

Modtager(e) > Ekspertpanel og referencegruppe

Resultater og justeringsmuligheder efter pilottest af måling blandt studerende

Dette notat gengiver hovedresultaterne af en række valideringsanalyser af pilotundersøgelsen af LEARN, som er gennemført af UFM og Danmarks Evalueringsinstitut (EVA). Analyserne viser overordnet, at spørgsmålene empirisk passer til den teoretiske model, og at de undersøgte skalaer generelt opfører sig ensartet på tværs af sektorer og baggrundskaraktistika. Der er dog visse spørgsmål, der giver anledning til opmærksomhed og steder, hvor skalaerne muligvis kan forbedres ved enten at tilføje eller fjerne spørgsmål.

I første afsnit af notatet uddybes de empiriske resultater og i anden del introduceres en række forslag til at justere spørgeskemaet.

1. Overordnede resultater

For at validere spørgeskemaet, er der gennemført forskellige analyser af det samlede spørgeskema. Det drejer sig blandt andet om analyser af skemaets eksterne validitet, indikatorernes styrke og enkeltspørgsmålenes validitet i forhold til indikatorerne (i denne sammenhæng også benævnt skalaer eller faktorer).

På indikatorniveau er der foretaget valideringsanalyser af hver af de ni indikatorer, som også indgår i LEARN, og herunder en Rasch-analyse, hvor det undersøges om spørgsmålene i hver af de ni indikatorer alene måler den skala, som spørgsmålene teoretisk tilhører. Der er også gennemført konfirmative faktoranalyser, der undersøger om alle skalaerne samlet passer til den teoretiske model. Analysemetoderne beskrives til sidst i notatet og uddybet i de vedlagte bilag.

Resultaterne viser overordnet set, at spørgeskemaet empirisk passer til de teoretiske forventninger. På skalaniveau har syv ud af ni skalaer tilfredsstillende værdier for intern konsistens. Resultaterne fra den konfirmative faktoranalyse bekræfter i store træk, at de teoretiske dimensioner i LEARN kan genfindes empirisk i de indsamlede pilotdata, og at de enkelte spørgsmål passer til de tilhørende skalaer.

Resultaterne af Rasch-analysen viser, at syv ud af ni skalaer er anvendelige uden yderligere justeringer. Denne vurdering er baseret på, hvordan data passer til den teoretiske model (item-fit) og om baggrundsforhold som køn, sektor og alder har en signifikant betydning for svarangivelsen (men ikke nødvendigvis niveauet). Der er kun én skala, hvor der er observeret signifikant sektorvariation i svarene på enkelte items.

Derudover viser resultaterne, at alle spørgeskemaets skalaer som forventet korrelerer med et spørgsmål om frafald på et signifikant niveau: Jo højere score i spørgeskemaet, desto mindre overvejer de studerende at droppe ud. Undtagelsen er som forventet skalaen for "Fragmenteret læring/Overfladelæring", hvor højere scorer hænger sammen med større frafaldssandsynlighed vender omvendt¹.

¹ Her korrelerer en højere score på overfladelæring positivt med overvejelser om frafald.

24. august 2018

**Uddannelses- og
Forskningsministeriet**
Analyse

Børsgade 4
Postboks 2135
1015 København K
Tel. 3392 9700
Fax 3332 3501
ufm@ufm.dk
www.ufm.dk

CVR-nr. 1680 5408

Tabel 1

Overordnede resultater fra validering af pilotundersøgelse

	Intern konsistens (reliabilitet)	Item-fit	Faktorloading	Baggrundsforholds betydning for skalaernes sammenhæng	Sammenhæng med frafald
Samlet vurdering	Syv ud af ni skalaer har tilfredsstillende værdier.	I otte ud af ni skalaer loader alle items på samme skala	Alle spørgsmål har tilfredsstillende faktorloadings.	Otte ud af ni skalaer har ingen signifikant sektorvarians.	Alle skalaer korrelerer som forventet med frafald, også på sektorniveau. Alle værdier er signifikante på 0,01-niveau
Læringsmiljø	Konfirmativ faktoranalyse viser, at de seks skalaer for læringsmiljø hver især har tilstrækkeligt høje faktorloadings. Tilsammen er der imidlertid ikke et optimalt model-fit. (RMSEA = 0,04; CFI = 0,62; SRMR = 0,31)				
Interesse og motivation	Tilfredsstillende (0,767) På hovedområdet er skalaen ikke tilfredsstillende på UNI SUND.	Tilfredsstillende	Over 0,4 (0,522-0,780)	Nej	Korrelerer negativt med frafald (-0,57)
Konstruktiv feedback	Tilfredsstillende (0,824)	Tilfredsstillende	Over 0,4 (0,632-0,816)	Nej	Korrelerer negativt med frafald (-0,31)
Læring og forståelse	Under grænseværdi (0,677)	Tilfredsstillende	Over 0,4 (0,534-0,647)	Signifikant sektorvariation på to spørgsmål	Korrelerer negativt med frafald (-0,39)
Overensstemmelse (Alignment)	Tilfredsstillende (0,768)	Tilfredsstillende	Over 0,4 (0,479-0,783)	Nej	Korrelerer negativt med frafald (-0,43)
Støtte fra Medstuderende	Under grænseværdi (0,686)	Tilfredsstillende	Over 0,4 (0,516-0,683)	Nej – dog findes ikke-signifikante forskelle mellem sektorerne.	Korrelerer negativt med frafald (-0,29)
Underviserinteraktion	Tilfredsstillende (0,784)	Tilfredsstillende	Over 0,4 (0,483-0,767)	Nej – dog findes ikke-signifikant sektorvariation. Der kan ses nogen effekt af køn på svarangivelsen for et enkelt item.	Korrelerer negativt med frafald (-0,39)
Lærings- tilgange	Konfirmativ faktoranalyse viser, at de tre tilgange genfindes empirisk i et tilstrækkeligt, men ikke optimalt fit. (RMSEA = 0,06; CFI = 0,76; SRMR = 0,11)				
Dybdelærings-tilgang	Tilfredsstillende (0,751)	Tilfredsstillende	Over 0,4 (0,597-0,748)	Nej	Korrelerer negativt med frafald (-0,35)
Organiseret læringstilgang	Tilfredsstillende (0,763)	Tilfredsstillende	Over 0,4 (0,662-0,721)	Nej	Korrelerer negativt med frafald (-0,25)
Overfladisk læringstilgang	Tilfredsstillende (0,785)	Tilfredsstillende	Over 0,4 (0,585-0,806)	Nej	Korrelerer positivt med frafald (0,32)
Analysemetoder	Faktoranalyse af skalaerne enkeltvis	Rasch-analyse (item-fit)	Konfirmativ faktoranalyse	Rasch-analyse (differentiel item-funktion)	Bivariat regressionsanalyse,
Grænseværdier	Cronbach's Alpha > 0,7	p > 0,01	Faktorloading > 0,4	Afhænger af antal spørgsmål i skalaen (fra p < 0,00125-0,0033)	p < 0,01

Anm.: Grænseværdier for model-fit: RMSEA < 0,06; CFI > 0,90; SRMR < 0,08

Kilde: Uddannelses- og Forskningsministeriet på baggrund af pilotundersøgelsen

2. Særlige opmærksomhedspunkter og justeringsmuligheder

Det fremgår af Tabel 1, at de tre skalaer for læringstilgange overordnet er statistisk stærke. Det er derfor vurderingen, at resultaterne ikke giver anledning til at justere disse skalaer. Resultaterne giver dog anledning til at overveje mulige justeringer af skalaerne for læringsmiljø. I det følgende gennemgås derfor kun resultaterne fra valideringsanalyserne for de seks skalaer for læringsmiljø. De mulige justeringer er opsummeret i Tabel 2.

Det fremgår af Tabel 1, at syv ud af ni skalaer opfylder konservative kriterier for målingsvaliditet og intern konsistens på overordnet niveau. De to skalaer, der ikke lever op til disse kriterier er *Støtte fra medstuderende* og *Læring og forståelse*, idet skalaernes interne konsistens ligger under grænseværdien på 0,7, jf. Tabel 1. Rasch-analysens resultater peger på, at de samme to skalaer har et forbedringspotentiale.

Uddannelses- og
Forskningsministeriet

En konfirmativ faktoranalyse af datamaterialet viser, at alle indikatorernes *faktorloading* er større end 0,4, hvilket er et udtryk for en tilstrækkelig grad af målingsvaliditet med de undersøgte skalaer. En eksplorativ faktoranalyse af den teoretiske model viser dog, at tre af spørgsmålene ikke loader entydigt på den teoretiske skala, og det anslås derfor, at der mulighed for at styrke spørgeskemaet ved at udelade disse tre spørgsmål. Denne vurdering er dog udelukkende baseret på et empirisk grundlag, og vurderingen er ikke teoretisk funderet.

Resultaterne indikerer, at det overordnet er muligt at forbedre skalaernes interne konsistens ved at *tilføje* spørgsmål.² Samtidig indikerer en eksplorativ faktoranalyse, at skalamodellen for læringsmiljøet kan blive mere klar ved at *fjerne* tre spørgsmål. Grundlaget for disse justeringsforslag uddybes i de bagvedliggende notater.

Der er flere forbehold ved at justere i spørgeskemaet. LEARN er allerede testet i både Danmark og en række andre lande, og den grundlæggende teoretiske model er derfor relativt robust. Det taler for at være forsigtig ved at ændre ved den grundlæggende model.

I forhold til at tilføje nye spørgsmål (f.eks. for at øge skalaernes reliabilitet) er der behov for at afveje et generelt hensyn om at begrænse antallet af spørgsmål i skemaet. Det skal derfor besluttes, om spørgeskemaet skal justeres på baggrund af analyserne af pilotdata, eller om den teoretiske model skal fastholdes.

For en del af justeringer er det undersøgt om de forbedrer eller svækker spørgeskemaet og skalaerne i forhold til de nævnte grænseværdier. Det gælder på nuværende tidspunkt ikke alle mulighederne.

I det følgende uddybes mulighederne for at justere de enkelte indikatorer.

² De spørgsmål, der foreslås tilføjet, stammer fra en række backup-spørgsmål, der blev testet i pilotundersøgelsen. Backup-spørgsmålene er formuleret af Danmarks Evalueringsinstitut med udgangspunkt i hhv. Motivated Strategies for Learning Questionnaire (MSLQ) og erfaringer fra lignende undersøgelser.

Interesse og motivation

Skalaen *Interesse og motivation* fungerer overordnet fint i forhold til intern konsistens, item-fit og faktorstruktur. Skalavalideringer på hovedområdeniveau viser dog lave værdier for skalaens interne konsistens på UNI. Valideringsanalyser af de udvidede skalaer viser, at den interne konsistens kan forbedres på hovedområdeniveau ved at tilføje spørgsmål til skalaen. Det må dog afvejes, om det er nødvendigt.

Skalaen har overordnet set fine værdier for hhv. intern konsistens og item-fit, og en konfirmativ faktoranalyse giver ikke anledning til justeringer. Der er heller ingen bemærkelsesværdige forskelle i, hvordan studerende svarer på tværs af baggrundsforhold.

På hovedområdeniveau er den interne konsistens ikke tilstrækkelig på det sundhedsvidenskabelige område på UNI (0,55). Valideringsanalyser af de udvidede skalaer viser, at den interne konsistens på hovedområdeniveau kan øges til et tilfredsstillende niveau for UNI (0,847) ved at tilføje tre spørgsmål.

Uddannelses- og
Forskningsministeriet

De tre spørgsmål, der kan tilføjes, er:

- *"Jeg tror, at jeg kommer til at bruge det, jeg lærer, når jeg har færdiggjort min uddannelse"*
- *"Jeg kan lide emnerne på uddannelsen"*
- *"Jeg tror, at jeg kommer til at bruge det, jeg lærer, senere på uddannelsen"*

Overvejelsen om at tilføje ekstra spørgsmål skal afvejes over for hensynet til længden på det samlede spørgeskema, og hvorvidt det kvalitativt vurderes, at spørgsmålene kan supplere den underliggende skala.

Konstruktiv feedback

Overordnet har skalaen *Konstruktiv feedback* tilfredsstillende intern konsistens, data passer til den teoretiske model og faktorloadings for alle spørgsmål er tilfredsstillende (0,632-0,816). En eksplorativ faktoranalyse af den teoretiske model viser, at ét af skalaens spørgsmål ikke loader på nogen af skalaerne, og at det kan styrke faktorstrukturen at udelade spørgsmålet. Dog risikerer det at gøre den samlede skala mindre robust at udelade spørgsmålet i forhold til fremtidige valideringsanalyser.

Spørgsmålet *"Det vi forventes at lave, hjælper mig til at skabe sammenhæng mellem det, jeg skal lære, og det jeg ved i forvejen"* loader ikke på nogen af de testede dimensioner. En konfirmativ faktoranalyse af en justeret model, hvor spørgsmålet er udeladt, bekræfter at det kan styrke faktorstrukturen at udelade spørgsmålet.

Justeringen vil dog medføre at skalaen kun har tre items, hvilket kan gøre skalaen mere sårbar i det endelige spørgeskema. Der er ikke foretaget skalavalideringer efter denne justering.

Læring og forståelse

Skalaen *Læring og forståelse* har et tilfredsstillende item-fit, og alle spørgsmål har en tilfredsstillende faktorloading. Værdierne for intern konsistens (0,677) er dog under den konventionelle grænseværdi på 0,7, og der er signifikant sektorvariation i besvarelsene. Valideringsanalyser af en udvidet skala viser, at den interne konsistens kan forbedres til tilfredsstillende værdier, hvis der tilføjes to spørgsmål. Dette vil også mindske sektorvariansen, uden at den dog forsvinder helt. Det kan overvejes, om sektorvariationen for det gældende spørgsmål kan forklares kvalitativt.

Rasch-analysen viser signifikant sektorvariation for to af skalaens spørgsmål:

- *"Undervisningen hjælper mig til at forholde mig kritisk til ny viden eller forskning, jeg bliver præsenteret for"*
- *"Uddannelsen har givet mig en fornemmelse af, hvordan fagpersoner arbejder på området (f.eks. forskere, eksperter, fagprofessionelle, praktikere og lign.)"*

Uddannelses- og
Forskningsministeriet

Valideringsanalyser af en udvidet skalaer viser, at skalaens interne konsistens opnår tilfredsstillende værdi ved tilføjelse af to spørgsmål til skalaen. Sektorvariationen mindskes til et ikke-signifikant niveau for det ene af spørgsmålene.

I en Rasch-analyse af de udvidede skalaer er der stadig signifikant sektorvariation på spørgsmålet *"Undervisningen hjælper mig til at forholde mig kritisk til ny viden eller forskning, jeg bliver præsenteret for."* Studerende fra UNI svarer signifikant højere på spørgsmålet end studerende fra PH og EA, hvilket muligvis kan forklares kvalitativt med udgangspunkt i forskelle mellem uddannelsestyperne.

De to spørgsmål, der kan tilføjes, er:

- *"Underviserne lægger vægt på, at vi kan anvende det lærte i en ny kontekst eller på praktiske problemstillinger"*
- *"Underviserne sætter pris på at vi kommer med nye ideer eller perspektiver i forhold til det, vi lærer"*

Hvis de to spørgsmål tilføjes, har skalaen fem spørgsmål. Den interne konsistens når tilfredsstillende værdi, og den signifikante sektorvarians forsvinder for et af to spørgsmål. Det må dog afvejes i forhold til spørgeskemaets samlede længde, og om det kvalitativt vurderes at spørgsmålene kan supplere den underliggende skala.

Overensstemmelse (Alignment)

Overordnet har skalaen *Overensstemmelse* tilfredsstillende intern konsistens, data passer til den teoretiske model, og faktorloadings for alle spørgsmål er tilfredsstillende (0,479-0,783). En eksplorativ faktoranalyse af den teoretiske model viser dog, at skalaen kan styrkes ved at udelade et spørgsmål, der loader lavt på skalaen.

En eksplorativ faktoranalyse af den teoretiske model viser, at spørgsmålet *"Det er tydeligt, hvad der forventes af det arbejde, der skal bedømmes (dvs. opgaver, projekter, eksamener og prøver)"* loader lavt (0,377) på skalaen for *Overensstemmelse*, og loader højere på *Konstruktiv feedback* (0,382). En konfirmativ fak-

toranalyse af en justeret model hvor spørgsmålet er udeladt bekræfter, at det kan styrke faktorstrukturen at udelade spørgsmålet.

Justeringen vil dog medføre, at skalaen kun har tre spørgsmål. Der er ikke foretaget skalavalideringer af de justerede skalaer.

Støtte fra medstuderende

Skalaen *Støtte fra medstuderende* har overordnet tilfredsstillende værdier for item-fit og faktorloading, og den eksplorative faktoranalyse har ikke givet anledning til justeringsforslag. Den interne konsistens for skalaen (0,686) er for lav i forhold til grænseværdien på 0,7, og der kan ses en vis sektorvariation i svarafgivelse. Valideringsanalyser af en udvidet udgave af skalaen viser, at det kan give tilfredsstillende intern konsistens og mindsket sektorvariation at tilføje tre ekstra spørgsmål.

De tre spørgsmål, der kan tilføjes, er:

- *"Jeg samarbejder ofte med andre studerende om at løse opgaver på studiet"*
- *"Jeg bruger ofte tid på at diskutere undervisningsmateriale med andre studerende"*
- *"Jeg prøver ofte at forklare pointer fra undervisningen til andre studerende"*

Uddannelses- og
Forskningsministeriet

Valideringsanalyser af den udvidede skala viser, at skalaen får tilfredsstillende værdi for intern konsistens (0,778) med de tre ekstra spørgsmål. Hvis spørgsmålene tilføjes, har skalaen seks spørgsmål, og der opnås intern konsistens. Det må dog afvejes i forhold til spørgeskemaets samlede længde, og om det kvalitativt vurderes at spørgsmålene kan supplere den underliggende skala.

Underviserinteraktion

Skalaen *Underviserinteraktion* fungerer overordnet fint, men der er forbedringspotentiale for et enkelt spørgsmål. Overordnet har skalaen tilfredsstillende intern konsistens, data passer til den teoretiske model og alle spørgsmål har en tilfredsstillende faktorloading. En eksplorativ faktoranalyse viser dog, at et af spørgsmålene loader lavt på skalaen. På baggrund af resultaterne kan et spørgsmål eventuelt fjernes, og det kan overvejes at tilføje spørgsmål til erstatning. Der er dog ikke foretaget valideringsanalyser af de justerede skalaer.

Rasch-analysen viser en ikke-signifikant sektorvariation i svarangivelserne for spørgsmålet *"Vi har i nogen grad indflydelse på, hvad underviserne lægger vægt på"*. Der er ikke foretaget analyser af sektorvariationen hvis spørgsmålet udelades, men valideringsanalyser af de udvidede skalaer viser, at sektorvariationen mindskes, hvis der tilføjes tre ekstra spørgsmål.

En eksplorativ faktoranalyse viser, at spørgsmålet *"Underviserne hjælper os med at forstå, hvordan man tænker og drager konklusioner på uddannelsen"* loader lavt på skalaen for underviserinteraktion (0,344) og samtidig krydsloader på skalaen for Læring og forståelse (0,328). Dette kan give anledning til at fjerne spørgsmålet.

Det spørgsmål, der kan udelades fra skalaen, er:

- *"Underviserne hjælper os med at forstå, hvordan man tænker og drager konklusioner på uddannelsen"*

De tre spørgsmål, der kan tilføjes, er:

- Underviserne interesserer sig for, om jeg udvikler mig fagligt
- Underviserne er gode til at inddrage de studerende i undervisningen
- Underviserne er tilgængelige for dialog uden for undervisningen

Justeringen vil medføre, at skalaen ændres substantielt. Der er ikke foretaget skavalideringer af de justerede skalaer, og det er derfor svært at vide, om en justering vil påvirke andre aspekter af skalaen. Justeringen må dog afvejes i forhold til spørgeskemaets samlede længde, og om det kvalitativt vurderes at spørgsmålene kan supplere den underliggende skala.

**Uddannelses- og
Forskningsministeriet**

Tabel 2

Mulige justeringer på baggrund af valideringsanalyserne

Skala	Mulig justering	Konsekvens
Interesse og motivation	1-3 Spørgsmål kan tilføjes for at øge intern konsistens på hovedområdeniveau: 1. "Jeg tror, at jeg kommer til at bruge det, jeg lærer, når jeg har færdiggjort min uddannelse" 2. "Jeg kan lide emnerne på uddannelsen" 3. "Jeg tror, at jeg kommer til at bruge det, jeg lærer, senere på uddannelsen"	
Konstruktiv feedback	Et spørgsmål kan fjernes for at styrke faktorstruktur (herefter tre spm i indikatoren): 1. "Det vi forventes at lave, hjælper mig til at skabe sammenhæng mellem det, jeg skal lære, og det jeg ved i forvejen"	
Læring og forståelse	1-2 spørgsmål kan tilføjes for at øge intern konsistens og mindske sektorvariation (herefter op til fem spm i indikatoren): 1. "Undervisningen hjælper mig til at forholde mig kritisk til ny viden eller forskning, jeg bliver præsenteret for" 2. "Uddannelsen har givet mig en fornemmelse af, hvordan fagpersoner arbejder på området (f.eks. forskere, eksperter, fagprofessionelle, praktikere og lign.)"	
Overensstemmelse (Alignment)	Et spørgsmål kan fjernes for at styrke faktorstruktur (herefter har indikatoren tre spm): 1. "Det er tydeligt, hvad der forventes af det arbejde, der skal bedømmes (dvs. opgaver, projekter, eksamener og prøver)"	
Støtte fra medstuderende	1-3 spørgsmål kan tilføjes for at øge intern konsistens og mindske sektorvariation (herefter op til seks spm): 1. "Jeg samarbejder ofte med andre studerende om at løse opgaver på stu-diet" 2. "Jeg bruger ofte tid på at diskutere undervisningsmateriale med andre studerende" 3. "Jeg prøver ofte at forklare pointer fra undervisningen til andre studerende"	
Underviserinteraktion	Et spørgsmål kan fjernes for at styrke faktorstruktur: 1. Underviserne hjælper os med at forstå, hvordan man tænker og drager konklusioner på uddannelsen" Tre spørgsmål kan tilføjes for at mindske sektorvariation og erstatte evt. fjernet spørgsmål: a. Underviserne interesserer sig for, om jeg udvikler mig fagligt b. Underviserne er gode til at inddrage de studerende i undervisningen c. - Underviserne er tilgængelige for dialog uden for undervisningen	

**Uddannelses- og
Forskningsministeriet**

Anm.: Der er ikke foretaget analyser af spørgeskemaet med udgangspunkt i justeringerne.

Kilde: Uddannelses- og Forskningsministeriet

Analysemetoder

Resultaterne fra pilotundersøgelsen er valideret i forskellige analyser.

Målingsvaliditet og intern konsistens

Skalaernes målingsvaliditet og interne konsistens er undersøgt med faktoranalyser af de ni skalaer, der fremgår af den teoretiske model. Analyserne er foretaget af Danmarks Evalueringsinstitut (EVA) og dokumenteret jf. vedlagte bilag. Målingsvaliditeten undersøges gennem item-item korrelation og item-skala korrelation for hver skala, både for de samlede resultater, og på hovedområde- og sektorniveau. Cronbach's alpha bruges til at vurdere skalaernes interne konsistens (reliabilitet).

Grænseværdierne, som EVA anvender til at vurdere, om reliabiliteten er tilfredsstillende, er hentet fra gældende litteratur på området:

Cronbach's alpha-værdi	> 0,70
Item-item-korrelationer	> 0,30
Item-skala-korrelationer	> 0,50

Uddannelses- og
Forskningsministeriet

I Tabel 1 er Cronbach's alpha-værdier over 0,67 kategoriseret som "tilstrækkelig".

I den gennemførte Rasch-analyse undersøges skalaernes interne konsistens også. Resultaterne stemmer overens med EVA's resultater, og de afrapporteres derfor ikke særskilt.

Regressionsanalyser af sammenhæng med frafaldsspørgsmål

EVA har gennemført regressionsanalyser for at belyse sammenhænge mellem LEARN-skalaerne og spørgsmålet om, hvorvidt respondenterne overvejer at stoppe på uddannelsen.

Rasch-analyse

Med en Rasch-analyse testes, om data passer ('fitter') til den teoretiske model. Overordnet er antagelsen bag, at sandsynligheden for at en person giver et bestemt svar på et givent item, er en funktion af forskellen mellem personens niveau på den latente variabel, der måles, og styrken af det givne item.

Med en Rasch-analyse undersøges det blandt andet, om spørgsmålene i de enkelte skalaer kun måler én latent variabel (item-fit), og at baggrundsvARIABLE som køn, sektor og alder ikke betinger vurderingen af den bagvedliggende variabel (DIF - differentiell itemfunktion). I vedlagte bilag uddybes beskrivelsen af Rasch-analysen og resultaterne af den.

Grænseværdierne for item-fit: $p < 0,001$

Grænseværdierne for DIF defineres gennem Bonferronis sandsynlighed, og varierer derfor afhængigt af, hvor mange spørgsmål, der er i skalaen.

For pilotundersøgelsen gælder eksempelvis følgende grænseværdier:

Skalaer med 3 items: $p < 0,0033$

Skalaer med 5 items: $p < 0,0020$

Konfirmative og eksplorative faktoranalyser

UFM har gennemført faktoranalyser af datamaterialet. Faktoranalyserne er foretaget i tre skridt. Først er der gennemført en konfirmativ analyse af de teoretiserede modeller. Herefter er modellen forsøgt forbedret gennem eksplorativ faktoranalyse. Endelig er modellerne testet i justerede versioner.

Alle analyser er foretaget på det fulde sample. Faktoranalysernes validitet vurderes ud fra konventionelle statistiske grænseværdier³:

Standardiserede faktorloadings > 0,40

RMSEA < 0,06

CFI > 0,90

SRMR < 0,08

Uddannelses- og
Forskningsministeriet

³ Jf. Bowen, N., & Guo, S. (2012). *Structural equation modeling*. New York: Oxford University Press