfølgelig bruge guiden til dialogen med læreren." – (Lærersurvey)*

Flere lærere betragter således observationsguiden som et godt grundlag for dialog og indføring i, hvad det er for nogle opmærksomhedspunkter i forbindelse med talblindhed, der er særligt vigtige. En lærer udtrykker det således:



*"[...] jeg vil absolut bruge observationsguiden til (som vejleder) at guide mine kolleger til tidligt at have et fokus på evt. talblindhed hos eleverne - selv i de mindre klasser, det der med, hvad det er, vi skal holde øje med." – (Lærersurvey)*

Derimod var der forbehold i forhold til anvendelse af observationsguiden til det oprindeligt påtænkte formål, nemlig at foretage den indledende identifikation af elever i risiko for talblindhed. En lærer formulerer det på denne måde:



*"Da man også kunne udvælge eleverne ud fra dagligdagen og andre test og prøver, brugte vi det, vi kender, og ikke særligt observationsguiden (den var i baghovedet)" – (Lærersurvey)*

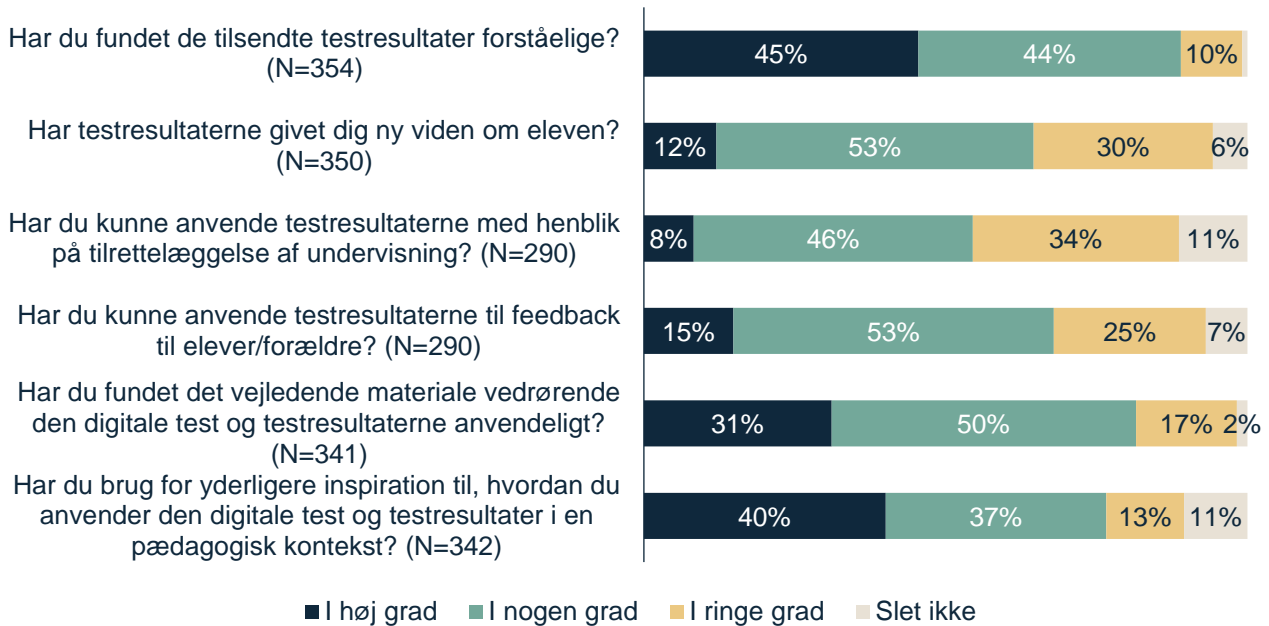
Med udgangspunkt i meldingerne fra spørgeskemaundersøgelsen, kan det overvejes at anvende observationsguiden til et bredere formål end oprindeligt påtænkt.

### 10.3.2 Lærernes erfaringer med den digitale test

Den digitale tests funktion er at skabe et generelt overblik over elevernes grundlæggende forståelse for tal og mængder ved at kombinere rigtighed med hastighed i besvarelsen. Opgaverne relaterer sig ikke direkte til undervisningens indhold, men snarere til nogle abstrakte konstrukter, der erfaringsmæssigt er koblet til tal- og mængdeforståelse. Testen er opbygget af meget lette opgaver, hvor det for langt de fleste elever bliver svarhastigheden, der er udslagsgivende for testresultatet.

I Figur 52 har lærerne vurderet specifikke forhold om testresultaterne fra den digitale test.

**Figur 52: Lærernes vurdering af resultaterne fra den digitale test**



I figuren ses en overordnet positiv holdning fra lærerne til den digitale test. Mere end halvdelen af lærerne finder i høj eller nogen grad testresultaterne forståelige og nyttige – og enkle at anvende. En lærer udtrykker, at de finder testen egnet som screeningsværktøj, fordi den er let og hurtig at anvende på en større gruppe elever. Yderligere fremhæves også elevernes oplevelse:



*"Det er meget nemt at gå til som vejleder. Det er udarbejdet i nogle formater, så børnene finder det motiverende og ikke for overvældende i mængde." – (Lærersurvey)*

Opgavernes lethed har den virkning, at alle elever – også de svageste – forventes at besvare langt størstedelen korrekt, og svarhastigheden får en afgørende rolle for testens resultat (scoren). Dette kan bidrage til at give eleverne en positiv oplevelse med at gennemføre testene. En lærer siger således om den digitale test:



*"Dejligt overskueligt materiale. Den digitale test er nem for børnene, og det er rart at kunne teste dem på en måde, så de ikke oplever nederlag" – (Lærersurvey)*

På trods af testresultaternes abstrakte karakter viser undersøgelsen endvidere, at over halvdelen af lærerne har kunnet anvende resultaterne i forbindelse med undervisningstilrettelæggelse. Undersøgelsen peger også på, at lærerne oplever et behov for yderligere inspiration til, hvordan de anvender testen og testresultaterne i en pædagogisk kontekst. En lærer beskriver:



”Jeg ønsker mere viden om, hvad det betyder, at en elev eksempelvis ikke kan se, om der er 4 prikker eller ej. Hvad er det for viden, jeg får om eleven der, og hvad gør jeg ved det.” – (Lærersurvey)

Dette eksempel kalder på yderligere viden om, hvordan man i undervisningen kan adressere de problemer, der konstateres ved de fem skalaer i testen. Det er vigtigt at pointere, at opgaverne i den digitale test i højere grad bør betragtes som indikatorer på talblindhed end som emner, der bør undervises i separat. Hvilke konkrete emner, der kan gøres til genstand for undervisning med henblik på at afhjælpe elevens vanskeligheder, kan samtaletesten i højere grad give et svar på.

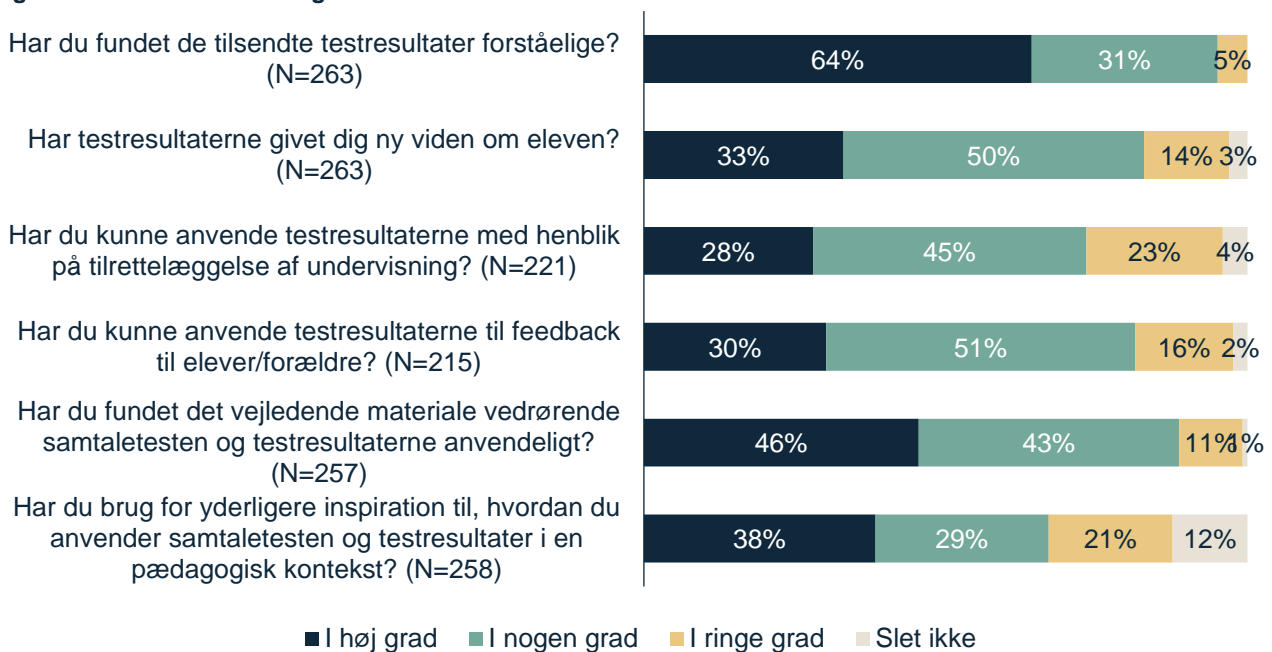
Undersøgelsen viser altså, at den digitale test fremstår som et færdigudviklet produkt, der er lettilgængeligt og enkelt at afvikle for både elever og lærere. Dog udestår en opgave med at vejlede om testresultaternes fortolkning og anvendelse i en pædagogisk kontekst.

### 10.3.3 Lærernes erfaringer med samtaletesten

I modsætning til den digitale test har samtaletesten i høj grad relation til den daglige undervisning og de opgaver, eleverne præsenteres for i den forbindelse. Opgavernes indhold spænder fra basal talforståelse og aritmetiske færdigheder til praktisk anvendelse af den mentale tallinje.

I Figur 53 har lærerne vurderet specifikke forhold om testresultaterne fra samtaletesten.

Figur 53: Lærernes vurdering af resultaterne fra samtaletesten



Samtaletestens relation til undervisningens indhold afspejler sig i lærernes besvarelser. Her angiver 3 ud af 4 lærere, at de har kunnet anvende testens resultater med henblik på undervisningstilrettelæggelse. En lærer beskriver, at:



”Der har været noget matematikstøtte for de fleste deltagende elever i talblindhedsprojektet. Indsatserne var primært tilrettelagt på baggrund af samtaletesten.” – (Lærersurvey)

Foruden samtaletestens opgaver bidrager testens dialogbaserede afviklingsform og umiddelbare tilbagemelding af resultater yderligere til, at læreren får en dybtgående indsigt i elevens udfordringer og svarprocesser.



*”Samtaletesten er bare et helt fantastisk værktøj til at opnå en dybere forståelse og indsigt i karakteren af elevens udfordringer. Selvom samtaletesten i sig selv giver virkelig meget nyttig information, så er det også rart, at den automatisk kommer med en konklusion på, om der er grund til yderligere bekymring for talblindhed, eller udfordringerne ikke skal betragtes i det lys.”*  
– (Lærersurvey)

Yderligere angiver langt de fleste (81 pct.), at samtaletestens resultater har kunnet anvendes som grundlag for feedback og dialog med forældre og elever. To ud af tre lærere mener dog stadig, at de har behov for yderligere inspiration til, hvordan samtaletesten og resultaterne kan anvendes i en pædagogisk kontekst.

Lærerne giver i det hele taget udtryk for, at der er behov for yderligere uddannelse i testafholdelsen samt fortolkning og feedback af testresultater. Nogle efterlyser egentlig efteruddannelse:



*”Det kræver meget mere uddannelse af lærere til at tilbagemelde i præcis denne test”*  
– (Lærersurvey)

Mens andre efterspørger kortere, operationelle forløb:



*”Hvor kunne det være godt, hvis I kunne holde en mini-workshop, hvor man bliver ”uddannet” i at tage disse tests”* – (Lærersurvey)

Alt i alt tyder undersøgelsen på, at samtaletesten anses som direkte anvendelig i forbindelse med tilrettelæggelse af lærerens indsats. Dog er der den udfordring, at samtaletesten kun kan anvendes meningsfuldt én gang pr. elev, da eleverne øjensynligt bliver trænet i at besvare de specifikke opgaver givet testens format, jf. afsnit 8.3.3.

### 10.3.4 Lærernes vurdering af talblindhedstesten

Et centralt spørgsmål er, om lærerne og skolerne ønsker at anvende et sådant testredskab, hvis det stilles til rådighed. Set i lyset af erfaringerne fra projektet er der en overvældende interesse herfor. 65 pct. svarer i meget høj grad og kun 1 pct. svarer, at de slet ikke vil benytte en talblindhedstest. Det samlede billede af lærernes vurdering af afprøvningen er således, at de har haft god nytte af de forskellige deltest og samtidig udtrykker en markant efterspørgsel efter en talblindhedstest i fremtiden, herunder også inspiration til, hvordan testen kan anvendes i en pædagogisk kontekst.

Figur 54: Lærernes vurdering af, om skolen vil benytte en talblindhedstest, hvis den stilles til rådighed (N=376)



Generelt giver lærerne udtryk for, at der er et presserende behov for at få adgang til en talblindhedstest som denne i folkeskolen. En matematikvejleder skriver således:



*”Testen skal gerne ud hurtigst muligt, så der faktisk er en reel talblindtest til rådighed for skolen. Flere elever, der har udfordringer med matematikken, kunne underviseren få glæde af at få testet af.” – (Lærersurvey)*

Flere lærere udtrykker endvidere, at de har fulgt udviklingsarbejdet om en talblindhedstest i mange år:



*”Samlet set er det et længe ventet og virkelig ønsket værktøj til at hjælpe os med at identificere mulige talblinde elever tidligere. Jeg håber meget, at testen bliver tilgængelig fremadrettet.” – (Lærersurvey)*

En lærer understreger, at en talblindhedstest som denne kan give grundlag for et bedre samarbejde med forældrene:



*”Vi vil meget gerne, at det udmunder i en egentlig test, for vi har stærkt brug for den. Det har betydet meget for min opfattelse af de elever, som kæmper hårdt med matematikken, og det har især også betydet meget for de forældre, som har været med. Pludselig gav det mening for dem, hvorfor det var så svært at hjælpe deres datter med matematikken. Vi håber, at vi får mere at vide omkring resultaterne, og at der evt. kommer en vejledning i forhold til, hvordan vi støtter de talblinde elever bedst muligt.” – (Lærersurvey)*

Der har også været enkeltstående kritiske kommentarer af det foreliggende testbatteri:



*”I den nuværende form er testen uklar, mangelfuld og ikke tilstrækkelig informativ. Talblindhedstesten skal forbedres væsentligt i forhold til forældre/elev på den ene side og lærer/skole/ledelse på den anden side.” – (Lærersurvey)*

Lærerens kritik er således ikke møntet på et spørgsmål om behovet for en talblindhedstest, men ambitioner om en forbedret udgave samt yderligere information og vejledning.

Langt de fleste tilbagemeldinger viser dog en klar positiv holdning til testen – også i den nuværende udformning:



*”Jeg synes testen har gjort, at man har fået vished i forhold til, om elever var talblinde. Det er et spørgsmål vi til tider også får fra forældre eller som vejleder fra kolleger. Her er det et godt værktøj, da svaret bliver givet ud fra konkrete og saglige begrundelser.” – (Lærersurvey)*

Samlet set viser spørgeskemaundersøgelsen, at der har været både gode og dårlige erfaringer med talblindhedstestens enkelte dele, men langt de fleste har været positive, og der er et udtalt behov for at få stillet en talblindhedstest til rådighed for skolerne.

## 11. Effektvurdering

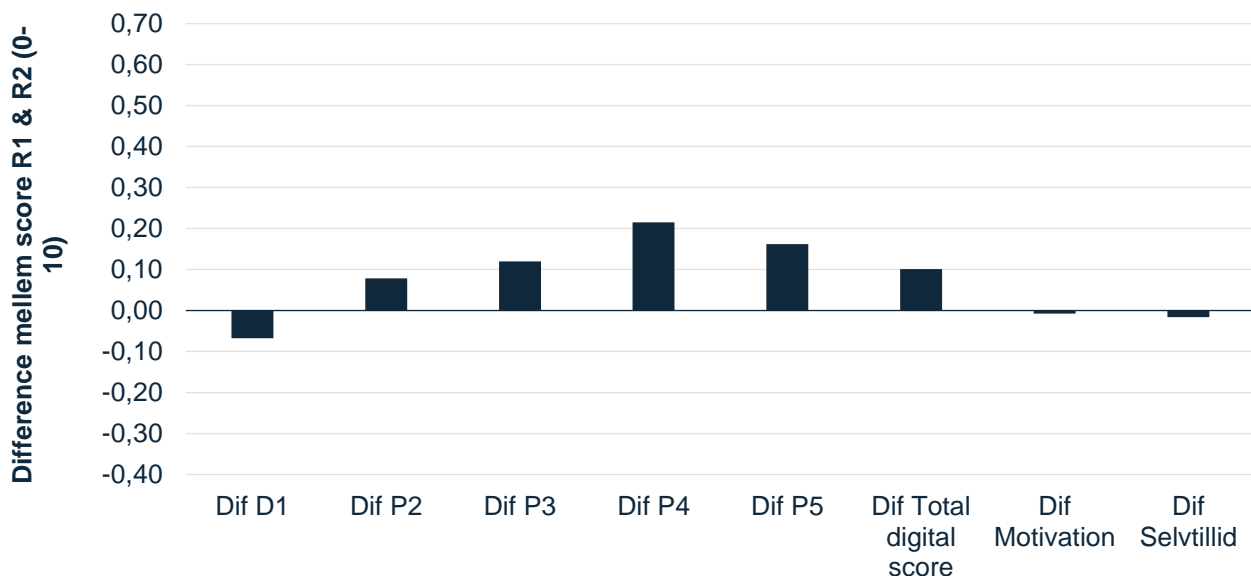
Effektvurderingen i denne undersøgelse består af en opgørelse af differencen mellem resultater fra hhv. runde 1 og 2, opgjort på samme skala. Målingen er således baseret på de elever, som har gennemført den digitale test (dvs. fulde besvarelser af alle fem delskalaer i den digitale test) samt motivation og selvtillidsspørgsmålene i fuldt omfang i både første og anden testrunde, samt hvor elevens lærer har besvaret spørgeskemaet. De fem delskalaer i den digitale test er:

- D1 (Delskala 1, scoret dikotomt): Sammenligning af mængder på tid (ANS)
- P2 (Delskala 2, scoret polytomt): Sammenligning af talsymboler
- P3 (Delskala 3, scoret polytomt): Identifikation af små mængder (Subitizing)
- P4 (Delskala 4, scoret polytomt): Sammenligning af tal og mængder
- P5 (Delskala 5, scoret polytomt): Addition af tal/mængder med forskellige repræsentationer

Af Figur 5 (s. 21) fremgår, hvordan items i de enkelte delskalaer er udformet. Resultaterne på hver skala er opgjort på en skala fra 0-10 med et gennemsnit på 5 (jf. Tabel 14 og teksten hertil på s. 45). Det samlede testresultat er opgjort som et gennemsnit af disse 5 delskalaresultater.

Der er i alt 919 elever, som har afleveret komplette besvarelser i begge testrunder. Det vil sige, at godt en fjerdedel af de elever, som har deltaget i projektet, inddrages i effektvurderingen.

Figur 55: Generel forskel i scorer mellem de to testrunder (N=919 elever)



I Figur 55 vises forskellen i gennemsnittet af elevscorer mellem runde 1 og 2 af den digitale test samt målingen af motivation og selvtillid. Fx betyder "Dif D1" differencen mellem resultatet på skala D1 i runde 1 og runde 2 af den digitale test. Når søjlen over "Dif Total digital score" er lige omkring 0,1 høj, betyder det, at de 919 elever gennemsnitligt scorede 0,1 mere i runde 2 end i runde 1.

Figuren afspejler, at der er forskel på, hvordan udviklingen er på de forskellige delskalaer mellem første og anden måling. Der hvor søjlerne viser positive værdier (hvor søjlen er over akse), har den gennemsnitlige score (der måles på en skala med et gennemsnit på 5) for anden måling været højere end den første. Det vil sige at der er målt en positiv udvikling i perioden – en progression. Omvendt hvis søjlen ligger under akse. Det ses fx, at udviklingen har været svagt negativ (ca. -0,08) for skala 1, der omhandler subitisme, hvorimod der ses en positiv udvikling for de øvrige



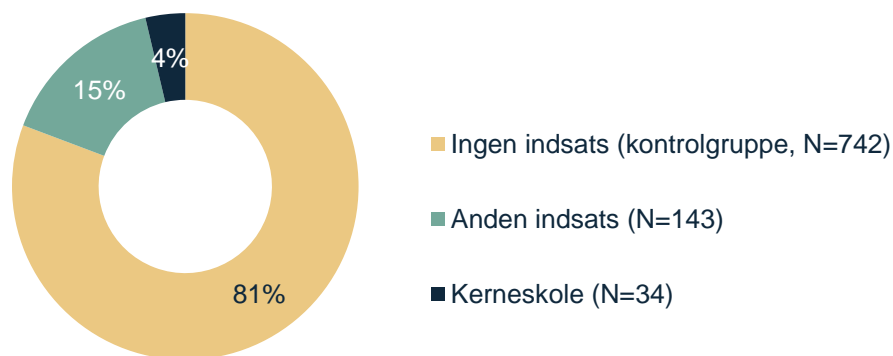
delskalaer. Den stærkeste positive udvikling spores på skala 4 (tilvækst på over 0,2), hvor talsymboler sammenlignes med antal prikker (jf. Figur 5, s. 21). For så vidt angår motivation og selvtillid spores stort set ingen ændringer.

## 11.1 Undersøgelles- og kontrolgrupper

Den oprindelige plan for effektmålingen tog afsæt i, at der arbejdes med to forskellige grupper af elever, der alle har gennemført begge testrunder: en indsatsgruppe (fra kerneskolerne, hvor der udføres en planlagt pædagogisk indsats mellem runde 1 og 2) og en kontrolgruppe (fra de øvrige deltagende skoler).

En forudsætning for, at de øvrige skoler kan fungere som kontrolgruppe for indsatsgruppen, er, at der ikke er udført særlige pædagogiske indsatser med de testede elever mellem de to testrunder. Som beskrevet i afsnit 10.2 er der dog en væsentlig del af de deltagende skoler, der i spørgeskemaundersøgelsen angav, at der er gennemført indsatser med elever fra projektet. På baggrund af spørgeskemaundersøgelsen har vi således identificeret en tredje elevgruppe. Som illustreret i Figur 56 stammer 15 pct. af datamaterialet fra ikke-kerne-skoleelever, som har gennemført den digitale test i både runde 1 og 2, og hvor læreren angiver, at der er gennemført en form for indsats imellem de to testrunder.

Figur 56: Oversigt over indsats- og kontrolgrupper (N=919 elever)



Der er således tre elevgrupper som grundlag for vores effektivvurdering:

- Indsatsgruppe 1: Kerneskooleelever
- Indsatsgruppe 2: Elever fra øvrige skoler, hvor det er angivet, at der er gennemført en form for indsats mellem testrunde 1 og 2
- Kontrolgruppen: Elever fra øvrige skoler, hvor det ikke er angivet, at der er gennemført særlige indsatser i perioden mellem de to test

Godt halvdelen af de elever, der har deltaget i både runde 1 og 2, har besvaret den digitale test i begge runder på en sådan måde, at de indgår i effektivvurderingen. Blandt kerneskolernes elever indgår 34, hvilket udgør lidt under halvdelen.

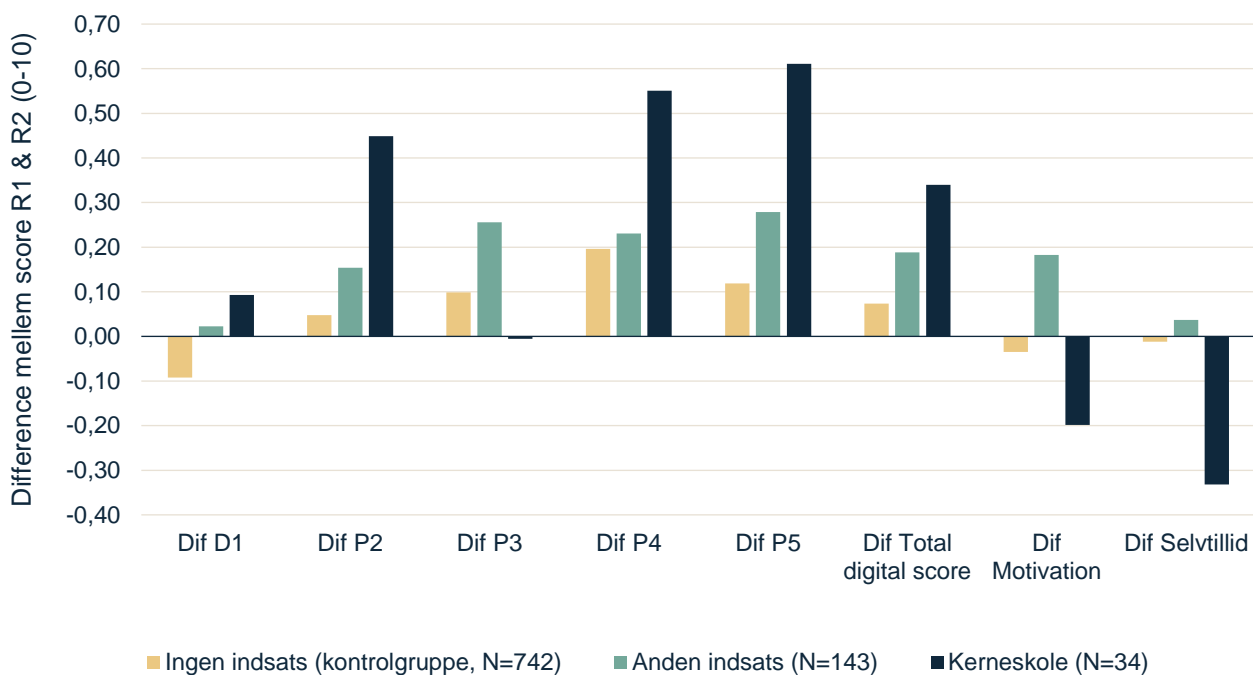
## 11.2 Effektivvurdering med den digitale test

Effektivvurderingen er planlagt til at foregå ved en sammenligning af kerneskooleeleverne med kontrolgruppen. Man skal dog være opmærksom på, at ud af de 84 kerneskooleelever, der deltog i både den digitale test og samtaletesten jf. Tabel 19, var det færre end en tredjedel (24 elever), hvor skolerne fandt det meningsfuldt at gennemføre en pædagogisk indsats mod talblindhed.

Under halvdelen af de 84 elever (34 elever jf. Figur 56) lever op til kriterierne for at deltage i effektvurderingen. Der er således et begrænset datagrundlag til rådighed for vurderingen af effekt, og yderligere ved vi ikke hvilke og hvor mange af de 34 elever, som indgår i effektvurderingen, der har modtaget indsatserne. Grundlaget for effektvurderingen er således af flere årsager spinkelt.

I Figur 57 ses differencen mellem resultaterne fra den digitale test i runde 1 og runde 2 fordelt på de tre grupper. Med det spinkle datagrundlag in mente fremgår det af figuren som om, der er en tydelig effekt – dvs. en forskel mellem kontrolgruppe og kerneskolerelever – på visse af de fem delskalaer. Særligt ved skala P2, P4 og P5 synes forskellen markant. På skala P5, hvor forskellen er størst, findes analysens stærkeste signifikans ( $\alpha < 0,1$ ). I praksis betyder det, at man med 90 pct. sikkerhed<sup>37</sup> kan sige, at de 34 elever på kerneskolerne har haft et positivt læringsudbytte mellem de to testrunder målt ved det, som den digitale test afprøver i deltest 5, og at dette læringsudbytte er betydeligt større end de øvrige skolers elever.

Figur 57: Effektvurdering fordelt på indsats- og kontrolgrupper



Som det fremgår af figuren, er der tydelige forskelle på gennemsnittet af kontrolgruppen og kerneskolereleverne, men det er antageligt på grund af datagrundlagets begrænsede størrelse, at der ikke kan påvises en statistisk signifikant forskel på et niveau, som normalt kræves ( $\alpha < 0,05$ ).

Hvis vi desuagtet ser på tendenserne i undersøgelsen, fremstår det tydeligt, at eleverne fra kerneskolere ligger markant højere på flertallet af de fem delskalaer og i alt (Dif Total digital score), hvilket betyder, at kerneskolernes elever har lært mere end de øvrige elever. Endvidere fremgår, at indsatsgruppe 2, dvs. elever fra projektets øvrige skoler, hvor læreren angiver, at der er gjort en indsats, har en systematisk større resultatscoremæssig tilvækst end kontrolgruppen. Det skal dog understreges, at forskellene for de enkelte skalaer heller ikke her er statistisk signifikante, men resultaterne indikerer samlet set, at indsatserne kan have haft tydelige effekter på elevernes matematikfaglige udbytte af indsatsen.

<sup>37</sup> Det normale krav til signifikans er 95 pct. sikkerhed ( $\alpha < 0,05$ ).

### 11.3 Effekten på motivation og selvtillid

Af Figur 57 fremgår i højre side den målte udvikling i motivation og selvtillid. Måden at opgøre motivation og selvtillid på er beskrevet i afsnit 6.4.5.

En negativ udvikling i disse to målinger, som det fremgår for kerneskolernes elever, er ikke nødvendigvis kendetegnende for elevens *oplevelse* af motivation og selvtillid, men kan være et udslag af en større grad af realisme i elevens selvvurdering i forbindelse med besvarelsen af indikatorspørgsmålene i undersøgelsen. En indikator på selvtillid i spørgeskemaet er fx: *Jeg er sikker på, at jeg kan få bedre matematikresultater end de fleste andre elever*. En elev, der mellem de to målinger er blevet gjort opmærksom på, at han/hun muligvis har talblindhed, vil således godt kunne svare mere negativt på dette spørgsmål anden gang det stilles, uagtet at eleven har haft en positiv oplevelse af at blive anerkendt i sine udfordringer og derigennem en større motivation til at lære nyt.

Når man har arbejdet med en problemstilling, bliver man mere bevidst om de udfordringer, man har. At eleverne scorer lavere anden gang, står altså ikke nødvendigvis i modsætning til, at de har øget motivation og selvtillid, men kan snarere være et udtryk for, at de bliver mere bevidste om egne styrker og svagheder.

## 12. Konklusion

Projektet har vist, at talblindhed i høj grad er et komplekst fænomen, der strækker sig over flere dimensioner. Særligt sammenhængen mellem talblindhed, matematikvanskeligheder og andre læringsudfordringer har i projektet vist sig at være ganske kompleks.

På baggrund af observationer og andre data i projektet kan følgende pointer fremhæves:

- Elever med talblindhed vil som hovedregel præstere lavt i generelle matematiktest målrettet klassetrinnet.
- Ikke alle elever med talblindhed har matematikvanskeligheder – der er elever med talblindhed, der hurtigt og sikkert løser basale aritmetikopgaver (dvs. scorer højt i alle talblindhedstest), men som alligevel udviser adfærd der karakteriserer talblindhed (fx tæller på fingrene). Dette er dog undtagelser, der bekræfter reglen.
- Elever, der har problemer med at løse basale aritmetikopgaver (dvs. scorer lavt i testene), har ofte faglige problemer på andre områder (fx det sproglige), jf. von Aster (2007). En stor andel af elever med talblindhed har således også andre vanskeligheder (komorbiditet, fx dysleksi).
- Indsatser, der gavner elever med talblindhed, gavner som regel også andre elever, der udviser svaghed i de samme typer af aritmetiske færdigheder.
- Effektivurderingen viser, at eleverne på kerneskolerne havde en større faglig tilvækst end de øvrige elever. Dette bekræfter den forskning der viser, at elever med talblindhed gennem målrettet undervisning og træning kan opøve strategier, der gør dem i stand til at klare sig i skolen og hverdagen.

### 12.1 Tragtmodellens elementer

Projektet har arbejdet med en modificeret udgave af tragtmodellen (se afsnit 6) med fokus på trin 2, den digitale test, og trin 3, samtaletesten. Det første trin i tragtmodellen, observationsguiden, var tænkt som grundlaget for udvælgelse af elever til den digitale test. I dette projekt er der i vidt omfang anvendt eksisterende data fra tidligere afholdte test og prøver. Yderligere har sidste trin i modellen ikke vist sig at være hensigtsmæssigt, da PPR ikke kan bidrage på den måde, som tragtmodellen forudsatte. Nedenfor gennemgås konklusioner fra afprøvningen af tragtmodellens fire trin.

#### 12.1.1 Observationsguiden

Undersøgelsen tyder på, at observationsguiden som første led næppe kommer til at fungere operationelt, da den er meget ressourcekrævende og omfatter flere dimensioner. Spørgeskemaundersøgelsen til lærere peger i retning af, at udbyttet isoleret set ikke står mål med indsatsen i det foreliggende format.

Derimod kan en tilrettet udgave vise sig som et værdifuldt analyseredskab senere i processen og i dialogen mellem vejledere og matematiklærere. Det kan overvejes, om observationsguiden bedre anvendes som vejledende/oplysende materiale i dialogen mellem lærere og vejledere om arbejdet med talblindhed, fx som supplement til samtaletesten, end som en selvstændig del af testforløbet.

#### 12.1.2 Den digitale test

Med få justeringer har den digitale test fungeret godt. Justeringerne har omfattet 1) frasortering af enkelte opgaver forud for afprøvningen, 2) funktionalitetsforbedringer fx ved indlejring af instruktionsvideo mellem hver deltest og 3) scoringsmodellen, hvor både rigtighed og svarhastighed indgår

i vurderingen af opgavebesvarelsen. Det er en stor kvalitet, at der indhentes et meget stort data-materiale (godt 300 opgaver pr. elev) på ganske kort tid (lav statistisk usikkerhed/høj statistisk reliabilitet). Blandt den digitale tests svagheder skal fremhæves, at test-retest-reliabiliteten kunne være højere ( $r=0,73$ ), hvilket sandsynligvis beror på følsomhed i forhold til instruktion, testafvikling og elevernes aktuelle dagform.

Analyserne peger i retning af, at de fem dimensioner måler lidt forskellige problemstillinger, idet skalaerne, der hver især fungerer homogent, kun har en begrænset korrelation parvis. De fem delskalaer vurderes alle som væsentlige indikatorer i forbindelse med talblindhed.

Lærerne har vurderet, at den digitale test er et brugbart og let anvendeligt redskab til at skabe overblik over tegn på talblindhed hos eleverne – dels ud fra den samlede score og dels ud fra de fem deltest – med en meget begrænset ressourceindsats. Både lærere og elever finder testen enkel at anvende og giver udtryk for, at den giver en hurtig og praktisk afdækning af en række komplekse problemstillinger. Lærerne beskriver elevernes oplevelse med testen som positiv – særligt opgavernes lethed medfører, at alle elever har mulighed for at få succesoplevelser ved besvarelse af opgaverne, hvor det således i høj grad bliver løsningshastigheden, der er afgørende for, om eleven scorer højt eller lavt.

Imidlertid skal det fremhæves, at den digitale test måler latente egenskaber, der relaterer sig til talblindhed, men ikke nødvendigvis kan relateres direkte til undervisningen. Opgaverne skal således ikke betragtes som indholdsangivende for pædagogiske indsatser, og testen bør derfor anvendes i kombination med undersøgelser, der i højere grad orienterer sig mod didaktisk indhold med henblik på at iværksætte passende pædagogiske indsatser til at afhjælpe elevens vanskeligheder. Derudover er der behov for videreudvikling af resultat anvendelsen for så vidt angår de enkelte delskalaer.

Samlet set vurderes det, at testen i sin nuværende form (dvs. opsætning, vejledning, scoring og resultatvisning) er klar til anvendelse i større skala, idet den giver skolen et udgangspunkt for at vurdere, om der er grundlag for mistanke om talblindhed og dermed om yderligere undersøgelse og testning bør foretages.

### 12.1.3 Samtaletesten

Samtaletestens formål er at give grundlag for en grundig analyse af elevens tegn på talblindhed – bl.a. som opfølgning på den digitale test – dels gennem observation og dialog ved gennemførelse af samtaletesten og dels på baggrund af samtaletestens resultater. Samtaletestens indhold og dialogbaserede afviklingsform skaber et fundament for, at læreren kan opnå brugbar indsigt i elevens strategier til håndtering af simple talforståelses- og regneopgaver. Lærerne vurderer, at det giver grundlag for en direkte anvendelse i forhold til didaktiske overvejelser og valg af pædagogiske indsatser. Samtaletesten var samlet set det redskab, som lærerne i spørgeskemaundersøgelsen gav udtryk for at have haft størst nytte af.

Valideringen mod den hollandske RD4-test viste, at der var stærke lighedspunkter, og at samtaletesten som helhed kan anses for at måle det, som er intentionen. Dette til trods for at analyse af besvarelsenerne viser, at enkelte opgaver ikke passede godt til modellen. Imidlertid udgør det en væsentlig begrænsning, at den med stor sandsynlighed kun kan anvendes én gang pr. elev med et retvisende resultat og kun til elever på 4. klassetrin. Selvom samtaletesten som helhed har fungeret overordentligt godt i projektet, vurderes der at være et stort potentiale ved en videreudvikling.

#### 12.1.4 PPR

Baseret på samarbejdet med PPR og tilbagemeldinger fra skolerne blev det vurderet, at det ikke forekom direkte foreneligt med PPR's opgaver og kompetencer at stå for diagnosticering af talblindhed. Dette bygger bl.a. på, at PPR ikke kan forventes at have tilstrækkelig matematikfaglig viden om og indsigt i elevernes faglige udvikling.

Det har dermed ikke været muligt at afprøve PPR's rolle som den oprindeligt var beskrevet i tragmodellen. Der blev gennemført en mindre afprøvning på to kerneskolers med henblik på at afprøve om skolerne selv (i det omfang det skønnes nødvendigt) vil kunne varetage de funktioner som PPR var tiltænkt i tragmodellen – evt. med støtte fra PPR.

## 12.2 Pædagogiske indsatser

Da udvikling af effektive pædagogiske indsatser til elever med talblindhed er en kompleks opgave, som kræver individuel tilpasning til den enkelte elevs forudsætninger og behov, kan den ikke løses på enkel vis. Erfaringerne fra projektet viser, at der ikke findes en universel løsning, én indsats, der er lige passende for alle elever med talblindhed.

Der kan derfor overordnet drages følgende konklusioner:

- Der er generelt stor usikkerhed om, hvad talblindhed er, og derfor også hvordan man kan tilrettelægge undervisning, som kan afhjælpe virkningerne heraf.
- Meget af det, som lærerne intuitivt gør, afspejler det, som eksperterne anbefaler. Imidlertid skaber det meget større sikkerhed, når indsatsen forankres i viden om elevens konkrete udfordringer samt teoretisk viden om didaktiske og pædagogiske virkemidler.
- Mange elever, der har talblindhed og/eller matematikvanskeligheder, har udviklet et anstrengt forhold til matematik.
- Det, som er godt for elever med talblindhed, er også godt for elever med andre matematikvanskeligheder i forbindelse med aritmetik, tal og mængder.

### 12.2.1 Samarbejdet med kerneskolerne

Erfaringer fra samarbejdet med kerneskolerne peger på en række didaktiske og pædagogiske principper, som kan danne grundlag for en indsats målrettet elevens konkrete udfordringer med talblindhed eller matematikvanskeligheder forbundet med aritmetik, tal og mængder.

Det drejer sig især om 1) mulighed for succesoplevelser, 2) positive relationer og rammer om matematikken, 3) samarbejde med elevens øvrige lærere, 4) brug af konkrete repræsentationer og gentagelser og 5) afdækning af elevens konkrete udfordringer, bl.a. ved hjælp af samtaletesten.

Konklusionen skal læses med afsæt i, at mængden af indsatsmateriale er begrænset, bl.a. grundet den reducerede afprøvningsperiode, og at elevgrundlaget har været reduceret i forhold til det, der oprindeligt var planlagt. Undersøgelsen ville naturligvis have været styrket i væsentlig grad, hvis skolernes vilkår havde været bedre, herunder hvis de tidsmæssige rammer havde givet mulighed for en afprøvningsperiode på tre måneder som oprindeligt planlagt og i øvrigt hvis et større antal elever i risiko for talblindhed havde været inddraget i afprøvningen af pædagogiske indsatser.

### 12.2.2 Effekten af pædagogiske indsatser

Effektvurderingen viser, at der overordnet set spores virkninger af at gøre en indsats. Materialets størrelse er imidlertid utilstrækkeligt i forhold til at påvise signifikante effekter og at sammenligne

de forskellige indsatser i de fire kerneskkolekommuner. Udvælgelsen af kerneskkoler, valg af målgruppe (4. klasse) samt de sundhedsmæssige udfordringer, som Covid-19 medførte i perioden, betød, at der var et forholdsvist begrænset materiale at analysere effekten ud fra. Dette gælder både, hvad angår antallet af elever med tegn på talblindhed, omfanget af elevernes udfordringer samt længden af afprøvningsperioden for pædagogiske indsatser.

Alt i alt peger undersøgelsens resultater dog på, at der har været positive effekter af de pædagogiske indsatser, omend størrelsen ikke kan opgøres signifikant.

## 13. Anbefalinger

I projektets opdrag udbedes anbefalinger på en række områder. Bestillingen omfatter *”En afsluttende rapport om projektet, der kan offentliggøres og indeholder: [...]*

- *anbefalinger til, hvilken status testen kan få,*
- *anbefalinger til videre udvikling, herunder udførelsen af en eventuel fase 2,*
- *anbefalinger til øvrige tiltag, der kan understøtte en indsats for talblindhed af høj kvalitet”*

Med udgangspunkt i projektets erfaringer samt aktuel forskning og viden på feltet gives neden for en række anbefalinger på de ønskede områder og begrundelser herfor.

### 13.1 Anbefalinger om talblindhedstestens status

Det fremgår af opgavebeskrivelsen, at projektet skal kunne *”... danne grundlag for at vurdere, om testen kan anvendes som en national test, der diagnosticerer elever med talblindhed, eller om testen skal have en anden status”*.

#### Anbefalinger:

1. Anvendelse af den digitale test som screeningredskab.
2. Anvendelse af samtaletesten som redskab til videre undersøgelse.
3. Grundlag for vurdering af rettigheder til pædagogisk og/eller økonomisk støtte

#### Kommentarer til anbefalinger:

- Ad. 1 Den digitale test vurderes, set i lyset af internationale erfaringer, at kunne anvendes i sin nuværende form som screeningredskab til elever ned til 1. klassetrin<sup>38</sup>.
- Ad. 2 Det anbefales at udvikle samtaletesten med en række lineære forløb målrettet de enkelte klassetrin, der kan anvendes til mere præcist at identificere vanskeligheder som typisk følger af talblindhed (jf. også afsnit 13.2, anbefaling 2).
- Ad. 3 Efter beregning af normer og cutscores, svarende til de enkelte klassetrin, vil resultatet af en screening med den digitale test i kombination med en samtaletest kunne indgå i et beslutningsgrundlag for tildeling af rettigheder, fx ekstra tid og/eller hjælpemidler i forbindelse med test og prøver samt specialpædagogisk støtte og/eller kompenserende pædagogiske foranstaltninger.

### 13.2 Anbefalinger til videre udvikling af testens elementer

#### Anbefalinger:



*” [Skolen vil benytte en talblindhedstest], hvis der følger hjælp/økonomi med til de elever, som bliver testet som talblinde. Eller hvis de får mulighed for at få længere tid, når de skal til prøve i matematik...” – (Citat fra lærersurvey)*

<sup>38</sup> | Butterworths dyscalculia screener er der opgaveformater, der minder om formaterne i den digitale test, og opgaver, som retter sig mod elever med højere fagligt niveau i matematik, fx simple additions- og multiplikationsopgaver. Butterworth anbefaler sin screener til elever fra 6-14 år. Butterworth, B. (2003). Dyscalculia screener. London, UK: GL Assessment.



1. Udvikling af et rapporteringsmodul til den digitale test, samtaletesten og observationsguiden.
2. Udvikling af flere lineære samtaletest målrettet forskellige klassetrin.
3. Udvikling af et it-baseret registreringssystem til samtaletesten.
4. Justering af observationsguiden.
5. Tilgængeliggørelse af supplerende redskaber.

### Kommentarer til anbefalinger:

- Ad. 1 Et rapporteringsmodul til den digitale test og samtaletesten vil give læreren/vejlederen mulighed for at hente elevens resultater umiddelbart efter gennemførelse af testene eller på et senere tidspunkt.
- Ad. 2 Erfaringerne har vist, at den nuværende samtaletest kun kan anvendes én gang, og yderligere er den målrettet 4. klassetrin. Der har vist sig behov for at kunne gennemføre samtaletesten flere gange for hver elev pr. klassetrin, bl.a. for at kunne følge elevens udvikling i løbet af året. Derfor anbefales det at videreudvikle samtaletesten som et redskab, der består af en række forskellige lineære opgavesæt målrettet hvert klassetrin, således at den kan anvendes flere gange i løbet af skoleåret på hvert af de udvalgte klassetrin.
- Ad. 3 Samtaletest bør afvikles ved fysisk tilstedeværelse fra både lærer og elev, hvor læreren løbende registrerer elevens besvarelser i et online registreringssystem, der gør det muligt at beregne og rapportere registreringer af elevsvar og en scoring af disse.
- Ad. 4 Det anbefales at tilpasse observationsguiden i forlængelse af de analyser, der blev gennemført i forbindelse med projektet, og i øvrigt anvende observationsguiden på en anden måde end beskrevet i tragtmodellen. Tilpasningen består i at opdele observationsguiden i en hhv. faglig og en adfærdsafdækkende del, at reducere omfanget af guiden til de spørgsmål, der fungerer bedst (jf. rapportens afsnit 8.1) samt at understøtte anvendelsen af observationsguiden gennem en it-baseret løsning på linje med den, der er anvendt i dette projekt. Endvidere anbefales, at observationsguiden anvendes som vejledende/oplysende materiale i dialogen mellem lærere og vejledere om arbejdet med talblindhed – fx som supplement til samtaletesten i stedet for en selvstændig del af testforløbet.
- Ad. 5 Det anbefales som hidtil at lade skolerne varetage opgaven om afklaring af elevernes vanskeligheder med de redskaber, som mange skoler allerede har til rådighed (CHIPS, Visuel/motorisk kodeprøve, mv.). Det anbefales, at PPR som hidtil kun indgår i samarbejdet om elever med tegn på talblindhed i en rådgivende rolle, evt. med hensyn til validering af undersøgelser foretaget af skolen. Det anbefales at samle og tilgængeliggøre (fx på EMU) de redskaber, som skolerne kan få brug for (CHIPS, Visuel/motorisk kodeprøve, mv.) ved afgørelse af, om elever har kognitive vanskeligheder.

## 13.3 Anbefalinger til øvrige tiltag

### Anbefalinger:

1. Ekspertrådgivning med afsæt i danske og internationale forskere.
2. Etablering af følgeforskning.
3. Uddannelse til skolerne om talblindhed og testresultater.
4. Udbud af pædagogisk materiale og undervisningsforløb.

**Kommentarer til anbefalinger:**

- Ad. 1 I nærværende projekt har der været etableret en ekspertgruppe, som løbende har givet værdifuldt input til analyser og pædagogiske indsatser samt feedback på fremdrift. Det har betydet, at projektets indhold og plan løbende er blevet kvalitetssikret og justeret efter behov. Denne bistand har været vigtig for at sikre projektets aktualitet og kvalitet. Det anbefales derfor, at denne model videreføres i en evt. fase 2 af projektet.
- Ad. 2 En fase 2 vil kunne tilvejebringe et unikt datamateriale i forhold til den internationale forskning i talblindhed. Det anbefales derfor, at der arbejdes struktureret med at sikre, at dette materiale stilles til rådighed for følgeforskning med henblik på at skabe ny viden om talblindhed, matematikvanskeligheder i forbindelse med talblindhed og pædagogiske virkemidler.
- Ad. 3 Igennem hele projektet har spørgsmål fra skolerne og dialogen med kerneskolerne vist, at selv de dygtigste og mest vidende lærere og matematikvejledere oplever et stort behov for sikker, opdateret og valideret viden om anvendelsen af test, fortolkning af testresultater og pædagogiske indsatser for elever med denne type vanskeligheder. Skoler, lærere og vejledere efterspørger både korte workshops/seminarer med specifikt indhold samt mere generelle og længerevarende uddannelsesforløb (jf. afsnit 10.3.3). En struktureret uddannelsesindsats med inspiration i den, der er foregået i de seneste årtier på læseområdet, vurderes at kunne løfte skolernes indsats på feltet.
- Ad. 4 I forlængelse af de afviklede projekter er der etableret et område på EMU, som samler materiale om talblindhed. Det anbefales at videreudvikle og udvide denne erfaringsbase, så det også kommer til at omfatte de mest anvendte analyseredskaber til afklaring af elevernes vanskeligheder (se fx afsnit 9.3.3), pædagogiske forløb og vejledningsmateriale som kan være til anvendelse, inspiration og nytte for skolernes arbejde med talblindhedsrelaterede matematikvanskeligheder.



*"Jeg håber, at der fremadrettet vil komme mere hjælp til, hvordan man så hjælper denne gruppe videre. På hvilke områder adskiller hjælpen sig fra de øvrige elever i vanskeligheder? Hvad skal afhjælpes gennem andre hjælpemidler, og hvad skal fortsat trænes og trænes? Og på hvilke områder vil behovet være det samme – måske bare kræve mere tid og træning?"*

*– (Citat fra lærersurvey)*

*"Vi mangler i den grad inspiration til, hvad vi skal gøre med de elever, som er særligt udfordret, og hvor talblindhed kunne komme på tale" – (Citat fra lærersurvey)*

## 13.4 Anbefalinger til udførelsen af en eventuel fase 2

Af opgavebeskrivelsen fremgår, at det nuværende projekt skal *"... bidrage til at afklare, om der er grundlag for at foretage et kontrolleret forsøg af test, testresultater, vejledningsmateriale og pædagogiske indsatser effekter i en fase 2. Fase 2 iværksættes kun, hvis det efter fase 1 vurderes, at testen og vejledningsmateriale samt de pædagogiske indsatser vurderes at have den fornødne forskningsmæssige kvalitet, som en national test af talblindhed (diagnostisk eller vejledende) skal have, for at der meningsfuldt kan gennemføres et kontrolleret forsøg herom."*



*"Det har været et fantastisk projekt, som har givet mig et nyt blik på talforståelse som et mere nuanceret begreb. Jeg kan benytte denne viden til min undervisning i klasserne samt med elever i vanskeligheder. Jeg håber projektet fortsætter, således at vi på skolerne kan blive endnu klogere på, hvordan vi hjælper denne elevgruppe på en konstruktiv måde." – (Citat fra lærer-survey)*

### 13.4.1 Overordnede anbefalinger vedr. fase 2

Der vurderes at være et solidt grundlag for det videre arbejde med talblindhed, og det anbefales derfor at gennemføre en fase 2 af projektet.

Ekspertter og forskning peger på, at talblindhed er en medfødt og varig neurologisk udviklingsforstyrrelse, som optræder i forskellige grader, men at virkningerne heraf (tegnene på talblindhed) som regel reduceres over tid og især ved en målrettet pædagogisk indsats. Således kan det i vidt omfang imødegås, at talblindhed bliver en hindring for en normal hverdag eller for videre uddannelse. Dette projekt (fase 1) har demonstreret, at disse tegn på talblindhed kan måles sikkert og pålideligt med en test, og på trods af et begrænset datagrundlag i fase 1 peger erfaringerne (ligesom de internationale erfaringer) entydigt i retning af, at det nytter at gøre en målrettet indsats med undervisning og træning.

Fase 2 anbefales at være målrettet udbredelse af kendskab til arbejdet med talblindhed og indsamling af erfaringer med effekterne af forskellige pædagogiske indsatser ved forløb af længere varighed. Det anbefales således at tilbyde alle interesserede grundskoler med elever på 1.-4. klassetrin adgang til at anvende testredskaberne. Dels for at udbrede viden om talblindhed, erfaring med test heraf og interesse for pædagogiske indsatser. Og dels for at kunne oparbejde et solidt datagrundlag for udregning af normer på de enkelte klassetrin.

Yderligere anbefales det, at der rekrutteres et stort antal skoler (med forskelligt elevgrundlag) til at deltage mere systematisk i storskalaafprøvning af pædagogiske indsatser målrettet forskellige tegn på talblindhed. Effekten af forskellige pædagogiske indsatser kan registreres ved hjælp af den digitale test (med fem delskalaer), der benyttes ved starten og afslutningen af forløbet (præ- og post-test) samt ved måling af langtidsvirkninger (fx et år senere). For elever hvor den digitale test viser forskellige tegn på talblindhed (fx problemer med talforståelse, subitisme, ANS osv.), kan resultaterne fra præ- og post-testen fra den digitale test bruges til at vurdere udviklingen indenfor de fem skalaer og i alt. Resultaterne fra post-testen og den målte udvikling kan yderligere suppleres af resultater fra en videreudviklet samtaletest (evt. kombineret med observationsguiden) med henblik på at belyse faglige styrker, svagheder og behov hos eleven. På den måde kan testresultaterne bruges til både at identificere, hvad der virker bedst i forhold til forskellige tegn på talblindhed, og til at identificere hvilke indsatser der vil være behov for fremover.

For at få det bedste udbytte af en fase 2 anbefales det, at der iværksættes kompetenceudvikling (kombination af workshops, webinarer og kurser/korte uddannelsesforløb) for at klæde lærere og vejledere bedst muligt på til opgaven med at afvikle test, fortolke resultater og anvende testresultater i forbindelse med valg af pædagogisk indsats, vurdering af effekt og feedback til elever.

### 13.4.2 Detaljerede anbefalinger vedr. fase 2

1. Udvidelse af målgruppen til at dække 1-4. klassetrin.
2. Justering af målgruppen for indsatsen.
3. Afprøvning af pædagogiske indsatser i et kontrolleret eksperiment i stor skala.

4. Længerevarende kontrolleret eksperiment.
5. Mulighed for fleksibel anvendelse af den digitale test og samtaletest.
6. Åben adgang for relevante skoler.
7. Udvikling af lettilgængelige multimodale vejledningsmaterialer.

### Kommentarer til anbefalinger:

- Ad. 1 Den aktuelle forskning i talblindhed og matematikvanskeligheder påpeger vigtigheden af at begynde arbejdet med eleverne tidligt. Erfaringer fra projektet peger på, at der er et ønske om at anvende talblindhedstest tidligere, dvs. på de mindre klassetrin, og dette vil være muligt med tilpasninger af de nuværende redskaber gennem beregning af klassetrinsnormer for den digitale test, observationsguide revideres som beskrevet i denne rapport og samtaletesten videreudvikles i forskellige versioner med opgaver tilpasset faglige niveauer på de forskellige klassetrin.
- Ad. 2 I den afgrænsning af målgruppen, der har været anvendt i projektet, er en væsentlig gruppe af elever med behov for støtte uden for målgruppens rammer. Anbefalingen omfatter en justering af målgruppen fra elever med den "rene talblindhed" (pDD) til elever med vanskeligheder som følge af tegn på talblindhed, inklusive elever med kombinerede vanskeligheder (komorbiditet), jf. afsnit 3. Endvidere er der elever i gruppen af talblinde, der ikke har matematikvanskeligheder. Anbefalingen går på i højere grad at målrette indsatsen til de elever, som har et behov for støtte.
- Ad. 3 Der er behov for et stort datamateriale for at afgøre kvaliteten og effekten af hver af de pædagogiske indsatser, der indgår i en eventuel fase 2. Skal undersøgelsen give statistisk sikre resultater bør mindst 100 elever afprøve hver pædagogisk indsats inden for det pågældende klassetrin. Hvis fire pædagogiske indsatser afprøves på hvert klassetrin fra 1.-4. klasse skal forsøgsgruppen altså omfatte mindst 1.600 elever og meget gerne flere. Hvis 5-6 pct. af eleverne antages at have tegn på talblindhed, betyder det en deltagelse på mindst 30.000 elever (svarende til fire årgange á 60 elever på ca. 125 skoler). Ud over de 4 eksperiment-grupper pr. klassetrin vil der være behov for at et større antal elever fra hvert klassetrin deltager i projektet som kontrolgrupper (gerne et tilsvarende antal skoler). Det anbefales at opfordre til, at hele klasser benytter den digitale test som screeningsredskab ved indgangen til en fase 2 for at få et sikkert overblik over standarder og normer for så vidt angår tegn på talblindhed på det specifikke klassetrin.
- Ad. 4 Med henblik på at indhente erfaringer for så vidt angår både kortsigtede/umiddelbare og langsigtede/varige effekter af pædagogiske indsatser foreslås det at lade eksperimentet strække sig over to skoleår. Visse indsatser giver hurtigt udslag i synlige effekter (målt ved forskellen mellem præ- og post-test), hvorimod andre kræver lang tids ihærdigt arbejde, før gevinsterne er målbare (effekterne måles ved at sammenligne med en tilsvarende test gennemført lang tid senere - en såkaldt retention-test<sup>39</sup>). Nyere forskning<sup>40</sup> viser, at der i høj grad mangler viden om, hvilke pædagogiske indsatser og virkemidler, der har de største

<sup>39</sup> Det foreslåede test-design svarer til det, der blev anvendt og beskrevet i Bonde, M. T., Makransky, G., Wandall, J., Larsen, M. V., Morsing, M., Jarmer, H., & Sommer, M. O. (2014). Improving biotech education through gamified laboratory simulations. *Nature biotechnology*, 32(7), 694-697.

<sup>40</sup> Lu, Y., Ma, M., Chen, G., & Zhou, X. (2021). Can abacus course eradicate developmental dyscalculia. *Psychology in the Schools*, 58(2), 235-251.

positive effekter fagligt såvel som personligt for elever med talblindhed – især på længere sigt<sup>41</sup>.

De umiddelbare effekter vurderes ved måling umiddelbart før og efter implementering (præ- og post-test) den pædagogiske indsats, der ønskes vurderet. For også at kunne vurdere langtidsvirkningerne anbefales at der også gennemføres en måling fx et år efter post-testen (en retention-test). Nogle elever vil til den tid gå i 5-6. klasse, hvilket ikke alle dele af talblindhedstesten er udviklet til at foretage målinger på. Imidlertid har den digitale test i dette projekt også været anvendt af elever fra 5-8. klassetrin på kerneskolerne, hvilket ikke har givet anledning til bekymringer eller problemer af nogen art. Samtaletesten vil dog kun kunne bruges meningsfuldt op til 4. klassetrin (det vil sige de yngre årgange ved projektets afslutning), men det er først og fremmest den digitale test der anbefales anvendt til effekt-målingen på langt sigt.

- Ad. 5 På grundlag af feedback fra skolerne vurderer vi, at testredskaberne (den digitale test, samtaletest og observationsguide) med fordel kan udbydes fleksibelt, så testene kan tages, når skolerne har behovet og på tidspunkter der passer ind i skolernes hverdag.



*"Det er vigtigt, at testen altid er "åben" og kan tilgås hele året. Har hørt fra dansk-kolleger, at det ikke er sådan for ordblindtesten. Det er nemlig vigtigt, at testen kan tages, når det er aktuelt for det enkelte barn." – (Citat fra lærersurvey)*

- Ad. 6 Der vurderes at være en stor interesse fra skolernes side i at kunne komme i gang med at anvende talblindhedstesten, selv når redskaberne er under udvikling. Det gennemførte projekt fik umiddelbart tilmeldinger fra 800 skoler, hvoraf halvdelen af skolerne gennemførte afprøvningen i større eller mindre omfang på trods af periodevis nedlukning af en stor del af skolerne i projektperioden relateret til COVID-19. Ydermere viser resultaterne af spørgeskemaundersøgelsen blandt lærerne, at der er stor interesse for at tage redskaber, ikke mindst en digital talblindhedstest, i brug. Når et it-system er etableret, betyder yderligere deltagelse marginalt ekstra ressourceforbrug, og det giver et forbedret datagrundlag for en evt. efterfølgende idriftsættelse. Det anbefales derfor at give alle relevante skoler åben adgang til alle redskaberne (særligt den digitale test og samtaletesten) i en fase 2.
- Ad. 7 Med henblik på at lette anvendelsen af redskaberne skal vejledninger i videst muligt omfang integreres i redskaberne. Ud over de manualer og tekststunge vejledninger, der er nødvendige for at sikre gennemsigtighed om regler og anvendelsen af testene<sup>42</sup>, anbefales det at begrænse omfanget af tekst i vejledende materialer for at øge anvendeligheden. Det kan gøres ved brug af multimodaliteter, fx illustrationer, lyd og video, så læseindsatsen begrænses. Erfaringerne fra projektets fase 1 er, at vejledende materiale kun anvendes, hvis det præsenteres på rette tid og sted, så det giver mening for læreren og ikke fordrer en betydelig læseindsats.



*"Vejledningerne til læreren kunne med fordel suppleres med billeder/video." – (Citat fra lærersurvey)*

<sup>41</sup> Haberstroh, S., & Schulte-Körne, G. (2019). The diagnosis and treatment of dyscalculia. Deutsches Ärzteblatt International, 116(7), 107.

<sup>42</sup> Jf. fx Vejledning til Ordblindetesten (seneste udgave kan hentes her [https://ordblindetest.nu/vejleder/pdf/Vejledning\\_til\\_Ordblindetesten\\_Version8.pdf](https://ordblindetest.nu/vejleder/pdf/Vejledning_til_Ordblindetesten_Version8.pdf))

## **Epinion København**

Ryesgade 3F  
2200 København N  
Denmark

T: +45 87 30 95 00

E: [copenhagen@epinionglobe.com](mailto:copenhagen@epinionglobe.com)

[www.epinionglobe.com](http://www.epinionglobe.com)

## **Epinion Aarhus**

Fellow Workspace  
Mariane Thomsens Gade 4B  
8000 Aarhus  
Denmark

T: +45 87 30 95 00

E: [aarhus@epinionglobe.com](mailto:aarhus@epinionglobe.com)