



Universiteit Leiden



Den Haag

Taal als medicijn

Een onderzoek naar de impact van het Nederlands taalvermogen op de fysieke en mentale gezondheid van participanten in het Instapeconomie project van de gemeente Den Haag.

Naam student: Pieter Sluijter (S2958333)

Begeleider: Dr. M.M.F. Collewet

Datum: 30/07/2025

Opleiding: MSc Public Administration: Economics and Governance

Woordenaantal: 18.890

Samenvatting

De taalvaardigheid van individuen is essentieel voor hun participatie aan de samenleving. Internationaal onderzoek laat een trend zien dat een beperkt taalvermogen samenhangt met een verslechterde mentale en fysieke gezondheid. Dit onderzoek richt zich op het begrijpen en verklaren van dit verband, waarbij de Fundamental Cause Theory (FCT) van Link en Phelan (1995) wordt toegepast als theoretisch kader. De FCT stelt dat een fundamentele veroorzaker structureel zorgt voor gezondheidsverschillen doordat het de toegang tot gezondheid bevorderende flexibele middelen beïnvloedt. Door middel van deze theorie wordt er onderzoek gedaan naar het verband dat ontstaat vanuit mechanismes en de inzet van desbetreffende flexibele middelen om een betere gezondheid te verkrijgen. In dit onderzoek wordt de toepasbaarheid van het Nederlandse taalvermogen als fundamentele veroorzaker van gezondheidsverschillen geanalyseerd binnen de context van het Instapeconomie project van de gemeente Den Haag. Door middel van een Structural Equation Model (SEM) worden de assumpties van de FCT getest en gewogen. Uit deze analyse komen de resultaten dat stress en de interpretatie van gezondheidsinformatie prominente mechanismes binnen het gehanteerde model zijn voor het verklaren van het verband tussen taalvermogen en gezondheidsverschillen. Door middel van robuustheidstesten zijn deze verbanden uitgebreid getest, die wijzen naar volatiliteit van het gehanteerde model door het lage aantal participanten aan de enquête, en de mogelijke alternatieve verklaarders voor de effecten. Hierdoor is het bewijs meer ondersteunend van aard, waardoor de waargenomen verbanden ondersteuning voor de hypothesen leveren in plaats van een causaal verband. Desondanks is er bewijs geproduceerd voor het verband tussen en de mogelijke toepasbaarheid van het effect dat taalvermogen op de gezondheid heeft van de participanten om te worden gezien als fundamentele verklaarder in de FCT.

Inhoudsopgave

| | |
|-----------------------------------------------------------------------------|----|
| 1. Inleiding | 3 |
| 2. Theoretisch kader | 5 |
| 2.1 De Fundamental Cause Theory..... | 5 |
| 2.1.1 De Fundamental Cause Theory..... | 5 |
| 2.1.2 Ontwikkelingen van de theorie | 6 |
| 2.2 Taalbeheersing | 8 |
| 2.2.1 Taalbeheersing als concept en de relatie tot gezondheid | 8 |
| 2.2.2 Taalbeheersing als fundamentele oorzaak..... | 9 |
| 2.2.3 Taalbeheersing als fundamentele oorzaak van SES | 11 |
| 2.2.4 Taalbeheersing als fundamentele oorzaak van gezondheidseffecten | 13 |
| 2.3 Hypotheses..... | 18 |
| 3. Het project Instapeconomie | 20 |
| 4. Methodologie | 22 |
| 4.1 Beschrijving dataset | 22 |
| 4.2 Gehanteerde variabelen..... | 23 |
| 4.2.1 Exogene variabelen | 23 |
| 4.2.2 Endogene variabelen | 25 |
| 4.2.3 Mechanismes | 26 |
| 4.3 Methode van analyse | 29 |
| 4.4 Over causaliteit | 33 |
| 4.5 Robuustheidstesten | 35 |
| 5. Analyse | 37 |
| 5.1 Descriptieve analyse..... | 37 |
| 5.2 SEM..... | 40 |
| 5.3 Robuustheidstesten | 45 |
| 5.4 Discussie | 48 |
| 6. Conclusie..... | 51 |
| 7. Literatuurlijst | 53 |
| 8. Bijlage..... | 58 |

1. Inleiding

Taalvaardigheid speelt een centrale rol binnen de functionaliteit van een samenleving. Naast dat taal het belangrijkste middel is voor mondelinge communicatie, is het ook cruciaal voor het delen en opnemen van schriftelijke informatie. Missende Nederlandse taalvaardigheid zorgt er daarom voor dat communiceren in de taal vermoeilijkt wordt, naast dat het interpreteren van Nederlandse informatiestromen vrijwel onmogelijk wordt voor iemand die het niet kan lezen.

Hierom is taalvermogen ook een belangrijke factor binnen het kunnen functioneren in de huidige maatschappij. Een literair laaggeletterde, iemand die niet goed genoeg Nederlands kan lezen om te functioneren, is bijvoorbeeld minder sociaal actief, brengt gemiddeld hogere zorgkosten met zich mee door onder andere slechtere gezondheidsvaardigheden, heeft structureel minder vaak een baan, en als ze een baan hebben betaald deze minder goed dan wanneer ze een beter taalniveau zouden hebben (PwC, 2018). Daarnaast is er duidelijk bewijs dat het spreken van de dominante taal voor migranten een betere mentale en fysieke gezondheid teweeg brengt, naast een betere sociaaleconomische positie in de samenleving (Buisman et al., 2024; Schachter et al., 2012; Clarke & Isphording, 2017).

Om deze verbanden te verklaren, worden aspecten uit de Fundamental Cause Theory (FCT) van Bruce G. Link en Jo Phelan (1995) gebruikt. Deze theorie, die oorspronkelijk het aanhoudende verband tussen sociaaleconomische status (SES) en gezondheidsverschillen verklaart, stelt dat een fundamentele veroorzaker constant zorgt voor deze gezondheidsverschillen doordat het toegang tot flexibel inzetbare middelen reguleert die bij inzet de gezondheid kunnen veranderen. Structureel zorgen fundamentele veroorzakers voor verschillen in de toegang tot deze middelen, waardoor er structureel gezondheidsverschillen tussen groepen veroorzaakt worden. De FCT is ontwikkeld op basis van SES en gezondheidsverschillen, maar legt een raamwerk waardoor andere fundamentele veroorzakers als taalvermogen kunnen worden ontwikkeld, mits de SES niet een te prominente verklaarder wordt voor de beschreven effecten van deze alternatieven.

Hierom is analyse van de FCT in de context van het Instapeconomie project van gemeente Den Haag potentievol. In dit project vervullen uitkeringsgerechtigden die door meervoudige problematiek een grote afstand tot de reguliere arbeidsmarkt hebben bij wijkbedrijven waardevol werk voor de wijk die laagdrempelig genoeg is dat de participanten

het kunnen blijven uitvoeren (Gemeente Den Haag, 2022). Dit project neemt daarmee langdurig uitkeringsgerechtigden aan die weinig uitzicht hadden op een reguliere baan en vaak vervreemd raakten van de samenleving, in een traject onder begeleiding van persoonlijke coaches om re-integratie in de samenleving te stimuleren (Gemeente Den Haag, 2022). Deze zeer diverse groep qua problematiek is een ogenschijnlijk homogene groep qua SES, omdat het gekenmerkt wordt door werkeloosheid en afhankelijkheid van een uitkering. Hierdoor is de introductie van een mogelijke alternatieve fundamentele veroorzaker in de context van deze groep potentievol, waardoor dit onderzoek het taalvermogen aandraagt als mogelijke fundamentele veroorzaker van gezondheidsverschillen. Om dit verband te onderzoeken en bewijs te leveren voor het wezen van taalvermogen als een fundamentele veroorzaker van gezondheidsverschillen is de centrale onderzoeksvraag als volgt:

Wat is het effect van een lagere beheersing van de Nederlandse taal op de gezondheid van participanten aan de Instapeconomie in Den Haag regio Zuidwest en omstreken?

Het analyseren van dit verband zal gebeuren op basis van een Structural Equation Model (SEM) dat indirecte verbanden analyseert die sterk theoretisch onderbouwd zijn. Deze indirecte verbanden zullen worden gekarakteriseerd door mechanismes, die uitgebreid zullen worden toegelicht in het hierop volgende theoretische kader. Voor deze analyse worden de uitkomsten van een jaarlijkse enquête onder de participanten van de Instapeconomie gebruikt.

Het begrijpen van het indirecte effect dat taal heeft op de gezondheid van Nederlandse inwoners is fundamenteel voor een goede allocatie van zorg. Het niet begrijpen van taal, zoals in het geval van laaggeletterden, zorgde in 2012 voor een gemeten 257 miljoen aan extra kosten (PwC, 2018). Deze kosten zijn de kosten die konden worden verweten aan een probleem met taal, waardoor het totale probleem dat niet alleen laaggeletterden betreft rond een verminderd taalvermogen waarschijnlijk substantieel hoger ligt. Dit zijn kosten die kunnen worden vermeden als er meer aandacht aan geletterdheid wordt gegeven en er daardoor praktijken worden gestimuleerd die gezondheid bevorderend zijn. Onderzoek naar waarom een verminderd taalvermogen de gezondheid verslechterd is daarmee maatschappelijk relevant omdat het kan leiden tot beter begrip van het probleem en daarmee maatwerk kan bevorderen.

De wetenschappelijke relevantie is tweezijdig. Ten eerste is er, buiten de literatuur en projecten rond laaggeletterdheid, weinig onderzoek over het effect van het Nederlands taalvermogen op de gezondheid. Daarnaast is het analyseren van het besproken effect in deze onderzoeksgroep uniek, waardoor het de potentie heeft om meer context te bieden aan

problemen die werkloosheid tot gevolg hebben. Ten tweede is dit onderzoek een waardevolle toevoeging aan de FCT omdat het bewijs levert voor een potentiële alternatieve fundamentele veroorzaker. Hiermee draagt het onderzoek bij aan een enorm impactvolle theorie doordat het bewijs levert voor de toepasbaarheid op andere verbanden.

De volgende hoofdstukken geven systematisch een antwoord op de hoofdvraag. In hoofdstuk 2 wordt de theorie achter de FCT gepresenteerd en uitgewerkt, waarna de potentie van het taalvermogen als fundamentele veroorzaker van gezondheidsverschillen en de mogelijke mechanismes daarvoor worden toegelicht. Dan presenteert hoofdstuk 3 de demografie van Den Haag Zuidwest en de context van het Instapeconomie project. Hoofdstuk 4 presenteert dan het gehanteerde SEM model en de variabelen, naast dat het de limitaties van het model in de context van dit onderzoek toelicht, en de gebruikte robuustheidstesten bespreekt. De uitkomsten van deze besproken methodieken worden dan uitgewerkt en bediscussieerd in hoofdstuk 5. Waarna er wordt afgesloten in de conclusie met een antwoord op de onderzoeksvraag en aanbevelingen voor vervolgonderzoek.

2. Theoretisch kader

In het volgende deel wordt de FCT uitgebreid behandeld, gevolgd door het beschrijven van de belangrijke uitbreidingen van de FCT, waarna het gebruik van taalniveau wordt toegelicht in plaats van de SES als fundamentele oorzaak van gezondheidsverschillen.

2.1 De Fundamental Cause Theory

De FCT, bedacht en uitgewerkt door Bruce G. Link en Jo Phelan (1995), is ontwikkeld om de constante associatie tussen SES en gezondheidsverschillen te verklaren. Volgens Link en Phelan worden deze constante verschillen veroorzaakt door het niveau aan flexibel inzetbare *resources*, oftewel middelen, dat een persoon heeft.

2.1.1 De Fundamental Cause Theory.

De correlatie tussen SES en gezondheidsverschillen was tijdens de opkomst van welvaartsstaten voorspeld om te vervagen doordat de levensstandaarden drastisch omhoog gingen, en risicofactoren als slecht sanitair en blootstelling aan ziektes steeds gelijk werd tussen sociale groepen (Kadushin, 1964, 75). Alleen was niks minder waar en bleef de samenhang tussen SES

en gezondheidsverschillen aanhouden, alleen nu door middel van andere mediators binnen de associatie (Link & Phelan, 1995, 86). In elk geanalyseerd geval bleken mensen met een hoger SES beter bestand tegen gezondheidsrisico's dan zij met een lager SES. Hierdoor hebben Link en Phelan (1995) de hypothese gesteld dat deze veranderingen in mediators zullen blijven aanhouden, omdat SES een fundamentele veroorzaker van gezondheidsverschillen is.

Hieruit vloeit de FCT, waarin Link en Phelan (1995) beredeneren dat mensen met een hogere SES de toegang tot én kennis hebben van middelen die kunnen worden gebruikt om risico's te vermijden of minimaliseren waaruit slechte gezondheid voortvloeit. Deze middelen worden geclassificeerd in vijf categorieën: geld, kennis, macht, aanzien, en sociale verbondenheid. De combinatie van deze middelen zorgen ervoor dat een individu beter in staat is om effectievere preventieve maatregelen te nemen, sneller toegang te verkrijgen tot de benodigde zorg en een betere reactie te hebben op gezondheidsbedreigingen, ongeacht de mechanismes van het huidige gezondheidsrisico.

Link en Phelan (1995, 88) stellen meerdere voorwaarden waardoor een bepaald aspect kan worden benoemd als een fundamentele veroorzaker van gezondheidsverschillen. Hierin zijn SES en gezondheidsuitkomsten niet de enige fundamentele oorzaak en afhankelijke variabele mogelijk, maar wel degene die de auteurs gebruikten. Deze voorwaarden stellen dat fundamentele veroorzakers gelinkt zijn aan meerdere gezondheidsuitkomsten, die worden veroorzaakt door meerdere risicofactor mechanismes. Daarnaast moeten er middelen beschikbaar zijn die gezondheidsrisico's kunnen minimaliseren of voorkomen. En tot slot moet de associatie tussen een fundamentele veroorzaker en gezondheid over tijd worden geproduceerd door vervanging van de ingrijpende mechanismes (Phelan et al., 2010, 29).

Kortom, FCT beschrijft aanhoudende relaties tussen een fundamentele oorzaak als SES en gezondheidsverschillen binnen sociale groepen. Hierin wordt de gezondheid op meerdere manieren beïnvloedt door de oorzaak die dit doet door middel van meerdere ingrijpende mechanismes. Deze ingrijpende mechanismes zorgen voor verschillen binnen de SES niveaus door de aanwezigheid van middelen die de gezondheidseffecten van een gezondheidsrisico verminderen, zoals een vaccin of medicijn. De FCT beschrijft de relatie tussen de verdeling van deze middelen, en hoe dit voornamelijk terecht komt bij zij die er genoeg geld, invloed, en kennis van hebben en zo consequent een betere gezondheid vergaren dan zij die dit niet hebben.

2.1.2 Ontwikkelingen van de theorie

De FCT is verder ontwikkeld door zijn auteurs door middel van implementatie van de sterke punten die worden geleverd door hun critici. Twee van deze critici zijn Jeremy Freese en Karen Lutfey (2005, 2010), die stellen dat de literatuur rondom FCT voornamelijk de aanwezigheid van het verband aantonen en niet waarom dit verband bestaat. Ze pleiten er verder voor dat middelen op zich een zeer ambigu concept is, waardoor de relatie tussen SES en gezondheid afhankelijk is van de interpretatie van de aanwezige middelen door de onderzoekers (Freese & Lutfey, 2010, 71). Door de ambiguïteit van de middelen en de missende nadruk op de waarom hebben de onderzoekers het concept van *metamechanismes* ingebracht binnen FCT. Met dit concept indiceren ze dat fundamentele oorzaak door middel van *spillovers*, *habitus en institutional processing* consequent mechanismes blijven produceren waardoor de gezondheid wordt beïnvloedt door SES. Freese en Lutfey (2010, 72-75) formuleren de drie concepten als volgt:

Spillovers ontstaan wanneer individuen vanuit sociale netwerken profiteren van de gezondheid bevorderende handelingen van anderen. Waar bijvoorbeeld door vervuiling of het marketen op vatbare groepen door bedrijven zorgt voor negatieve gezondheidseffecten van de getroffen personen zonder dat er een bewuste keuze voor wordt gemaakt. *Habitus* verwijst naar de onbewuste en automatische neigingen en interpretaties die voortkomen uit iemands sociale positie. Verschillende SES leiden tot uiteenlopende preferenties, zoals het kiezen van goedkoper en natriumrijk voedsel vel bewerkt is boven verse groenten, met als gevolg negatieve gezondheidseffecten. *Institutional processing* betreft de rol die instituties nemen in het toegankelijk maken en de inzet van middelen die een impact hebben op de gezondheid. Hierin geven instituties de gezondheidskeuzes van individuen vorm, waardoor bijvoorbeeld gezondheidsverschillen in groepen ontstaan door verschillen in naleving van beleid of adviezen.

De toevoeging van deze metamechanismen verdiept ons begrip van de interactie tussen de verschillende werkingen binnen FCT. Ze laten zien hoe deze mechanismen elkaar versterken en in stand houden, waardoor sociaaleconomische verschillen consistent bijdragen aan gezondheidsverschillen. Dankzij deze metamechanismen begrijpen we beter hoe en waarom bepaalde gezondheidsresultaten keer op keer terugkeren, ondanks dat de specifieke risicofactoren en contexten veranderen. Bovendien helpt dit kader om te verklaren waarom veranderingen in een enkel mechanisme vaak niet volstaan om de achterliggende structurele ongelijkheden in gezondheid te doorbreken. Kortom, de integratie van deze metamechanismen is cruciaal voor een volledig begrip van FCT, omdat ze een samenhangend beeld schetsen van hoe structurele ongelijkheden hardnekkig blijven bestaan, zelfs als de concrete omstandigheden veranderen.

Een verdere waardevolle toevoeging aan de FCT literatuur is gemaakt door Clouston en Link (2021) om gevolgen van een slechte mentale gezondheid te definiëren door gebruik van het begrip "*Deaths of despair*" (DoD). In DoD wordt er verwezen naar gezondheidseffecten en doden door onder andere suicide, en drugs- en alcoholmisbruik (Clouston & Link, 2021, 15). De beredening hierachter is dat men door het gevoel van wanhoop dodelijk tot zeer schadelijk gedrag voor hun gezondheid gaan vertonen, waarin studies hebben aangetoond dat een lage SES gerelateerd is aan verhoogde psychologische klachten als depressie (zie bijvoorbeeld Lee & Singh (2021) of Reiss (2013)). Hierdoor wordt ook de mentale last van verschillen in SES aangetoond waardoor een nieuwe dimensie aan mechanismes worden aangetoond binnen de theorie.

2.2 Taalbeheersing

Dit onderzoek betoogt dat missende taalbeheersing van de dominante taal binnen een land een fundamentele oorzaak is van gezondheidsverschillen. Door middel van een conceptualisering van taalbeheersing, gevolgd door twee routes waardoor taalbeheersing als een fundamentele oorzaak kan worden gezien, wordt dit stapsgewijs beredeneerd.

2.2.1 Taalbeheersing als concept en de relatie tot gezondheid

Taalbeheersing, of beter gesteld de onbeheersing hiervan, zal binnen dit onderzoek drieledig worden gezien. Ten eerste is er het begrip van taalbeheersing op zich, waarnaast twee specifieke en bekende vormen van taal- onbeheersing worden bekeken. Enerzijds is er een groep die vaak luisterend en sprekend vaardig zijn maar moeite hebben met het lezen en schrijven van de taal. Deze groep wordt gedefinieerd als laaggeletterden, waarvan het lees- en schrijfniveau niet goed genoeg is om volledig te participeren in de samenleving (Stichting Lezen en Schrijven, z.d.) Anderzijds is er een groep anderstaligen die te maken hebben met een aanpassing aan de Nederlandse taal vanuit hun moedertaal. Deze groep is naast laaggeletterd ook onkundig in het spreken van en luisteren naar het Nederlands, waardoor sociale interacties in desbetreffende taal moeilijk zijn.

De uitsplitsing van deze twee groepen is noodzakelijk voor het begrip van de mechanismes die het verband tussen taalbeheersing en gezondheid verklaren. Beide groepen hebben een verlaagd lees- en schrijfvermogen, maar de anderstaligen hebben een sociale

achterstand in de Nederlandssprekende maatschappij ten overstaande van de laaggeletterden. Hierdoor volgt de veronderstelling dat er andere gezondheidseffecten zijn van taal(on)beheersing voor anderstaligen dan laaggeletterden, omdat de anderstaligen meer sociale effecten ondervinden van de taalachterstand door de missende conversatiekunde.

De relatie tussen taalbeheersing en gezondheid is voor beide groepen onderzocht. De laaggeletterde groep omvat ook veel anderstaligen binnen Nederland in de onderzoeken, omdat het binnen het concept iedereen die ouder dan zestien is en geen 2F niveau Nederlands beheerst hieronder valt, waardoor de uitsplitsing uit de literatuur minder kan worden gemaakt. Onderzoeksbureau PwC (2018) heeft in opdracht van Stichting Lezen en Schrijven een literair en financieel onderzoek gestart naar de maatschappelijke kosten van laaggeletterdheid. De Nederlandse laaggeletterden vertonen consequent een ongezondere levensstijl, een hogere kans op hartfalen bij laaggeletterde diabetespatiënten, en een hogere frequentie aan chronische ziekten dan niet-laaggeletterden (PwC, 2018, 22). Dit komt onder andere doordat hun lees- en schrijfproblemen ervoor zorgen dat medicatie verkeerd wordt toegediend en er onzekerheden ontstaan over hun behandelingen, of dat ze minder zelfredzaam zijn (PwC, 2018).

De groep anderstaligen vertonen in de academische literatuur consequent een relatie tussen taal en verschillende gezondheidsvormen, waar taalbeheersing van migranten centraal staat. Schachter et al.(2012) hebben bijvoorbeeld bevonden dat in de VS Aziatische en Latino migranten die vloeiend Engels spreken naast hun eigen taal beter scoren op zelf gerapporteerde fysieke en mentale gezondheid dan immigranten die alleen de oorspronkelijke taal of alleen Engels spreken. Dit wordt bevestigd door Clarke en Ispording (2017), die in Australië een significant negatief effect hebben bevonden tussen het niet beheersen van Engels door immigranten en hun zelf gerapporteerde gezondheid. En in Noorwegen hebben recentelijk Kjøllesdal et al. (2023) bevonden dat een slechtere Noorse taalbeheersing samenhangt met onder andere een slechtere zelf gerapporteerde gezondheid, chronische rug- en nek klachten, slechtere slaap en mentale klachten. Deze drie onderzoeken zijn slechts een greep uit de beschikbare literatuur, maar vertonen wel een eenzijdig beeld in de beoogde relatie van dit onderzoek tussen taalbeheersing en zelf gerapporteerde gezondheid.

2.2.2 Taalbeheersing als fundamentele oorzaak

Eerst een terugblik op de door Link & Phelan (1995, 88) geformuleerde essentiële aspecten van een fundamentele oorzaak, waarin ze stellen dat een fundamentele oorzaak de verdeling van middelen betreft die bepalen hoe capabel mensen zijn om risico's te vermijden voor morbiditeit

en sterfte. Hierin zijn de middelen sociaal en economisch van aard en kunnen ze op meerdere manieren ingezet worden in veelal verschillende situaties en ziektebeelden, waardoor fundamentele oorzaken effect hebben op de gezondheid zelfs als de risicofactoren drastisch veranderen. Aanvankelijk werd SES binnen de FCT gebruikt als introductie van de fundamentele oorzaken van gezondheidsverschillen, waarbij de auteurs deze relatie onderbouwden met toenemend empirisch bewijs.¹ Maar zoals eerder vermeld is de FCT het verband tussen een specifieke exogene variabele en een set van endogene variabelen, waardoor een verband ook gelegd kan worden tussen andere exogene variabelen en endogene variabelen.

Phelan en Link (2015) hebben twee routes geformuleerd, gebaseerd op het beargumenteren van racisme als een fundamentele oorzaak, waardoor een exogene variabele een fundamentele oorzaak van een set van endogene variabelen is. Deze routes zullen worden beschreven en uitgewerkt met de beoogde exogene variabele van dit onderzoek: taalbeheersing. De eerste route is het beschrijven van taalbeheersing als een fundamentele oorzaak van SES. Deze route beredeneert dat SES een fundamentele oorzaak is van gezondheidsverschillen, zoals de eerdere werken van de auteurs hebben laten zien, en daardoor als mediator verklaart dat taalbeheersing een fundamentele oorzaak van gezondheid is. De tweede route beschrijft een direct verband tussen taalbeheersing en gezondheidsuitkomsten, zonder de tussenstap van SES dus.

Binnen de groepen laaggeletterden en anderstaligen verdient opleiding aandacht als alternatieve verklaring voor sociaal-maatschappelijke uitkomsten. Opleidingsniveau is inderdaad een sterke voorspeller van arbeidsmarktpositie, integratie en gezondheid (CBS, 2020 & 2024; de Vries & Lautenbach, 2023; Van der Heide & Rademakers, 2015; Willems et al., 2022; Nivel, 2024). Tegelijkertijd tonen dezelfde studies aan dat taalbeheersing onafhankelijk bijdraagt aan deze uitkomsten. De relatie tussen taalvaardigheid en opleiding wordt vaak als complementair beschouwd: beperkte taal- én opleidingsniveaus versterken elkaar in het reduceren van arbeidsparticipatie en gezondheidsvaardigheden.

Vanuit een causale benadering valt echter te verdedigen dat beperkte taalbeheersing voorafgaat aan lager onderwijsniveau. Laaggeletterden met Nederlands als moedertaal ondervinden beperkingen in het formeel leren, wat hun opleidingskansen verkleint. Bij anderstaligen hangt de beheersing van het Nederlands sterk samen met eerder verworven onderwijsniveau, maar ook hier kan de taalvaardigheid de toegang tot vervolgonderwijs

¹ Dit bewijs wordt uitgebreid gepresenteerd in de uiteenzettingen van Phelan et al. (2010) en Clouston & Link (2021) gebaseerd op een selectie van onderzoeken afkomstig uit de ruim 8000 citaties op het originele werk over FCTs van Link & Phelan (1995)

belemmeren (de Vries & Lautenbach, 2023). Hoewel cognitieve capaciteiten een rol spelen, is dit concept te abstract om de relatie tussen taalvaardigheid, opleiding en uitkomsten adequaat te verklaren. Gezien de onafhankelijke effecten van taalvaardigheid op sociaaleconomische en gezondheidsuitkomsten, is taalbeheersing op zichzelf te beschouwen als een fundamentele veroorzaker.

2.2.3 Taalbeheersing als fundamentele oorzaak van SES

Betreffend de eerste route, er zijn aantoonbare verschillen binnen componenten die samenkomen in SES in Nederland. Nederlandse inwoners met het laagste taalniveau hadden bijvoorbeeld in 2023 ruwweg een 20%-punt lagere arbeidsmarktparticipatie dan zij met het hoogste taalniveau (Buisman et al., 2024, 69). Daarnaast hebben Nederlanders met de laagste taalvaardigheden ongeveer 10%-punt minder vaste arbeidscontracten dan zij met de hoogste taalvaardigheden, wat indiceert dat deze groep andere contractvormen heeft dan de meer taalbeheerste groep waarin vaak een andere vorm van aanstelling aanwezig is (Buisman et al., 2024, 73). Afsluitend zijn Nederlanders met de laagste taalvaardigheden significant minder in leidinggevende posities, zijn er meer werkeloos geweest in de laatste vijf jaar, en zijn ze veel meer aanwezig in lagere inkomens kwintielen dan alle andere taalniveaus (Buisman et al., 2024, 74-78). Kortom, in Nederland hebben mensen met een lagere taalbeheersing slechtere financiële uitkomsten dan mensen met een hogere taalbeheersing en op dat vlak een lagere SES.

Met betrekking tot de sociale aspecten van SES blijkt dat scholieren met een lage taalbeheersing in 2022 nauwelijks doorstromen naar hogere niveaus van het middelbaar onderwijs, terwijl scholieren met een goede taalbeheersing juist vrijwel uitsluitend in die richting afstromen (Inspectie van het Onderwijs, 2023). Verder laten migranten in Nederland een verband zien waarin taalbeheersing samenhangt met meer sociale contacten met Nederlandssprekenden, een sterker gevoel van verbondenheid met de Nederlandse samenleving en meer tevredenheid in hun leven (Het Kenniscentrum Arbeidsmigranten, 2023; Rijk & Lubbers, 2020; Schmeets & Cornips, 2023). Deze effecten zijn ook longitudinaal bevestigd: Chiswick en Wang (2019) vonden op basis van gegevens uit de periode 1991–2002 dat beheersing van de Nederlandse taal significant bijdraagt aan betere arbeidsmarktitkomsten en hogere inkomens, zelfs wanneer gecontroleerd wordt voor opleidingsniveau. Verder hebben de auteurs bevonden dat meer sociale contacten met Nederlandssprekenden samenhangt met betere arbeidsuitkomsten, wat slaat op een verband tussen Nederlandse taalbeheersing en zo betere connecties met SES.

Hierin wordt het effect van taalbeheersing op SES uitkomsten beschreven aan de hand van *Human Capital (HC)* assumpties. Hierin wordt taalbeheersing als een van de aspecten van HC gezien doordat het alle drie de vereisten ervan vervult, zoals beschreven in Chiswick en Miller (2001). Ten eerste wordt de productiviteit van individuen vergroot en de consumptie ontwikkeld door transactiekosten verlaagd als de taalbeheersing toeneemt. Wanneer een individu een betere taalbeheersing verkrijgt, worden kosten van communicatie lager en zo de uitkomsten positiever. Daarnaast zorgt een betere taalbeheersing ook voor een hogere productiviteit bij de andere aspecten van HC, doordat de productiviteit van individuen met een lagere taalbeheersing in een land altijd lager ligt dan die van een individu met een hogere taalbeheersing als de andere HC componenten hetzelfde zijn (Chiswick & Miller, 2001, 392). Ten tweede is de taalbeheersing een aspect van een persoon en kan het niet los worden gezien van die persoon. Ten derde wordt taalbeheersing verbeterd ten koste van andere materialen, want voor het verbeteren van de taal worden tijd en andere middelen ingeleverd door de lerende persoon en zij die investeren voor de verbetering van desbetreffende persoon. Doordat alle drie de vereisten van HC worden vervuld zien we vanaf nu taalbeheersing als een component van HC. Een hogere HC hangt samen met een hogere productiviteit, wat een individu aantrekkelijker maakt voor werkgevers en zo betere arbeidsuitkomsten geeft.

Taalbeheersing presenteert zo zichzelf als een fundamentele veroorzaker van SES omdat het flexibel inzetbare middelen presenteert die kunnen worden ingezet om de SES te bevorderen. Het bevat meerdere structurele factoren die betrekking hebben op meerdere uitkomsten waaruit de SES tot stand komt, waaronder het positieve verband tussen taalbeheersing en het hebben van werk, de betaling van dit werk, en de algehele satisfactie van dit werk die worden veroorzaakt door de componenten gesteld binnen de HC theorie. Verder zorgt taalbeheersing voor verschillen in inzetbare middelen, aangetoond door bijvoorbeeld het verband tussen de interactieterm van taalbeheersing met sociale Nederlandssprekende contacten en de arbeidsuitkomsten van individuen. Daarnaast is het argument te maken dat alle middelen makkelijker te vergaren en in te zetten bij een betere taalbeheersing, doordat communicatiekosten lager worden bij het vergaren van beheersing in de dominante taal van een samenleving. Verder is er het feit dat de Nederlandse taal wordt geprefereerd binnen de Nederlandse samenleving, waardoor het beheersen van Nederlands een voordeel oplevert over mensen van een lager niveau wanneer alle andere aspecten hetzelfde blijven.

Afsluitend is er ook een vergelijkbaar patroon en rationale in HC bevonden die de associatie tussen taalbeheersing en SES over de tijd beschrijft. Hierin wordt bijvoorbeeld door laaggeletterdheid en onbeheersing in leesvaardigheid een slechtere adaptatie bewerkstelligt tot

de steeds meer digitaliserende samenleving, wat zo een nieuw mechanisme aantoont in de relatie tussen taalbeheersing en SES (Non et al., 2021). Deze aspecten tonen bewijs voor de bewering dat taalbeheersing een fundamentele oorzaak van SES is, en zo ook van gezondheidsuitkomsten sinds SES een fundamentele oorzaak daarvan is.

2.2.4 Taalbeheersing als fundamentele oorzaak van gezondheidseffecten

De tweede route beschrijft een direct verband tussen taalbeheersing en gezondheidsuitkomsten, zonder de tussenstap van SES dus. Buisman et al. (2024, 92) tonen aan dat Nederlanders in 2023 met het laagste taalniveau bijna drie keer zo vaak aangeven dat ze een matige tot slechte gezondheid hebben dan hun medebewoners met het hoogste taalniveau. Verder is de literatuur vermeld in alinea 2.2.1 ook vrij eentonig, en laat het voor laaggeletterden én anderstaligen het beeld zien dat het niet beheersen van de dominante taal in een land leidt tot slechtere gezondheidsuitkomsten.

Om deze relatie te kunnen beschrijven als fundamentele oorzaak, moeten de gestelde voorwaarden van Link en Phelan (1995) worden voldaan. Ten eerste is het noodzakelijk om te beschrijven wat voor soort flexibele middelen kunnen worden ingezet om de gezondheidsuitkomsten te verbeteren. Hierin is het belangrijk om te kunnen beschrijven dat deze flexibele middelen vergroten naar mate de taalbeheersing beter wordt. Daarnaast moeten deze flexibele middelen los staan van SES. En ten tweede moeten deze flexibele middelen uitkomen in ongelijkheden in meerdere gezondheidsuitkomsten, omdat de flexibiliteit van de middelen relevant zijn voor veel uiteenlopende situaties en zo de gezondheid kunnen beïnvloeden. Uit de vijf middelen die Link en Phelan (1995) aanduiden als kenmerkend voor fundamentele oorzaken van gezondheidsuitkomsten, blijken in het geval van taalbeheersing met name drie middelen relevant: aanzien, sociale contacten en kennis.

Aanzien, ofwel persoonlijke status, wordt door Weber beschreven als de “eer of eerbied die verbonden is aan een persoon of sociale status en ongelijk verdeeld is binnen een sociale groep” (Phelan & Link, 2015, 317). Dit flexibel middel wordt al veel beschreven in de relatie tussen taalbeheersing en SES door de link met de arbeidsmarkt, maar bevat ook sociale middelen die daarbuiten vallen. Voor laaggeletterden is er bijvoorbeeld een bepaald taboe gevormd omtrent het missen van de basis taalvaardigheden waardoor ze door schaamte het minder vertellen aan naasten, terwijl een significant deel van de Nederlandse bevolking laaggeletterd is (Stichting Lezen & Schrijven et al., 2014). Dit wordt beaamd in andere delen

van de wereld, waar in de VS was bevonden onder Afrikaans-Amerikanen dat schaamte voor laaggeletterdheid, waar 67.2% van de laaggeletterden met lagere gezondheidsvaardigheden het zelfs nooit aan de partner heeft verteld (Parikh et al., 1996). Hierdoor wordt herkenning van laaggeletterdheid en eventuele verbetering van de taalbeheersing geremd door het lage persoonlijke aanzien ervan, waardoor de positieve gevolgen van taalverbetering uitblijven.

Een slechte taalbeheersing door anderstaligheid of spreken met een accent heeft ook een effect op het aanzien van een persoon voor anderen. Zo constateerden Dekker, Duarte en Loerts (2021) dat kinderen al op jonge leeftijd een voorkeur ontwikkelen voor Nederlands en Engels boven zowel het Marokkaans als Nederlands dat wordt gesproken met een Marokkaans accent. Verder is er bevonden door Pew Research Center (2024) dat 96% van de Nederlanders de beheersing van de Nederlandse taal het meest belangrijke hoofdbestanddeel van de Nederlandse identiteit vindt. Verder is er in de Nederlandse politiek ook veel nadruk gelegd op Nederlandse taalbeheersing bij de integratie van migranten. Hieruit is te halen dat de Nederlandse bevolking het spreken van de taal hoger in aanzien heeft, waardoor een verlaagd aanzien samenhangt met het niet spreken of het spreken met een accent van Nederlands. Op basis van de beschreven verschillen in aanzien, die voortkomen uit de schaamte bij laaggeletterden en de koppeling tussen de Nederlandse taal en identiteit, kan worden gesteld dat aanzien een flexibel inzetbaar middel is in de relatie tussen taalbeheersing en gezondheidsuitkomsten. Mensen met veel aanzien kunnen zorgen voor individuele en groepsmatige voordelen in de gezondheid, door bijvoorbeeld meer attentie naar een probleem te vestigen.

Het middel **sociale contacten** betreft de aanwezigheid van anderen in de omgeving die kunnen bijdragen aan de gezondheid. In onderzoek naar laaggeletterden is bevonden dat er minder vrijwilligerswerk wordt gedaan door deze groep en dat ze vaker in een wijk wonen met een lagere sociale status in tegenstelling tot geletterden (Christoffels et al., 2016). Het wonen in een wijk met een lage sociale status blijkt geassocieerd te zijn met lagere onderwijsniveaus en een verhoogde kans op voortijdig schoolverlaten onder jongeren uit deze wijken (Van de Looij-Jansen et al., 2014). Verder is er een grotere kans op ernstig norm overschrijdend gedrag, wat vaak uitmondt in criminaliteit doordat er vaak slechtere sociale voorzieningen aanwezig zijn die de kans op een vergrijp verkleinen (Van de Looij-Jansen et al., 2014). Daarnaast is bevonden dat arme en niet-arme laaggeletterden lager scoren in het sociaal vertrouwen in mensen dan arme niet-laaggeletterden (Christoffels et al., 2016, 33). Dit laatste toont zo ook de openheid naar buitenstaanders aan, en zo mogelijke contacten. Afsluitend is er bevonden dat anderstaligen ook significant minder sociaal kapitaal ondervinden, wat indiceert dat ze over het algemeen minder sociale contacten hebben in de omgeving (Schmeets & Cornips, 2023). Deze

lagere hoeveelheid aan sociale contacten zorgen er ook voor dat ze minder hulp en incentief krijgen om te assimileren, wat sociale interactie vermindert (Wang et al., 2018, 78). Sociale interactie is hierbij een klassiek mechanisme gebruikt in de FCT door Link en Phelan (1995), waarin het hebben van minder mensen die je kunnen helpen of eventuele potentiële positieve spillovers bewerkstelligen ervoor zorgt dat er een slechtere gezondheid tot stand komt.

En afsluitend het middel **kennis**, wat de relatie tussen gezondheidsvaardigheden en taal omvat. Binnen dit kader biedt de *Health Literacy Theory* waardevolle inzichten. Deze theorie definieert gezondheidsvaardigheden als “de cognitieve en sociale vaardigheden, evenals de capaciteiten van een individu om toegang te krijgen tot gezondheidsinformatie en dan te begrijpen, en daarna te gebruiken op manieren die een goede gezondheid bevorderen en behouden” als betekenis heeft (Kickbusch & Maag, 2008, 205). Hierin komen deze vaardigheden en capaciteiten samen in de gezondheidsvaardigheden van een individu, waarin deze vaardigheden door de World Health Organization (WHO) als een significantere determinant van gezondheid dan SES wordt beschreven (Rademakers, 2014). In een veelal geletterde wereld van medicatie en gezondheid bevorderende handelingen is het begrijpen van de gesproken taal in een land fundamenteel voor een adequate uitvoering daarvan.

Dat laaggeletterden en anderstaligen hierdoor een achterstand opbouwen op zij die het Nederlands volledig begrijpen is daarmee ook begrijpelijk. Ten eerste zijn laaggeletterden vaak lager opgeleid, wat twee hoofddeterminanten zijn van een lagere score van gezondheidsvaardigheden (Fransen et al., 2014). Ten tweede brengen mensen met een lagere score op gezondheidsvaardigheden aanzienlijk vaker een bezoek aan de huisarts en het ziekenhuis dan mensen met een hogere score. (PwC, 2018, 23). Dit verband wordt toegeschreven aan meerdere aspecten, laaggeletterden zijn vaak onzekerder over hun behandeling, ze zijn minder zelfredzaam, er is minder gebruik van preventieve zorg onder de groepen, ze hebben minder gezonde levensstijlen, en er ontstaan veel misopvattingen over medicijngebruik die mogelijk ernstige gezondheidsgevolgen hebben door de moeite met het lezen (PwC, 2018, 22). Voor anderstaligen is dit effect vergelijkbaar, de inputs van het gezondheidsstelsel zijn minder sterk door het taalverschil waardoor de gezondheid een negatief gevolg ondervindt van anderstaligheid (Clarke & Isphording, 2017). Doordat de laaggeletterdheid en anderstaligheid zorgen voor verminderde effecten van de inputs van het gezondheidsstelsel, wordt kennis hier als een flexibel middel gezien omdat deze groepen minder ervan opnemen en zo zorgen voor slechtere gezondheden.

De flexibel inzetbare middel zijn hierbij duidelijk en kunnen worden getransleerd in mechanismes. Meerdere van de volgende mechanismes zijn al aan het licht gekomen in eerdere stukken maar zullen hier expliciet worden benoemd.

Het eerste mechanisme is **schaamte**, zoals eerder benoemd is er een stigma rond laaggeletterdheid waardoor veel mensen met een lagere taalbeheersing geen hulp vragen bij lezen en de taal. Door deze schaamte wordt vaak de laaggeletterdheid niet benoemd bij bijvoorbeeld een dokter, waardoor de dokters ‘dure woorden’ blijven gebruiken in het consult (Easton et al., 2013). Hierdoor ontstaat onder de *Communication Accommodation Theory* (CAT) sociale afstand tussen de laaggeletterde en de dokter wat zo bijdraagt aan verminderde interesse en een lagere informatieopname dan bij makkelijker taalgebruik (Zhang & Giles, 2018). Deze verminderde informatieopname zorgt voor verminderde gezondheidsvaardigheden omdat de aangeraden zelfzorg naar aanleiding van het consult minder goed wordt geïnterpreteerd.

Stress is het tweede mechanisme wat alom is geassocieerd met veel verschillende ziekten en negatieve gezondheidsuitkomsten onder alle bevolkingsgroepen bij langdurig hogere waarden van stress (Dhabhar, 2014). Dit mechanisme is aanwezig bij laaggeletterden en anderstaligen omdat beiden te maken krijgen met onbegrijpelijke taal wat zo zorgt voor persoonlijke problemen die leiden tot meer stress (Zhang & Giles, 2018).

Het derde mechanisme betreft **stress** ontstaan uit **marginalisatie**. Onder de Minority Stress Theory (MST) wordt gesteld dat de mentale gezondheid lijdt onder gevoelens van uitsluiting door de dominante groep in een land (Frost & Meyer, 2023). Voor anderstaligen is deze relatie vrij direct en meer gebaseerd op discriminatie door anderstaligheid, omdat het niet spreken van de Nederlandse taal een signaal is dat een persoon niet in Nederland is opgegroeid en zo behoort tot een minderheid. Voor laaggeletterden is het een meer gecompliceerde relatie die gebaseerd is op de zelfafzondering door de schaamte van het niet kunnen lezen en het lage vertrouwen in de medemens (PwC, 2018). Hierdoor marginaliseren laaggeletterden zich sneller en zonderen ze zich af, wat een van de padstellingen is waaronder volgens de MST stress wordt gegenereerd en zo zorgt voor hogere stressniveaus en de gevolgen van dien (Frost & Meyer, 2023).

Inactiviteit wordt gebruikt als het mechanisme dat de politieke, maatschappelijke en sportieve participatie dat is opgemerkt bij laaggeletterden benoemd (PwC, 2018). Door de afstand tussen de informatie die wordt geleverd vanuit de politiek, verenigingen en sportaccommodaties, en de informatie die kan worden geïnterpreteerd door de laaggeletterden en anderstaligen wordt de participatie daaraan ontmoedigd. De politiek wordt bijvoorbeeld veel

schriftelijk behandeld en bevat debatten met hogere intensiteit waardoor een lagere taalbeheersing daarin ontmoedigend kan werken, wat overeenkomt met de gedachten van de CAT. Daarnaast is taal een horde voor vrijwilligerswerk omdat er enerzijds kennis moet zijn van deze banen wat veelal eerst virtueel gebeurt, en daarnaast zijn veel werkzaamheden binnen deze functies geletterd qua aard. En sporten is ontmoedigend omdat mensen met een lager taalniveau enerzijds over het algemeen armer zijn en daarom meer problemen met betalen hebben, en anderzijds geen basis instructies binnen sportscholen begrijpen en zich vaker verloren voelen in onbekende sportaccommodaties (Wieland et al., 2015). Deze drie vormen van inactiviteit zorgen voor meerdere gezondheidseffecten, omdat ze samen een lagere sociale participatie laten zien wat samenhangt met een lagere fysieke activiteit (Torres et al., 2024, 802). En dit zorgt voor een slechtere gezondheid voor mensen met een lager taalniveau omdat fysieke activiteit een zeer belangrijke determinant is van de gezondheid.²

Afsluitend is de **interpretatie** van **gezondheid inputs** belangrijk om gezondheidsverschillen door taalbeheersing te verklaren. Zoals Clarke en Ispording (2017) beschrijven, zorgt taalbeheersing voor minder transactiekosten door communicatie barrières en maakt daarmee de efficiëntie van gezondheid inputs vanuit meerdere kanten effectiever. Onder andere door deze communicatie barrières is er zoals benoemd meer miscommunicatie met laaggeletterden bij dokters, waardoor ze minder vertrouwen hebben in de behandeling en vaker naar het ziekenhuis moeten (PwC, 2018). Voor medicatie zorgt dit zoals benoemd voor meer complicaties uit verkeerd gebruik, zoals bij verkeerde toepassing van medicatie bij obesitaspatiënten en de volgende orgaanproblemen (PwC, 2018). Daarnaast is er door de overheid ook een actor die gezondheid bevorderend gedrag aanstuurt, wat met moeilijk taalgebruik of geschreven campagnes kan zorgen voor gezondheidsverschillen door misinterpretatie onder laaggeletterden en anderstaligen. Motivaction (2024), een onderzoeksbureau voor onder andere de Nederlandse overheid, heeft bijvoorbeeld bevonden dat tijdens de corona pandemie veel misinformatie ontstond door de gebruikte complexe en juridische taal, waardoor bijvoorbeeld de uitspraak ‘Samen alleen’ voor veel verwarring zorgde. En dan is het vanzelfsprekend dat dit vaker gebeurt dan in dit geval alleen, omdat de overheid in Nederland veel stimulans probeert af te geven omtrent de gezondheid.

De mechanismes en middelen buiten die afkomstig uit SES komen samen in het beeld gebaseerd op de voorwaarden van Link en Phelan (1995) dat taalbeheersing een fundamentele veroorzaker is van gezondheidsverschillen. De middelen en mechanismes tonen aan dat ze een

² Geen citaat omdat dit algemene kennis is.

effect hebben op meerdere gezondheidsuitkomsten, omdat ze flexibel inzetbaar zijn in meerdere situaties die gezondheidsrisicos op een slechtere gezondheid verminderen of vergroten. Daarnaast zijn er meerdere risicofactoren, zoals beschreven in de mechanismes omdat er veel aspecten zijn waardoor de mechanismes de gezondheid beïnvloeden. Ter illustratie, de mechanismes tonen enerzijds effecten op persoonlijke blokkades tot het zoeken van hulp door ervaren schaamte (habitus), terwijl taal ook de interpretatie van informatie vermindert (vermindert dus de institutional processing doordat het veelal professionals zijn die de informatie aanvoeren), en zelfs misinformatie gecreëerd door anderen veel voorkomt bij mensen met lagere taalvaardigheden.

Hierdoor worden alle metamechanismes beschreven, waar het gedrag afkomstig uit schaamte gezondheidsverschillen door habitus veroorzaken, terwijl verminderde interpretatie van professionals een vermindert effect van institutional processing aantoont, en misinformatie vanuit anderen een spillover creëert omdat de gezondheid bevorderende informatie er mee verwatert en zo ongelijke verschillen aanwakkert. Verder zorgen zoals benoemd de mechanismes tot het verminderen van of veroorzaken van risicofactoren die gezondheidsuitkomsten beïnvloeden. En afsluitend zijn de mechanismes zodanig sociaal van aard dat het zich in meerdere situaties kan repliceren, waar bijvoorbeeld vroeger informatie voornamelijk werd verspreid door kranten wat zorgt voor problemen voor laaggeletterden terwijl nu enorm veel gezondheid bevorderende gedragingen worden gedeeld die soms conflicterend zijn en kunnen zorgen voor overweldiging van de ontvangers.

Kortom, de mechanismes vanuit taalbeheersing op de gezondheid zijn gelinkt aan meerdere gezondheidsuitkomsten en meerdere risicofactor mechanismes, waarin flexibele middelen te vinden zijn die risicos kunnen minimaliseren of voorkomen, en dat deze mechanismes over de tijd veranderen, wat alle door Link en Phelan (1995) gestelde voorwaarden zijn voor het werken als een fundamentele veroorzaker van gezondheid.

2.3 Hypotheses

Vanuit de theorie zijn er meerdere hypothesen ontstaan die antwoord bieden op de onderzoeksvraag: *Wat is het effect van een lagere beheersing van de Nederlandse taal op de gezondheid van participanten aan de instapeconomie in Den Haag regio Zuidwest en omstreken?*

Binnen deze onderzoeksvraag was de uitsplitsing tussen anderstaligen en laaggeletterden belangrijk. Hierin is gesteld dat er een verschil in effecten wordt verwacht

tussen deze groepen, omdat de redenatie luidt dat anderstaligen andere sociale effecten zullen ervaren dan laaggeletterden omdat anderstaligen de taal niet kunnen spreken terwijl laaggeletterden met een Nederlandse achtergrond dat wel kunnen. Hierdoor duidt de eerste hypothese het verschil in gezondheidseffecten voorspelt door de verschillen in de mate van taalbeheersing.

De verdere hypothesen hebben meer betrekking op de FCT en de daaruit gevolgde veronderstellingen die zijn gemaakt in de benoemde theorie. Deze hypothesen betreffen de twee routes waardoor taalbeheersing de gezondheidsverschillen verklaren, indirect door de link met SES en direct door middel van de geformuleerde mechanismes. Daarnaast zijn de assumpties van FCT een hypothese op zich, waardoor hier ook hypothesen uit volgen. Hierom zijn de hypothesen als volgt:

1. Het effect van anderstaligheid op de gezondheid van participanten verschilt van dat van laaggeletterdheid.
2. Er is een negatief effect van slechtere taalbeheersing op de gezondheid van participanten, wat wordt gemedieerd door SES.
3. Er is een negatief effect van slechtere taalbeheersing op de gezondheid van participanten, wat wordt gemedieerd door de ervaren schaamte omtrent de onbeheersing van de taal.
4. Er is een negatief effect van slechtere taalbeheersing op de gezondheid van participanten, wat wordt gemedieerd door de ervaren stress.
5. Er is een negatief effect van slechtere taalbeheersing op de gezondheid van participanten, wat wordt gemedieerd door marginalisatie.
6. Er is een negatief effect van slechtere taalbeheersing op de gezondheid van participanten, wat wordt gemedieerd door inactiviteit.
7. Er is een negatief effect van slechtere taalbeheersing op de gezondheid van participanten, wat wordt gemedieerd door het interpretatievermogen van gezondheid outputs.
8. Er zijn meerdere mechanismes waardoorheen taalbeheersing de gezondheidsverschillen verklaart.
9. De taalbeheersing hangt samen met meerdere gezondheidsfactoren door meerdere routes.

3. Het project Instapeconomie

Den Haag Zuidwest wordt gekenmerkt door een stapeling aan sociaaleconomische uitdagingen. Zo ligt het percentage bijstandsontvangers met 15% ruim boven het stedelijk gemiddelde van 7% (Nationaal Programma Den Haag Zuidwest, 2023). De wijk telt verder een hoger aandeel laagopgeleiden (46% tegenover 30%), een lager gemiddeld inkomen (€19.500 tegenover €27.500), en meer chronisch zieken (30% tegenover 20%) dan de rest van Den Haag. Ook op sociaal-maatschappelijk vlak scoort de regio slechter door bijvoorbeeld het gevoel van onveiligheid dat er hoger ligt (9% tegenover 5%), meer bewoners ervaren criminaliteit (23% tegenover 18%), en de verkiezingsopkomst is aanzienlijk lager (29% tegenover 43%).

Om deze problematiek aan te pakken, is het Nationaal Programma Den Haag Zuidwest opgezet, wat een twintigjarig samenwerkingsprogramma is met als doel de cijfers van de gezondheid, educatie, veiligheid, en inkomens van de regio op het Haagse gemiddelde te brengen (Nationaal Programma Den Haag Zuidwest, 2023). Binnen deze bredere aanpak is het project Instapeconomie opgestart, wat gericht is op uitkeringsgerechtigden met een langdurige afstand tot de arbeidsmarkt. Deze doelgroep bestaat voornamelijk uit mensen die langdurig in de bijstand zitten en door meervoudige problematiek nauwelijks aansluiting vinden bij regulier werk en daardoor een verminderde maatschappelijke participatie hebben (Gemeente Den Haag, 2022).

De Instapeconomie biedt via wijkbedrijven laagdrempelige, maatschappelijke werkplekken aan die zowel aansluiten bij de persoonlijke capaciteiten van de deelnemers en ook waardevol zijn voor de wijk. Werkzaamheden waaraan je hierbij moet denken zijn soep koken voor ouderen en het bezorgen hiervan, het uitvoeren van afvalrondes door de wijk, het repareren van fietsen en het openen van een wijkrestaurant. Door deze inzet ontstaat enerzijds waarde voor de wijk doordat de werkzaamheden waarde meebrengen voor de omgeving, en anderzijds persoonlijke vooruitgang voor de participanten op gebieden als taalvaardigheid, zingeving in het dagelijks leven, sociale verbondenheid, en mentale, fysieke en financiële gezondheid (Gemeente Den Haag, 2022)

Het project sluit bovendien aan bij het bredere arbeidsmarktbeleid van gemeente Den Haag, dat zich in deze tijden van krapte richt op activering van onbenutte potentiële arbeiders.

De ambitie van de gemeente hierin is om iedereen perspectief te bieden op duurzame bestaanszekerheid (Gemeente Den Haag, 2023). De Instapeconomie draagt hieraan bij doordat het inactieve inwoners via laagdrempelig werk in de wijkbedrijven betreft bij de samenleving, wat in sommige gevallen zelfs leidt tot doorstroming naar regulier werk.

De werving van de deelnemers aan de Instapeconomie wordt gedaan onder langdurig bijstandsgerechtigden voor wie de stap naar betaald werk onhaalbaar lijkt. Binnen het project krijgen zij persoonlijke begeleiding van medewerkers van de wijkbedrijven, waarbij het werken niet alleen de focus krijgt, maar ook ondersteuning bij praktische problemen wordt geleverd. Hierbij kun je denken aan het helpen bij de aanvraag van een DigiD of het aanpakken van financiële zorgen van de deelnemers.

De werkzaamheden die de participanten leveren worden beloond met een vrijwilligerstoelage die is gebaseerd op het aantal uren dat een participant maakt, wat deelnemers motiveert tot het werken van meer uren. Deze vorm van beloning is bewust gekozen, omdat een vrijwilligerstoelage complementair is aan een bijstandsuitkering terwijl een normaal loon supplementair is door de wetten rondom bijstand. Een loon zou in de praktijk kunnen zorgen voor een inkomensverlies en ook hogere kosten bij de wijkbedrijven. Ter illustratie, als een ontvanger van bijstand inkomsten verkrijgt vanuit loon wordt de bijstand met 100% van het verkregen loon gekort (Steunpunt Uitkeringen, z.d.). Hierdoor loont het maken van een lager aantal uren, wat het geval is in het Instapeconomie project, niet en wordt er in de bijstand gekort. Daarnaast kan bij teveel looninkomsten de inkomsten uit toeslagen wegvallen, wat ook de netto inkomsten verlaagt. Dit inkomstenverlies wordt ontweken door het gebruik van een vrijwilligerstoelage, wat geen effect heeft op de hoogte van de bijstand (Steunpunt Uitkeringen, z.d.). En voor de wijkbedrijven betekent dit dat er een lager bedrag nodig is om de participanten te betalen dan bij de keuze voor loon. Hierdoor is de vrijwilligerstoelage financieel gunstig voor participanten én wijkbedrijven.

De deelnemersgroep is divers in demografie, ze verschillen in bijvoorbeeld leeftijd, geslacht, gezinssamenstelling en moedertaal (Gemeente Den Haag, 2022). Maar ze zijn homogeen in SES, omdat alle deelnemers afhankelijk zijn van uitkeringen en geen betaald werk hebben naast hun inzet bij de wijkbedrijven. Hierom bevinden ze zich in een vergelijkbare sociaaleconomische positie, wat deze doelgroep interessant maakt voor een uiteenzetting doordat verschillen in gezondheidsuitkomsten dan eventueel zouden kunnen worden toegeschreven aan het Nederlandse taalvermogen.

4. Methodologie

Dit hoofdstuk bespreekt de methodologie van het onderzoek. De indeling hiervan is dat eerst de datasetverzameling wordt beschreven, waarna de gekozen variabelen worden gemotiveerd, en wordt afgesloten door de analysemethodes toe te lichten, de causaliteit te bespreken en robuustheidstesten te presenteren. Vanaf dit hoofdstuk wordt de bijlage veelvuldig gebruikt, wanneer er een B in de nummering van de grafiek, tabel, of figuur zit, dan wordt er verwezen naar de bijlage.

4.1 Beschrijving dataset

De beantwoording van de onderzoeksvraag en bijbehorende hypothesen vindt plaats aan de hand van een kwantitatieve analyse gebaseerd op de gegevens die zijn verzameld onder deelnemers van het Instapeconomieproject in Den Haag. De data is afkomstig van twee lichtingen (2024 en 2025) waarbij deelnemers van vijf wijkcentra worden bevestigd. Deze wijkinitiatieven zijn Gewoon Sociaal Moerwijk, Wijk vol Waarde Vrederust, Lichtpuntjes Mariahoeve, Stagehuis Schilderswijk, en Wijkz Morgenstond.

De eerste drie organisaties waren al aanwezig bij de afname van de eerste lichte enquête in 2024. Participanten van Stagehuis en Wijkz zijn sinds dit jaar voor het eerst meegenomen in de afname van de enquête, echter zijn deze ten tijde van de oplevering van dit onderzoek niet aan de dataset toegevoegd vanwege verschillende moeilijkheden tijdens de dataverzameling in de opstartfase van deze initiatieven.

Een deel van de participanten hebben in 2024 én in 2025 de enquête ingevuld, hiervoor zijn voor enkele wijkcentra dezelfde identificatiecodes gebruikt zodat tijdsverschillen kunnen worden herkend. Deze identificatiecodes zijn anoniem en alleen bekend bij de coördinatoren van de wijkcentra. Alleen zijn deze identificatiecodes niet uniform bijgehouden en toegewezen, waardoor enkele toewijzingen zijn gedaan op basis van leeftijd (t_{2024+1}) en moedertaal van de participanten. De enquêtes die de participanten hebben beantwoord is gelijk gebleven aan die van lichte 2024 voor lichte 2025.

Door deze praktijken is eventueel longitudinaal onderzoek mogelijk, alleen in de onderzoeksopzet van dit onderzoek onwenselijk door de kleine hoeveelheid panel-observaties en gecompliceerdheid van het gebruikte model. Daarnaast is de mogelijke bias van het toeschrijven van panel-observaties aan participanten op basis van de genoemde aspecten door de onderzoekers een verder argument voor het niet uitvoeren van een panel analyse. In tabel

4.1 is uitgewerkt hoeveel enquêtes er per lichting zijn afgenomen per wijkcentrum, en hoeveel van deze participanten in twee enquête lichtingen aanwezig zijn.

De enquêtes die de participanten hebben beantwoord bestaan uit de thema's: achtergrondinformatie, motivatie, taal, gezondheid, financiën, opleiding en talent en vaardigheden, dagelijks leven, zingeving en perspectief, en sociaal netwerk. Hierin zijn er open en gesloten vragen aanwezig, waar de gesloten vragen binaire, ordinale, en nominale waarden bevatten. Voor dit onderzoek zijn voornamelijk de thema's 'taal' en 'gezondheid' belangrijk, maar er zullen meerdere vragen uit andere thema's worden aangehaald voor de analyse wat zal gebeuren in de hierop volgende paragrafen.

Tabel 4.1: Aantal enquêtes per wijkcentrum per lichting, en panel-observaties.

| | Lichting 2024 | Lichting 2025 | Panel-observaties |
|-----------------|---------------|---------------|-------------------|
| Wijkcentras | | | |
| Gewoon Sociaal | 15 | 11 | 7 |
| Wijk vol Waarde | 22 | 41 | 17 |
| Lichtpuntjes | 14 | 25 | 7 |

4.2 Gehanteerde variabelen

In dit stuk worden de gekozen variabelen gemotiveerd en ingericht. Hierin zullen meerdere transformaties worden toegelicht en zal afsluitend de beschrijvende statistiek van de gekozen variabelen worden geanalyseerd. Eerst zullen afzonderlijk de endogene en exogene variabelen worden besproken, waarna de mechanismes worden besproken en de gekozen variabelen daarvoor toegelicht.

4.2.1 Exogene variabelen

De exogene variabelen betreffen de taalbeheersing van de participanten. Hierin is eerder een uitsplitsing gemaakt tussen laaggeletterdheid en anderstaligheid. De formulering voor de variabele voor de beschrijving van laaggeletterdheid is dat laaggeletterden kunnen worden herkend door de moeite met lezen en schrijven. Dit zal worden berekend door middel van een combinatie van vragen, waarin wordt verondersteld dat het twijfelen aan de capaciteiten in het lezen en schrijven al aantoont dat de participant laaggeletterd is omdat de vragen het lezen van een poster/flyer of bijlage van een gemeentebrief betreffen, en het schrijven van een SMS of

Whatsapp of een brief naar de gemeente betreffen. Deze vragen samen vormen een proxy voor het lees en schrijfvermogen, waarin een ander antwoord dan het meest positieve een signaal geeft voor moeite met lezen en schrijven en zo een indicator voor laaggeletterdheid is. Hierdoor is de exogene binaire variabele *Laaggeletterd* geformuleerd als volgt:

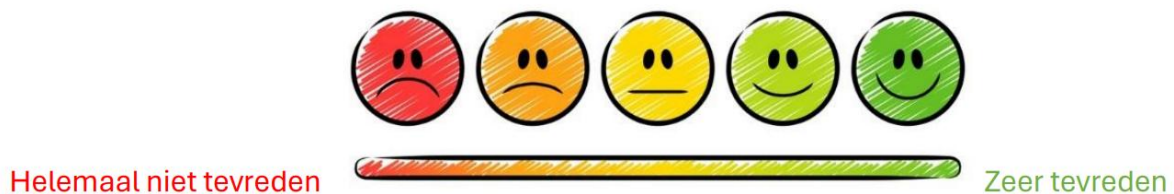
- Laaggeletterd = 1 als één van de volgende vragen een antwoord heeft dat anders is dan “Ja, helemaal” en 0 als geen van de andere opties die hieronder worden gespecificeerd worden benoemd.
 - 11. Kan je een Nederlandse poster/flyer lezen? & 12. Kan je een brief in de bijlage van de gemeente Den Haag lezen?
 - “Nee, dat lukt niet”
 - “Ja, een beetje”
 - 15. Kan je Whatsapp of SMS berichten schrijven in het Nederlands? & 16. Kan je een brief schrijven aan de gemeente in het Nederlands?
 - “Nee, dat lukt niet”
 - “Ja, maar met moeite”

Daarnaast is er de variabele *anderstalig*, wat de effecten van het missende spreekvermogen van het Nederlands door het hebben van een andere moedertaal aantoont. Dit toont de sociale effecten van een slechtere taalbeheersing aan. Hierin is de assumptie gelijk aan dat van laaggeletterden dat de gestelde vragen een vrij basis niveau hebben, waardoor het twijfelen aan één aspect rond gespreksvoering moeite met het spreken van het Nederlands aantoont. Hierom is de binaire variabele *Anderstalig* geformuleerd als volgt:

- *Anderstalig* = 1 als moedertaal niet (alleen) Nederlands is en één van de volgende vragen een antwoord heeft dat anders is dan “Ja, makkelijk” voor vraag 9 of “Ja, helemaal” voor vraag 13 en 14, en 0 als dat niet zo is. De andere opties worden hieronder gespecificeerd.
 - 9. Kan je een brood bestellen bij de bakker in het Nederlands?
 - “Nee, dat lukt niet”
 - “Ja, maar zonder hele zinnen”
 - 13. Kan je een gesprek hebben bij de koffie in het Nederlands? & 14. Kan je een zakelijk gesprek hebben in het Nederlands?
 - “Nee, dat lukt niet”
 - “Ja, maar met moeite”

En afsluitend wordt de variabele *Tevredenheid taalvermogen* gebruikt, wat afhangt van enquêtevraag 6 dat de zelf gerapporteerde tevredenheid met het Nederlandse taalvermogen afbeeld. Deze vraag luidt “Hoe tevreden ben je over hoe je de Nederlandse taal begrijpt?” en hanteert de scoremogelijkheid zoals getoond in figuur 4.1. De codering is 1 bij de optie onder de rode smiley en 9 bij de optie onder de groene smiley. Hierin wordt de ruimte tussen de smileys gezien als een even getal, waar bijvoorbeeld de ruimte tussen de rode en oranje smiley wordt geclassificeerd als 2. Hierdoor heeft deze variabele 9 punten. Deze codering zal bij alle volgende variabelen die hetzelfde antwoordmodel gebruiken gelden.

Figuur 4.1: Antwoordmodel enquêtevraag 6.

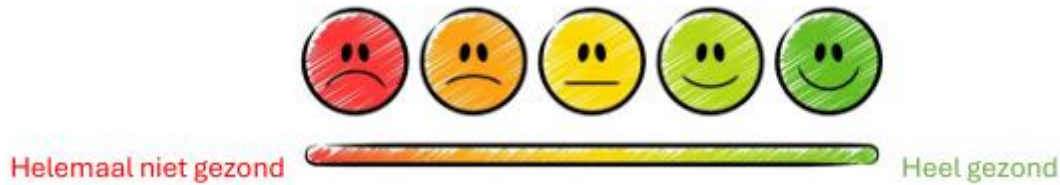


De combinatie van deze variabelen tonen de verschillende concepten aan die zijn behandeld en vanuit waar wordt beredeneerd in dit onderzoek. De variabele *Laaggeletterd* bespreekt de effecten van het niet hebben van een adequaat schrijf- en leesniveau, waarna de variabele *Anderstalig* inhoud geeft aan het sociale aspect van het niet sprekend beheersen van een taal, en de algemene taalvraag *Tevredenheid taalvermogen* de zelfreflectie van participanten meet van het begrip van de Nederlandse taal. Hiermee zijn verschillende indicatoren voor taalbeheersing besproken én geoperationaliseerd.

4.2.2 Endogene variabelen

De endogene variabelen in dit onderzoek betreffen de gezondheid en zullen op meerdere manieren worden geïllustreerd. De eerste endogene variabele is *Gezondheid*, gedefinieerd door vraag 18 “Hoe gezond voel jij je”, wat wordt beantwoord met de opties die te zien zijn in figuur 4.2. Deze variabele heeft 9 punten, waar 1= “Helemaal niet gezond” en het meest linkse is, en 9= “Heel gezond” en het meest rechtse is.

Figuur 4.2: Antwoordschaal vraag 18



Voor de tweede variabele van gezondheid wordt een variabele genomen in de vorm van de frequentie van ziekte. In de FCT is er de voorwaarde dat een fundamentele oorzaak betrekking heeft tot een set van gezondheidsuitkomsten, waardoor een afhankelijke variabele die alle instanties van ziekte bij een participant meeneemt in lijn is met de besproken theorie van gezondheidseffecten. Dit komt voort uit de assumptie dat iemand die vaker ziek is vaak een slechtere gezondheid ervaart. Hierdoor is de variabele *Hoeveelheid ziek* meegenomen, gebaseerd op vraag 23 “Hoe vaak ben je ongeveer ziek in het jaar?”. Deze vraag zal worden gecodeerd als 1= “Nooit”, 2= “1 of 2 keer”, 3= “3 of 4 keer”, 4= “5 of 6 keer”, 5= “7 keer of vaker”. Deze vraag is ordinaal van aard en heeft een oplopende status waardoor de hoogte een slechtere gezondheid zal weergeven.

De laatste afhankelijke variabele weergeeft de mentale gezondheid, wat gebaseerd is op het zelf gerapporteerde geluk van de respondenten. Hiervoor wordt vraag 53 “Hoe gelukkig ben je met je leven op dit moment” genomen, waarin het antwoordmodel zoals te zien is in Figuur 4.3 wordt gebruikt. De variabele zal *Geluk* heten en wordt vergelijkbaar gecodeerd als de variabele *Gezondheid* met 1= “Helemaal niet gelukkig” en 9= “Heel gelukkig”.

Figuur 4.3: Antwoordschaal vraag 53



4.2.3 Mechanismes

Zoals besproken zijn er meerdere mechanismes waardoor de gezondheidsverschillen door anderstaligheid, laaggeletterdheid, en een missend taalvermogen worden veroorzaakt in de twee

routes. De eerste route besprak het verband tussen taalbeheersing en SES, waar SES wordt gezien als een fundamentele oorzaak van gezondheidsverschillen. Binnen deze populatie zijn aspecten die vaak de SES opmaken homogeen, omdat alle participanten geen werk hebben buiten de werkzaamheden van het programma en geld verkrijgen vanuit de staat. Wel zijn er bepaalde aspecten die nog steeds kunnen worden meegenomen in de analyse, namelijk wat hun hoogst afgeronde opleiding is, hoeveel geld ze te besteden hebben per week, en hoeveel moeite ze hebben met rondkomen per week. Hierdoor zijn er variabelen die in de analyse worden opgenomen die het mechanisme vanuit SES afbeelden. Alle coderingen van de variabelen staan in tabel B4.1.

De eerste proxy voor SES is zoals vermeld de hoogst afgeronde opleiding en heet *Educatie*, afkomstig van vraag 38 “Wat is je hoogst afgeronde opleiding?”. *Educatie* is een prominente factor binnen de SES van een individu, waardoor dit wordt gebruikt als proxy voor SES. Deze proxy heeft een oplopende schaal in moeilijkheid van studie, waardoor hogere waarden een hoger niveau van educatie weergeeft. Verder worden twee vragen, 28 “Hoeveel kan je per week ongeveer besteden, naast je vaste lasten?” en 29 “Ervaar je uitdagingen met rondkomen?”, gebruikt voor het bevatten van de hoeveelheid geld dat de participanten te besteden hebben per week en de verplichtingen die daar tegenover staan. De eerst genoemde optie wordt in de variabele *Besteden* opgenomen en de tweede in de variabele *Rondkomen*. Deze twee variabelen proberen de dimensie van besteedbaar geld en de daar tegenover staande betalingsverplichtingen te illustreren. Beide variabelen laten bij hogere waarden een positiever beeld zien van de financiële situatie, waardoor hogere waarden overeenkomen met een positievere SES. Doordat hogere waarden van deze variabelen overeenkomen met een hogere SES, wordt hypothese 2 die het SES mechanisme beschrijft bevestigd als een slechter taalniveau correspondeert met een lagere SES en daarna een slechtere gezondheid.

De tweede beschreven route bevat meerdere mechanismes die al eerder zijn benoemd, hiervoor worden meerdere variabelen gebruikt om het verband dat het medieert te bevatten. Ten eerste de stress afkomstig van taalproblemen, wat enerzijds directe stress is en anderzijds afkomstig is van marginalisatie. Zoals weergegeven in tabel B4.1 laat vraag 25 “Hoeveel stress ervaar je in een week?” van de enquête een zeer directe afspiegeling zien van de ervaren stress van de respondent. Hierdoor is de *Stress* variabele een accurate weergave van het mechanisme waarin een hogere schaal minder stress weergeeft. Hypothese 4 beschrijft dit gestelde mechanisme en wordt bevestigd wanneer een lager taalniveau correspondeert met lagere ervaren stress en zo een slechtere gezondheid.

De stress van marginalisatie wordt aangetoond door middel van de proxy variabele *marginalisatie* en wordt weergegeven door vraag 47 “Zou jij graag contacten willen met mensen in jouw omgeving?”. In de enquête was geen directe vraag omtrent het gevoel van uitsluiting, waardoor er is gekozen voor een variabele die inspeelt op de drang van mensen die zich mogelijk ongewild gemarginaliseerd voelen om contact te krijgen met anderen in hun omgeving. Hierdoor is de huidige variabele gekozen, die binair is door de twee opties waarvan de optie gecodeerd met 1 poogt om marginalisatie te illustreren en zo het mechanisme van het stresseffect van marginalisatie op de gezondheid. Dit mechanisme betreft hypothese 5 en bevestigt desbetreffende hypothese als een lager taalniveau overeenkomt met een hogere kans om “Ja” te beantwoorden op de vraag en uitkomt op een slechtere gezondheid.

Verder is er het mechanisme rond de veroorzaakte inactiviteit door taalbarrières van respondenten. Zoals beschreven is dit mechanisme driedig, omdat het politieke, sociale en sportieve inactiviteit laat zien. Er is voor dit mechanisme gekozen om alleen vraag 20 “Doe je aan sport activiteiten?” te gebruiken dat de sportieve inactiviteit bevat. Deze keuze is gemaakt doordat er geen enquêtevraag omtrent de politieke inactiviteit was en een vraag van eventuele sociale activiteit alleen mogelijk was door het gebruik van een vraag die ging over hoeveel men naar buiten ging. Deze laatste optie is niet meegenomen in dit mechanisme omdat tijdens de afname van de enquête ambigue redenen werden gegeven waarom men naar buiten ging en veel verwarring was wat viel onder de vraag. Daarnaast baseerden veel participanten hun antwoorden op hoe vaak ze naar de supermarkt gingen, wat in sommige gevallen een sociaal uitje kan weerspiegelen maar ver weg ligt van de kern van dit mechanisme. Hierdoor is alleen het sportaspect genomen, wat als variabele is benoemd als *Sport*. Deze variabele heeft een oplopende schaal, waardoor een hogere waarde van de variabele correspondeert met meer sportactiviteiten. Desbetreffende variabele beschrijft hypothese 6 en bevestigt dit als een slechter taalniveau leidt tot minder sportactiviteiten en zo een lagere waarde van het mechanisme en daarna uitkomt tot een slechtere gezondheid.

Afsluitend zal de variabele *Dokter*, die afkomstig is uit vraag 10 “Kan je goed uitleggen aan de dokter als er iets is in het Nederlands?”, worden gebruikt om meerdere mechanismes te beschrijven. Deze vraag omvat ook een taalvraag zoals de vragen gebruikt voor de variabelen *Anderstalig* en *Laaggeletterd*, maar werd daar niet voor gebruikt omdat de andere taalvragen meer sociale interacties beschrijven terwijl *Dokter* een weerspiegeling geeft van de interpretatie van het advies dat medische professionals geven aan de participanten. Hierdoor betreft deze vraag meer gezondheidsvaardigheden doordat dit het interpretatievermogen weergeeft in een gesprek met de dokter. Hierom correspondeert deze variabele met hypothese 7 die het

interpretatievermogen van gezondheid outputs op de gezondheid poogt te illustreren, die wordt bevestigd bij als de taalvariabelen bij een niveau dat een slechter taalniveau reflecteert corresponderen met lagere waarden van *Dokter* en daarna een slechtere gezondheid.

In de operationalisering van de mechanismes komt geen variabele terug die correspondeert met de schaamte die de participanten ervaren door een slechter Nederlands taalvermogen, wat hypothese 3 beschrijft. Dit komt doordat deze hypothese een vrij complex verband beschrijft die de gestelde enquêtevragen niet kunnen bevatten. Er is echter een argument te maken dat de variabele *Dokter* enig indirect bewijs kan leveren voor deze hypothese. Het mechanisme van schaamte door beperkte taalvaardigheid hangt namelijk ogenschijnlijk sterk samen met het vermogen om effectief te kunnen communiceren met een medische professional. Wanneer tijdens een medisch gesprek moeilijk taalgebruik wordt gehanteerd, kan dit leiden tot vervreemding of een dissociatie van de verstrekte informatie, zoals is beschreven in het theoretisch kader. Hierdoor worden gezondheid inputs mogelijk minder goed geïnterpreteerd, wat de toegang tot en de effectiviteit van zorg negatief beïnvloedt. Alleen is de uitleg van het schaamtmechanisme ook gebaseerd op schaamte in het dagelijkse leven, waardoor de variabele *Dokter* slechts een klein deel ervan kan bevatten. En daarnaast is het directere verband van het niet kunnen snappen van de medische professional en daarom slechtere gezondheidsuitkomsten, wat wordt besproken in hypothese 7, logischer en directer dan het mogelijke verband in hypothese 3. Hierdoor wordt er geconcludeerd dat hypothese 3 niet kan worden verworpen of bevestigd in dit onderzoek door een gemis aan bruikbare variabelen.

4.3 Methode van analyse

In dit onderzoek wordt gebruik gemaakt van het Structural Equation Modeling (SEM). De keuze voor een SEM is gemotiveerd door de complexe verbanden die binnen dit onderzoek centraal staan. Voor de uitleg van SEM gebruikt dit onderzoek de uiteenzetting van English (2006). Het SEM model schat voor elk geselecteerd verband een waarde voor elke endogene variabele, die ook exogeen mogen zijn in een andere vergelijking in het model (English, 2006, 969). Hierin worden directe en indirecte effecten geraamd en is SEM door het pad analyse aspect een analysetechniek dat causale verbanden aantoonst wanneer de geformuleerde paden onder adequate assumpties zijn vastgesteld, waardoor het ook causaal modelleren wordt

genoemd. Wel moet er bij SEM rekening worden gehouden met verminderde normaliteit, missende data en uitschieters (English, 2006, 969).

In essentie is SEM een analysemethode waarmee meerdere relaties tussen variabelen gelijktijdig worden geschat. Dit doet het door middel van OLS schattingen die gelijktijdig worden uitgevoerd in plaats van dat ze los van elkaar worden geanalyseerd, zoals bij een reguliere OLS gebeurt. Het gelijktijdige schatten houdt in dat het geformuleerde model, waarin een variabele zowel endogeen als exogeen kan zijn in verschillende vergelijkingen, volledig wordt meegenomen in eenzelfde schatting. Door het gelijktijdige schatten wordt er rekening gehouden met indirecte effecten en onderlinge verbanden tussen endogene variabelen die tegelijk worden geschat. Hierdoor wordt er gecontroleerd voor meetfouten terwijl de volledige hypothetische opbouw gelijktijdig wordt getest, wat waardevol is voor het snappen van onderlinge verbanden in het model (Nifadkar et al., 2012, 1158)

Op basis van de FCT wordt in dit onderzoek geanalyseerd hoe taalbeheersing de gezondheid beïnvloedt door middel van onderliggende mechanismes. Een persoon zijn gezondheid is een variabele die door veel verschillende aspecten tegelijk wordt veroorzaakt, waardoor afzonderlijke berekening van elk mechanisme geen actueel beeld schetst van het beredeneerde verband.

Hierom is SEM geschikt, omdat het een model is dat meerdere processen van analyse gelijktijdig kan uitvoeren en zo een passender model kan schetsen (English, 2006, 840). Daarnaast is de in dit onderzoek gehanteerde analyse vanuit mechanismes passend voor SEM, omdat het in zijn simpelste vorm pad analyse kan toepassen en voornamelijk goed werkt wanneer de endogene variabele in één equatie de exogene in de ander is (English, 2006, 840). Dit is fundamenteel voor dit onderzoek, sinds de volledige redentatie in dit onderzoek vanuit FCT gebaseerd is op complexere indirecte verbanden lopend door mechanismes. Hiermee gaat de getoonde theorie uit van een volledig indirect verband, omdat de redentatie vanuit de FCT een effect op en vanuit de besproken mechanismes beschrijft en geen direct verband, waardoor in deze analyse een volledig mediatie model wordt gehanteerd.

Doordat de simultane schatting van het verband gewenst is, is de keuze van het gebruik van SEM boven een klassieke mediatie-analyse gegrond om de besproken theorie te testen. De in een klassieke mediatie-analyse gebruikte regressie analyses kunnen slechts één vergelijking gelijktijdig uitvoeren, terwijl een SEM simultaan een set van regressies uitvoert die collectief de variatie in de data bevatten (Ballen & Salehi, 2021, 4). Hierdoor is er door de in het theoretisch kader beredeneerde onderlinge verbanden tussen de gebruikte variabelen meer reden om een SEM uit te voeren als mediatieanalyse dan een klassieker model.

In het door dit onderzoek gebruikte model gebeurt de schatting in twee stappen: eerst wordt het effect van taalvariabelen X_n op de mechanismes M_i geschat, waarna het effect van de mechanismes M_i gelijktijdig op de afzonderlijke gezondheidsvariabelen Y_j wordt geschat. Het regressiemodel wordt hieronder geïllustreerd:

Stap 1, effect van X_n op M_i :

$$M_i = \alpha_i + \beta_i X_n + \varepsilon_i, \quad i = 1, 2, \dots, 7, \quad n = 1, 2, 3.$$

Stap 2, effect van M_i op Y_j :

$$Y_j = \gamma_j + \sum_{i=1}^7 \Theta_{ij} M_i + \eta_j, \quad j = 1, 2, 3.$$

Deze vergelijking is compact opgesteld maar omvat alle vergelijkingen die gelijktijdig worden berekend. In stap 1 wordt voor elk van de zeven mechanismes M_i een OLS regressie uitgevoerd met één van de drie taalvariabelen X_n . Hierin is β de geschatte coëfficiënt van het effect van de taalvariabele op het mechanisme M_i , waardoor het de gezochte waarde is omdat het een verband schetst tussen de taalvariabele en mechanismes. De constante α en de storingsfactor ε zijn hierin voor elk mechanisme i uniek, omdat elk mechanisme afzonderlijk wordt geschat. In deze stap is er geen effect van de gelijktijdigheid van de berekeningen, omdat één taalvariabele als enige verschuivingen veroorzaakt in meerdere mechanismes.

Stap 2 laat een complexer verband zien omdat hier alle verschillende mechanismes in elkaar samen komen door een gelijktijdige schatting van de drie gezondheidsvariabelen Y . Hierin wordt het eigen effect dat elk van de beschreven mechanismes heeft op elke gezondheidsvariabele in relatie tot elkaar gelijktijdig geschat. De hieruit ontstaande coëfficiënt Θ_{ij} beeldt deze simultane schatting uit en is ook de gezochte coëfficiënt om het verband tussen de mechanismes i en gezondheidsvariabelen j aan te tonen. Het toont een matrix van drie bij zeven coëfficiënten aan die het effect van elk mechanisme op elke afzonderlijke gezondheidsvariabele laat zien. Daarnaast is er in deze stap ook aanwezigheid van de constante en storingsfactor per gezondheidsvariabele j in de vorm van respectievelijk γ_j en η_j .

De samengevoegde vergelijking van deze twee stappen is essentieel voor het begrijpen van het gevraagde effect, daarom wordt dat hieronder geschreven.

Samengevoegde vorm:

$$Y_j = \gamma_j + \sum_{i=1}^7 \Theta_{ij} (\alpha_i + \beta_i X_n + \varepsilon_i) + \eta_j, \quad i = 1, 2, \dots, 7, \quad n = 1, 2, 3$$

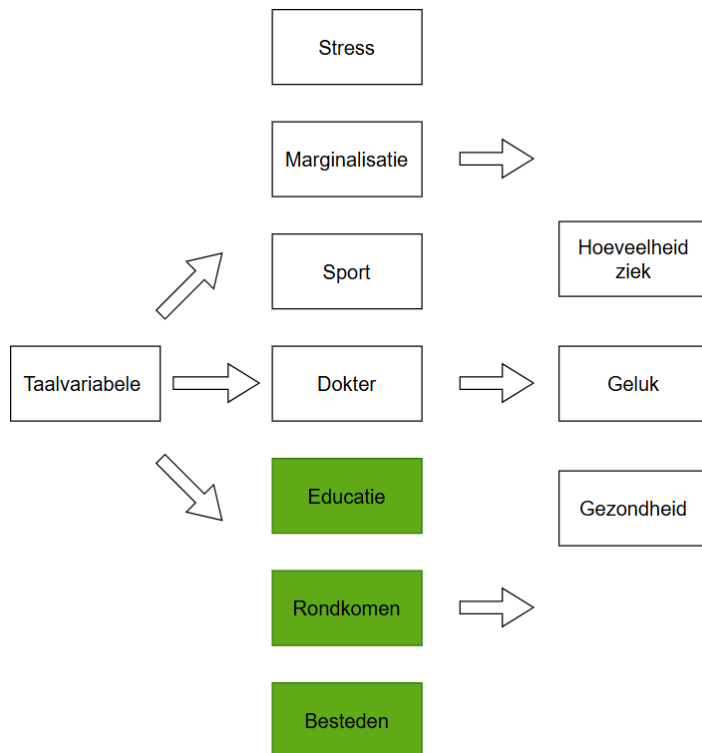
$$j = 1, 2, 3.$$

In deze equatie zijn in principe alleen de variabelen Θ_{ij} en β_i afhankelijk van de hoogtes van de onderzochte variabelen X_n en Θ_{ij} , omdat de anderen constantes en errortermen van de equatie op zich aangeven. Vanuit deze samengevoegde vorm is te zien dat de hoogte van de gezondheidsvariabele Y_j afhangt van de taalvariabele X_n door de som van alle coëfficiënten Θ_{ij} keer de coëfficiënten van de het effect dat de taalvariabele op de mechanismes had β_i . Hiermee laat de combinatie van deze twee coëfficiënten, mits significant, het gemedieerde effect zien van X_n op Y_j door de mechanismes en zo eventueel bewijs zien voor de gestelde hypothesen en gehanteerde theorie. De analyse van deze coëfficiënten laat een duidelijk beeld zien van de relatie tussen de taalbeheersing van de participanten en de ervaren gezondheid.

In dit onderzoek wordt constant uitgegaan van twee routes waardoor taalbeheersing de gezondheid beïnvloedt, de route door SES en de route vanuit de mechanismes. Het SES is op veel manieren verweven met de mechanismes waardoor mogelijk bias kan ontstaan, waardoor het schatten van het model met en zonder SES een gepaste manier is om eventuele verwevenheid van de variabelen die SES uitbeelden en de mechanismes te verklaren. Hierdoor is er per taalvariabele één model met een i van 1, 2, 3, 4 ipv 1, 2, ..., 7 omdat er minder variabelen in de mechanismelaag zitten door het weglaten van de SES variabelen.

Het geschetste model wordt weergegeven in figuur 4.4. Hierin kan je zien dat het model vanuit de gekozen taalvariabele X_n voor die berekening een effect heeft op elk mechanisme in de tweede lijn, waarna elk mechanisme uit de tweede lijn een effect heeft op elke gezondheidsvariabele in de derde lijn. Dit is het model dat wordt getest samen met de gekozen variabelen. De SES variabelen zijn groen afgebeeld, omdat deze afhankelijk van het niet of wel meenemen van deze variabelen wegvallen of meedoen in de regressie.

Figuur 4.4: Hoofdmodel.



4.4 Over causaliteit

Zoals besproken door English (2006, 969-971) kan een SEM analyse causale uitspraken maken over verbanden als het gemaakte model is gebaseerd op empirisch bewezen paden. Daarnaast moeten de variabelen die deze paden opmaken volledig en geschikt zijn zodat elke variabele daadwerkelijk een beschreven pad beschrijft. Afsluitend moet er genoeg data aanwezig zijn om een correcte schatting van het model te maken, waar bij complexe modellen meer observaties nodig zijn voor correcte causale inferenties. In dit onderzoek zijn er alleen meerdere voorwaarden waaraan niet kan worden voldaan, waardoor er geen causale inferenties kunnen worden gemaakt en er louter gesproken kan worden over eventueel bewijs voor verbanden.

Er zijn bijvoorbeeld meerdere problemen op te merken aan de datakwaliteit van dit onderzoek. Ten eerste is de enquête gemaakt buiten dit onderzoek om, waardoor veel beschreven effecten niet exact kunnen worden gemeten en er meerdere variabelen afhankelijk zijn van een proxy die in enige instanties is gemaakt op basis van inferenties die kort door de bocht kunnen worden genoemd. Hierdoor zijn enkele variabelen vatbaar voor alternatieve verklaringen in plaats van het beoogde effect dat het hoort af te beelden.

Verder is er een vermoeden van een verstoring in de werking van het model door weggelaten variabelen. Een erg prominent voorbeeld hiervan is het taalniveau van Engels. Nederland is het best Engelssprekende land per wereld waarvan Engels niet de hoofdtaal is (World Population Review, 2025). Kijkend naar de besproken theorie, zijn enkele mechanismes berust op het kunnen communiceren met een ander en het kunnen lezen van informatie, wat in Nederland vrij vaak ook in het Engels gebeurt waardoor het Nederlandse taalvermogen vermoedelijk minder effect heeft op de gezondheid dan er waarschijnlijk zou zijn als de respondenten geen Engels spraken. Hierdoor is het niet kunnen meten van Engels taalniveau van de respondenten een duidelijk blijk van een verstorende werking in het model, ofwel een omitted variable.

Ten tweede zijn er meerdere problemen aan te merken waardoor er structureel andere interpretaties en zo storing in de validiteit van de gestelde vragen in de enquête mogelijk zijn. De enquête wordt bijvoorbeeld afgenomen in twee talen, waardoor eventuele verschillen in taal tussen de Nederlandse en Arabische versie kan worden verklaard door mogelijke vertalingsfouten. Daarnaast zijn er ook dingen aan te merken voor de niet Nederlands sprekenden die in het Arabisch antwoord gaven of een tolk gebruikten om antwoord te geven. Voor de Arabische antwoorden moesten de enquêteafnemers gebruik maken van vertalingssoftware op de geschreven antwoorden om deze in te voeren. Voor degenen die een tolk gebruikten is samenwerking met anderen waargenomen tijdens het invullen van de enquête, naast mogelijk sociaalwenselijke antwoorden omdat er vaak een tolk én een afnemer tegenover deze participanten zaten. Hierdoor is er een aannemelijke kans op structurele verschillen in de tendens van antwoorden tussen verschillende groepen.

De gestelde vragen zijn in meerdere gevallen ambigue, waardoor veel verschillende interpretaties van deze vragen kunnen zorgen voor verminderde normaliteit. Hierbij is er bijvoorbeeld geen mogelijkheid tot het vergelijken van welke taal mensen als Moedertaal hebben, omdat meerdere participanten tot wel 3 talen hier hebben ingevuld. Ook zijn er enkele vragen in willekeur niet ingevuld, waardoor enkele participanten volledig worden uitgesloten door het SEM model in de huidige vorm wat zo resulteert in een kleinere onderzoeksgroep.

Verder zijn er ook punten binnen de analysemethode die leiden tot een onvermogen om causale verbanden te illustreren. Ten eerste is de dataset bijna gehalveerd ten overstaande van wat was beoogt. Doordat enkele wijkcentra geen agendaruimte hadden om tijdig de enquêtes te laten afnemen, of minder participanten hadden dan eerder was beoogt, is er een veel kleinere onderzoeksgroep overgebleven dan in de opstartende fase van dit onderzoek werd verwacht.

Dit is nadelig voor een SEM model, omdat dit model veel regressies simultaan uitvoert en complexer is waardoor meer observaties gewenst is.

Afsluitend zijn er door de kleiner uitgevallen onderzoeksgroep geen controlevariabelen toegevoegd aan de gehanteerde modellen. Deze keuze is gebaseerd op de al aanwezige complexiteit van het model, waardoor convergentie ervan tijdens de berekening onwaarschijnlijker wordt bij toevoeging van verdere variabelen en kan leiden tot onstabiele schattingen en slechte toepasbaarheid van het model. Hierdoor worden eventuele alternatieve verklaringen van gezondheidsverschillen als leeftijd niet in het model gebruikt wat bijdraagt aan het onvermogen om causale verbanden vast te stellen.

Er zijn wel maatregelen getroffen tegen de gestelde factoren die de validiteit van dit onderzoek ondermijnen. Dit onderzoek maakt gebruik van drie exogene taalvariabelen, en drie endogene gezondheidsvariabelen. Door het gebruik van meerdere variabelen wordt een waargenomen effect van het taalniveau op de gezondheid minder waarschijnlijk veroorzaakt door kans. Daarnaast zorgt de diversiteit in deze variabelen voor een grotere context van de uitkomsten, waardoor een geschetst effect beter kan worden begrepen. Verder wordt er om de enkele missende uitkomsten te compenseren gebruik gemaakt van `method(mlmv)` in STATA, wat een schatting maakt wat zou moeten staan op de plek van een willekeurig missende waarde op basis van de andere aanwezige variabelen. Hierdoor worden er minder participanten uitgesloten die wél waarden hebben voor de andere waarden dan degene die mist. Ook zijn de variabelen afkomstig vanuit theoretische assumpties waardoor er een sterk fundament zit aan de gestelde relaties en zo meer validiteit ontstaat. Afsluitend wordt er uitgebreid de robuustheid van het model getest door middel van meerdere methodes die worden beschreven in de volgende paragraaf.

4.5 Robuustheidstesten

Zoals eerder beschreven zijn er enkele factoren die eventuele validiteit kunnen aantasten, waardoor het testen van de robuustheid van eventuele bevindingen uit de analyse belangrijk zijn voor het maken van betrouwbare inferenties. Ten eerste is er in het model geen rekening gehouden met eventuele correlatie tussen de fouten van het gebruikte model. Hierom zal de eerste robuustheidscheck een Seemingly Unrelated Regression (SUR) uitvoeren zodat hier wel rekening mee wordt gehouden. Tijdens het uitvoeren van deze test wordt onderzocht of de verbanden die eerder zijn bevonden standhouden wanneer er rekening wordt gehouden met de

onderlinge samenhang van fouttermen. Afwijkende uitkomsten van het SEM model geeft correlatie van deze fouttermen aan, wat duidt op foutief geschatte standaardfouten in het standaardmodel en zo wijst naar onbetrouwbare toetsing van effecten (UCLA, z.d). Wanneer de geobserveerde verbanden na de test standhouden, dan bevestigt het dat de conclusies niet afhangen van eventueel ruis. De SUR test is daarmee complementair aan SEM omdat het validiteit toevoegt aan de gemaakte inferenties als de uitkomsten overeen komen.

De keuze voor SEM boven SUR als methode van analyse is te rechtvaardigen door de complexiteit van de variabelen, waarvan veel worden gemeten via proxies die in SEM gecorrigeerd worden op meetfouten, iets wat SUR niet ondersteunt. Daarnaast is SEM praktischer in aanvullende opties, waardoor bijvoorbeeld de method(mlmv) uit de volgende test gebruikt kan worden en zo beter om kan worden gegaan met missende waardes.

De tweede robuustheidstest baseert zich daarmee op het gebruik van method(mlmv) in de modellen om minder subjecten met ontbrekende data uit te sluiten van de analyse. Deze keuze leidt tot meer statistische kracht van schatting omdat het zorgt voor meer participanten, maar door de werking van method(mlmv) kan er eventuele bias in de uitkomsten ontstaan. Dit komt doordat deze methode ervan uitgaat dat missende waardes willekeurig ontbreken (Statacorp LLC., 2025). Dit onderzoek gaat ervan uit dat de ontbrekende waarden volledig willekeurig ontbreken door eventuele fouten bij invoering en misinterpretatie bij het invullen van de enquête door de participanten, waardoor het testen van deze veronderstelling voor de robuustheid van de uitkomsten noodzakelijk is.

Een verdere robuustheidstest wordt gefocust op de eerder vermelde keuze voor een volledige in plaats van een gedeeltelijke mediatie analyse. Doordat er gebruik wordt gemaakt van louter de mechanismes en niet een direct verband tussen de taal- en gezondheidsvariabelen kan het zo zijn dat dit directe verband zorgt voor een andere uitkomst dan eerder omdat de mechanismes minder sterke voorspellers zijn dan het directe verband. Wanneer dit in het SEM model wordt gekoppeld aan standaard OLS regressies van de afzonderlijke taalvariabelen op de afzonderlijke gezondheidsvariabelen, kan er gekeken worden of bijvoorbeeld de correlatie tussen enkele variabelen wordt weggenomen door de mechanismes. Als dit het geval is, laat het model zien dat er variatie over wordt genomen door de mechanismes, wat bewijs geeft voor de bepleitte relatie tussen taal en gezondheid dat door mechanismes loopt. Dit zorgt daarmee voor een betere robuustheid van de inferenties die worden gemaakt.

Afsluitend is er een robuustheidstest nodig voor de eerder gemaakte notie dat eventueel bias is ontstaan vanuit de vertaling van de enquête naar de Arabische taal. Hierdoor is het noodzakelijk dat de laatste robuustheidstest kijkt naar het effect van het spreken van Arabisch

als moedertaal op het model. Een kleine kanttekening is dat er inderdaad ook Arabisch sprekkenden zijn die de enquête in het Nederlands hebben gedaan, alleen hiervan is geen data waardoor deze methode de enige mogelijke is om een mogelijk verband te testen. In deze test wordt Arabisch eerst gebruikt als taalvariabele, waardoor het kijkt of er eventueel grote verschillen ontstaan in de relaties met de mechanismes. Daarna wordt Arabisch als een mechanisme beschouwt, om te kijken of het een deel van de voorspellingskracht van de andere mechanismes wegneemt. De uitkomsten worden dan vergeleken met de hoofdmodellen om te kijken wat de veranderingen zijn wanneer we alleen kijken naar deze subgroep. Naast het bekijken van andere effecten, kan het ook meer diepte brengen aan de inzichten die zijn opgedaan uit de hoofdanalyse als er afwijkende uitkomsten zijn. Hierom brengt het naast robuustheid ook verdere mogelijkheden tot inzichten afkomstig uit dit onderzoek.

5. Analyse

In de analyse wordt eerst een descriptieve analyse van de gebruikte variabelen uitgevoerd, waarna de uitkomsten van de SEM analyse zullen worden getoond en de daaruit ontstaande bevindingen beschreven, waarna deze bevindingen worden geverifieerd door middel van de beschreven robuustheidstesten.

5.1 Descriptieve analyse

In tabel B5.1 worden de belangrijkste statistieken van de gebruikte variabelen in deze analyse getoond. De eerste twee variabelen zijn de samengestelde taalvariabelen die vrij vergelijkbaar zijn gecodeerd. Het verschil in de observaties tussen de twee is te verklaren doordat enkele respondenten niet hebben geantwoord op de leesgerichte taalvragen, maar wel op de spraakgerichte vragen.

Zoals te zien is in tabel B5.1 heeft de variabele *Laaggeletterd* in het gemiddelde bijna een waarde van 0.1 hoger dan de variabele *Anderstalig*. Deze observatie geeft aan dat er meer mensen worden geclassificeerd als laaggeletterd dan als anderstalig, waardoor kan worden opgemerkt dat er ook een redelijk groot deel laaggeletterden het Nederlands als moedertaal hebben. Dit wordt bevestigd door figuur B5.1, waar is te zien dat er 13 participanten alleen laaggeletterd zijn en twee alleen anderstalig. Wel zijn de twee anderstaligen die niet laaggeletterd zijn interessant, omdat ze dan hebben aangegeven problemen te hebben met het

spreken van de taal, maar niet met het lezen van de taal omdat ze anders in de groep “Beiden” zouden komen.

Dat ongeveer 66% van de participanten laaggeletterd is volgens de in deze studie gebruikte classificering, is grofweg vijf keer groter dan het percentage Nederlandse inwoners die moeite hebben met het lezen van de taal (CBS, z.d.; Stichting Lezen en Schrijven, z.d.). Dit toont aan dat de demografie van de onderzoeksgroep daarin sterk afwijkt van de Nederlandse bevolking. Daarnaast is door de grote groep anderstaligen in het onderzoek te stellen dat de onderzoeksgroep ook vaker een andere moedertaal heeft dan het Nederlands ten overstaande van de doorsnee bevolking.³

Hierdoor is de bevinding dat het grootste deel van de groep voor de variabele *Tevredenheid taalvermogen* tussen de 5 en 9 hebben gescoord interessant. Want ongeveer 65% van de groep heeft wel aangegeven moeilijkheden te hebben met de Nederlandse taal, waardoor een bandbreedte die vrijwel volledig in het positieve domein van de klassering in deze variabele ligt een wellicht te positief beeld geeft in de variabele. Ondanks dat is er te zien dat respondenten ook hebben gereageerd met de meest negatieve waarde, waardoor er wel genoeg variatie verwacht is in de variabele in het negatieve domein. Om deze verwachting te bekrachtigen is er een histogram gemaakt die te zien is in figuur B5.2. Hierin is te zien dat er inderdaad een sterk tendens is in deze vraag om positiever te antwoorden, maar er wel enige waarden die een negatievere betekenis hebben aanwezig zijn. Ook is er met deze histogram duidelijk dat de participanten de opties tussen de smileys in de vraag vermijden, wat begrijpelijk is sinds er geen initiële uitleg is geweest over de antwoordopties.

De mechanismes laten meerdere dingen zien. De variabele *Stress* laat met een gemiddelde van 5.04 zien dat de participanten zich gemiddeld gestresst voelen: 31.5% ervaren heftigere stress tegenover de 34.6% die lichtere tot geen stress ervaren dan het gemiddelde niveau dat is afgebeeld met score 5. Verder begeert 51.7% van de participanten meer sociaal contact, afgebeeld in de variabele *Marginalisatie*. Deze vraag wordt gekenmerkt met de meeste missende observaties, wat kan duiden op een ander gevoel opgewekt door de vraag of een eventueel probleem met de vraagstelling waardoor mensen de gestelde vraag waarop deze variabele is gebaseerd niet willen of kunnen beantwoorden. De variabele *Sport* is daarentegen een van de meest complete met 128 observaties. Deze variabele laat zien dat de gemiddelde participant minder sport dan 1x in de week, wat een vrij inactieve groep qua sport afbeeld. De

³ Deze opmerking is gemaakt op basis van de persoonlijke ervaring omdat er geen directe cijfers over anderstaligheid in Nederland te vinden zijn.

variabele *Dokter* laat zien met een gemiddelde van 1.42 dat de participanten een sterker tendens hebben richting het beantwoorden dat ze een medisch professional goed kunnen verstaan tegenover het niet tot matig begrijpen wat ze zeggen. Hierin heeft 42.5% een ander antwoord gegeven dan het meest positieve antwoord, waar 15.8% aangaf helemaal niet te kunnen spreken met de professional en 26.8% het een beetje te kunnen.

Verder laten de mechanismes die het SES beschrijven enkele interessante waardes zien. Ten eerste kan je door het maximum van *Educatie* zien dat er geen participant een PhD opleiding heeft genoten omdat de hoogste waarde een acht is. Van de respondenten heeft 15.9% geen enkele studie afgerond, 21.4% alleen de basisschool, 32.5% een diploma van de middelbare school, en 30.2% een voortgezette studie genoten. Dit is beduidend lager dan het landelijke gemiddelde, waar de participanten van dit project met 37.3% afronding van de basisschool of geen afgeronde studie ruim vier keer hoger is dan het landelijke gemiddelde van 7.9% (Ministerie van Onderwijs, Cultuur en Wetenschap, z.d.). De variabele *Besteden* laat met een gemiddelde van 2.7 zien dat de onderzoeksgroep gemiddeld tussen de 50 en de 150 euro per week te besteden heeft. En de variabele *Rondkomen* laat met een gemiddelde van 4.6 zien dat ze hierbij een tendens hebben om moeite te hebben om rond te komen met dit budget.

De gezondheidsvariabelen laten meerdere aspecten zien van de gezondheid van de respondenten. De participanten zijn zoals te zien aan het gemiddelde van de variabele *Hoeveelheid ziek* ongeveer 3 of 4 keer ziek per jaar. Figuur B5.3 laat echter zien dat deze verdeling niet normaal is en men het vaakst “1 of 2 keer” als optie hebben gekozen. Veel participanten hebben ook aangegeven erg vaak ziek te zijn, met “7 keer of vaker” zijnde een hoge frequentie. De variabele *Geluk* schetst een groep die gemiddeld zichzelf gelukkig acht met een 6.6. Dit wordt versterkt door het figuur B5.4, waarin te zien is dat er een hoge frequentie aan participanten hebben aangegeven zich gemiddeld gelukkig te voelen. Afsluitend laat de variabele *Gezondheid* met een gemiddelde van 5.3 zien dat het tendens van de participanten is om zichzelf iets gezonder dan neutraal te schatten. Dit wordt bevestigd door figuur B5.5, waarin een normale verdeling te zien is met een tendens naar de meer positieve reflecties van gezondheid. Verder is in deze histogrammen ook weer sprake van het tendens om de smileys in het antwoordmodel te gebruiken, zoals te zien is in de lage frequentie van even getallen ten overstaande van oneven getallen.

Afsluitend is het interessant om te kijken of er een initieel verband is tussen de afzonderlijke taalvariabelen en de afzonderlijke gezondheidsvariabelen. Dit verband is afgebeeld in Tabel B5.2 in een OLS regressie. Hierin is te zien dat de *Tevredenheid taalvermogen* in geen instantie een significant direct effect heeft, terwijl *Laaggeletterd* een

significant direct positief effect heeft op *Hoeveelheid ziek* en *Geluk* en *Anderstalig* alleen een significant positief effect op *Hoeveelheid ziek*. Deze regressies brengen de bevindingen gemaakt in de SEM analyse en de robuustheidstesten in contrast. En daarnaast laat het ook meteen een onverwachts verband zien, waar bijvoorbeeld volgens deze regressies het laaggeletterd zijn zorgt voor een positief effect op het ervaren geluk van de participanten.

5.2 SEM

In dit hoofdstuk zijn de resultaten uit de SEM analyse, afgebeeld in tabel B5.3, gedeeld. Deze tabel is te lezen als volgt, de eerste kolom onder “Effect X_n op mechanisme” betreft de waarden van de gekozen X op het mechanisme dat eronder is vermeld. De coëfficiënt die dan aan die rij correspondeert is de β_i waarde in stap 1 en de samengevoegde vergelijking. De tweede kolom onder “Effect mechanisme op Y ” betreft het effect van de mechanismes die worden genoemd in de eerste kolom op de gezondheidsvariabelen in de tweede kolom. In die rij staat de coëfficiënt die correspondeert met Θ_{ij} uit stap 2 en de samengevoegde vergelijking. Dus de eerste rij nummers geven het geschatte effect van de taalvariabelen op het mechanisme *Stress* weer, waarna de tweede rij de standaarddeviatie daarvan weergeeft. De derde rij geeft daarna het effect van mechanisme *Stress* weer op gezondheidsvariabele *Hoeveelheid ziek* en zo voort. Dus het mechanisme dat boven de gezondheidsvariabelen staat is het mechanisme dat de gezondheidsvariabelen beïnvloedt. De verdere kolommen geven aan welke taalvariabele X_n is gebruikt en “Met SES” en “Zonder SES” slaan op het wel of niet gebruiken van de SES mechanismes. In dit hoofdstuk wordt stapsgewijs elk mechanisme behandeld en de effecten die daardoorheen lopen.

Daarom wordt begonnen met mechanisme *Stress*, waarin in eerste instantie enig verband waar te nemen is met de taalvariabele *laaggeletterd* en in alle gevallen een effect heeft op de gezondheidsvariabelen. Deze variabele vertoont bij hogere waarden mindere stressniveaus, waardoor daarbij rekening moet worden gehouden met de interpretatie van de variabele. De variabele *laaggeletterd* laat met én zonder SES een met 90% significant positief effect zien op het stressniveau met een waarde van 0.710 met en 0.737 zonder de SES variabelen. Dit betekent dat het laaggeletterd zijn ervoor zorgt dat respondenten van de enquête gemiddeld 0.710 of 0.737 hogere waarden geven aan de variabele *Stress* dan mensen die niet laaggeletterd zijn.

Het vermelde effect laat dus zien dat laaggeletterden minder stress ervaren dan niet laaggeletterden, wat niet volgens de in dit onderzoek gepresenteerde redenatie is. Dit effect heeft alleen een bandbreedte met de standaarddeviatie die over het nulpunt heen gaat, waardoor

dit positieve effect niet met een 90% significantie kan worden bevestigd. Daarnaast laten *Anderstalig* en *Tevredenheid taalvermogen* insignificant effecten zien, waardoor het algehele verband tussen de taalbeheersing van de Nederlandse taal en het mechanisme Stress niet kan worden geduid.

Stress heeft wel als mechanisme een significant verband getoond met elke gezondheidsvariabele. Het laat zien dat *Stress* voor elk model een negatief effect heeft op de hoeveelheid dat participanten ziek worden. Voor elke stap dat *Stress* omhoog gaat in de negen puntschaal, zorgt het afhankelijk van het gebruikte model voor een gemiddeld negatief effect van 0.116 tot 0.136 op de hoeveelheid dat men ziek is. Dit heeft als betekenis dat een hogere waarde van *Stress* zorgt voor een lagere waarde van *Hoeveelheid ziek*, wat het effect toont dat minder stress leidt tot minder ziekte met een betrouwbaarheidsniveau van 95%. De variabele *Stress* toont ook een positief verband op 99% betrouwbaarheidsniveau met de variabelen *Geluk* en *Gezondheid*, afhankelijk van het gebruikte model van respectievelijk gemiddeld 0.473 tot 0.474 en 0.547 tot 0.558 per trede van de *Stress* variabele. Dit laat dus een zeer sterk positief effect zien van *Stress* op het geluksniveau en zelf gerapporteerde gezondheidsniveau. Hierdoor laten deze verbanden ook zien dat hogere waarden van de variabele *Stress*, en zo lagere niveaus van stress, zorgen voor een betere mentale en algemene gezondheid van de respondenten. Door deze verbanden kan er worden gesteld dat het mechanisme *Stress* bij hogere waarden een betere gezondheid, afgebeeld door de meerdere gezondheidsvariabelen, teweeg brengt en zo betekent dat er bewijs is dat meer stress een slechtere gezondheid veroorzaakt. Deze bevindingen zijn lijn met de voorspellingen, alleen kan er door de vermelde insignificante tot averechts van voorspelde verbanden tussen de taalvariabelen en het mechanisme *Stress* geen uitspraak worden gemaakt over hypothese vier.

Daarna wordt de variabele *Dokter* besproken, waarbij het verband tussen de taalvariabelen en het mechanisme allemaal met een 99% betrouwbaarheidsniveau significant zijn. De variabelen *Laaggeletterd* en *Anderstalig* laten beiden een negatief verband zien met de variabele *Dokter* van respectievelijk gemiddeld 0.820 en 0.932. Dit houdt in dat wanneer participanten laaggeletterd of anderstalig zijn, ze in beide gevallen bijna een volledige waarde van *Dokter* lager scoren dan wanneer ze het niet zijn. Dus het bevat dat ze minder goed kunnen spreken met de medische professionals en zo het advies minder goed kunnen interpreteren. Dit verband wordt bevestigd door de positief gecodeerde *Tevredenheid taalvermogen* die een positief verband heeft met de variabele *Dokter*. Dit verband houdt in dat participanten per hogere trede van de 9 puntschaal van *Tevredenheid taalvermogen* gemiddeld 0.215 hoger scoren dan mensen van de vorige trede voor de variabele *Dokter*. Kortom, dit laat zien dat bij

beter taalvermogen de participanten hoger scoren voor de variabele *Dokter* wat zo een positief verband laat zien tussen het beheersen van de Nederlandse taal en het kunnen spreken met medische professionals, wat het getheoretiseerde verband is.

Het mechanisme wordt daarna in verband gebracht met de gezondheidsvariabelen, waar te zien is dat het voor elk model een significant negatief verband heeft met de variabele *Hoeveelheid ziek* en geen significant verband heeft met de andere twee gezondheidsvariabelen. Het negatieve verband met de variabele *Hoeveelheid ziek* is bij alle schattingen betrouwbaar op een 95% niveau en zorgt afhankelijk van het gebruikte model voor een 0.323 tot 0.392 verlaging in de variabele *Hoeveelheid ziek* per trede van *Dokter*. Dit effect laat zien dat een hogere waarde van *Dokter*, wat een betere gezondheidsvaardigheid weerspiegelt, ervoor zorgt dat de participanten minder ziek worden.

Deze twee bevindingen resulteren samen in de bevinding dat een betere taalbeheersing positief samenhangt met een betere interpretatie van gezondheid outputs van medische professionals, wat daarna weer samenhangt met minder ziekte per jaar bij de respondenten. Dit betekent dus ook dat het omgekeerde waar is en bij een slechtere taalbeheersing er door middel van de lagere interpretatievermogens van gezondheid outputs mensen vaker ziek worden. Hierdoor kan er met een 95% betrouwbaarheidsinterval hypothese zeven voorlopig en gedeeltelijk worden aangenomen wat dit genoemde effect beschrijft. De voorlopige notie slaat erop dat er veel kanttekeningen zijn gezet in hoofdstuk 4.4 van dit stuk waardoor er de nodige robuustheidstesten moeten worden uitgevoerd om enig betrouwbaar verband te bevestigen. En de hypothese wordt gedeeltelijk aangenomen omdat alle taalvariabelen significant zijn in het verband maar alleen de gezondheidsvariabele *Hoeveelheid ziek* een verband heeft met het mechanisme en niet met de andere twee gezondheidsvariabelen.

Verdergaand heeft alleen de taalvariabele *Anderstalig* een significant verband met het mechanisme *Marginalisatie* met een 95% betrouwbaarheidsinterval. Dit verband is positief en heeft een waarde van 0.201 met SES mechanismes en 0.203 zonder, wat inhoudt dat mensen die anderstalig zijn gemiddeld 20%-punt vaker aangeven dat ze behoefte hebben aan sociale contacten dan mensen die niet anderstalig zijn. Dit is zeer interessant, doordat de variabele *Laaggeletterd* niet significant is en lager van waarde, en de achterliggende theorie zich voornamelijk richt op discriminatie door anderstaligheid. Het genoemde effect geeft bewijs voor de redentatie en de achterliggende MST omdat de variabele die de sociale taalmoeilijkheden vermeldt zo een ander effect laat zien dan de variabele die alleen moeite met lezen en schrijven aantoont. Hierdoor heeft dit verband wezenlijk bewijs geleverd voor hypothese een, waarin het verschil in het effect van anderstaligheid tegenover laaggeletterdheid

op de gezondheid van respondenten wordt besproken. Alleen heeft de variabele *Marginalisatie* geen enkel significant effect op de gezondheidsvariabelen, waardoor het louter blijft bij enig bewijs voor de eerste stap in het beredeneerde effect bij taalvariabele *Anderstalig*.

Het mechanisme *Sport* heeft in geen enkele instantie een significant verband met de taalvariabelen, de hoeveelheid dat mensen ziek worden of het geluk dat mensen voelen in hun leven. Wel heeft het voor elk model een positief significant verband met de variabele *Gezondheid* op een betrouwbaarheidsniveau van 95%. Dit verband houdt in participanten die vaker sporten per niveau van de vraag afhankelijk van het gebruikte model gemiddeld 0.300 tot 0.306 hoger scoren voor de variabele *Gezondheid*. Dit heeft als betekenis dat het vaker sporten in alle modellen leidt tot een betere gezondheid, wat bewijs levert voor het effect van het mechanisme op de gezondheid van participanten. Maar door het missende effect van taal op het mechanisme is er geen significant bewijs voor hypothese zes in deze analyse.

De variabelen die het SES mechanisme beschrijven laten meerdere effecten zien, zo hangen de taalvariabelen samen met *Educatie* en *Besteden* en is er geen enkel verband met of vanuit *Rondkomen*. *Educatie* hangt negatief samen met *Laaggeletterd* en *Anderstalig* op een 95% betrouwbaarheidsinterval met respectievelijk een waarde van 0.817 en 0.517. Dit houdt in dat als een individu laaggeletterd of anderstalig is er gemiddeld een lagere educatie aan samenhangt tegenover de groepen die dit niet zijn wat een positief effect van taalvermogen op educatie laat zien. Dit effect wordt bevestigd door *Tevredenheid taalvermogen* door de met 95% betrouwbare positieve verband met *Educatie* met de waarde 0.317 wat betekent dat ruwweg per drie treden van de 9-punten schaal van *Tevredenheid taalvermogen* de participanten een hoger afgerond onderwijsniveau aangeven. Dit bevestigt de bevinding afkomstig uit het verband tussen *Laaggeletterd* en *Anderstalig* met *Educatie* en toont een positief verband tussen beter taalvermogen en geleerdheid, wat in zijn beurt een positieve afspiegeling is van SES. Deze positieve afspiegeling wordt echter gecontrasteerd door het positieve verband tussen *Anderstalig* en *Laaggeletterd*, en *Besteden*, wat betekent dat participanten met taalproblemen meer te besteden hebben dan participanten zonder taalproblemen. Respectievelijk zorgen *Laaggeletterd* en *Anderstalig* voor de gemiddeld 0.565 en 0.975 hogere score op de vraag hoeveel de participanten te besteden hebben per week met een 95% en 99% betrouwbaarheidsinterval. Daarnaast laat *Tevredenheid taalvermogen* een negatief verband met *Besteden* zien van -0.231 met een betrouwbaarheidsinterval van 99%, wat betekent dat per ruwweg vier treden van de variabele *Tevredenheid taalvermogen* participanten de interval van hun wekelijks te besteden inkomen met vijftig euro verlagen. Dit betekent in de beredeneerde interpretatie dat vanuit het mechanisme *Besteden* participanten met een beter taalvermogen een

lagere SES hebben. Daarnaast is het ook waardevol om te vermelden dat er degelijke verschillen zijn in de effecten van *Anderstalig* en *Laaggeletterd* voor beide variabelen, waar bijvoorbeeld bij anderstaligheid gemiddeld een hoger bestedingsniveau uitkomt dan laaggeletterdheid, en er een lagere educatie uitkomt bij laaggeletterdheid dan anderstaligheid. Dit geeft bewijs voor de bewering dat laaggeletterdheid een lager SES tot uitkomst heeft dan anderstaligheid omdat ze én een negatiever effect op educatie hebben én een lager positief effect op besteding hebben dan de anderstaligen.

Het effect van de SES mechanismes op de gezondheidsvariabelen laten ook aparte verbanden zien. Naast dat voor vrijwel alle gezondheidsvariabelen nergens een verband is met de SES variabelen, hangt het educatieniveau positief samen met de hoeveelheid dat men jaarlijks ziek wordt. Dit effect heeft alleen een 90% betrouwbaarheidsniveau en geeft geen definitieve directie van het verband, maar heeft wel een positief gemiddeld effect van 0.098 tot 0.099 per trede, afhankelijk van de gebruikte taalvariabele op de hoeveelheid dat de participanten ziek worden. Dit verband laat zien dat mensen die WO opgeleid zijn bijna een trede hoger score op *Hoeveelheid ziek* dan mensen die geen educatie hebben gevolgd, wat een effect is dat niet is voorspeld en ook niet was verwacht. Dit effect laat ook zien dat mensen met een hoger SES een slechtere gezondheid hebben, wat tegen de bevindingen vanuit de FCT ingaan en ook tegen de daardoor opgestelde hypothese twee. Deze bevinding geeft bewijs tegen hypothese twee, maar geeft door de zwakke verbanden geen substantieel bewijs voor de verwerping van de hypothese.

Wat verder interessant is aan de toevoeging van de SES mechanismes aan het model, is dat er niet veel verandert aan de verbanden. In geen enkele in tabel B5.3 zichtbare instantie is er een wisselende significantie te zien voor elke gebruikte taalvariabele X_n als de SES mechanismes worden toegevoegd. Wel zijn er wisselingen in de coëfficiënten te zien bij het verband van enkele mechanismes op de gezondheidsvariabelen Y_j . Deze wisselingen zijn alleen vrij klein van aard en verschillen niet in de bandbreedte van de schatting ten overstaande van die zonder SES mechanismes.

Het ontbreken van flinke verschuivingen door toevoeging van de SES mechanismes kan meerdere dingen betekenen voor de uitkomsten. Enerzijds kan het betekenen dat door de weinige en laag significante verbanden die er in de SES mechanismes überhaupt aanwezig zijn deze variabelen weinig unieke variantie verklaren en daarom geen storing ontstaat erdoor. Dit biedt dan enige ondersteuning voor de eerder gedeelde gedachte dat de onderzoeksgroep homogeen is in SES, waardoor verdere toevoeging van SES variabelen geen substantiële variantie meer kunnen verklaren omdat er weinig verschillen binnen de onderzoeksgroep zijn

op dat vlak. Anderzijds kan het betekenen dat de waargenomen verbanden robuuster zijn, omdat ze minder vatbaar zijn voor mogelijk weggelaten variabelen en modelaanpassingen. De precieze oorzaak is niet af te leiden van de uitkomsten, maar de tweede notie zal verder grondig worden getest in de hierop volgende robuustheidstesten.

5.3 Robuustheidstesten

Ter ondersteuning van de validiteit en interpretatie van de bevindingen worden in dit gedeelte de uitkomsten van de robuustheidstesten besproken. De gebruikte testen zijn besproken in alinea 4.5 en zijn opgesteld om te verifiëren of de gevonden verbanden standhouden onder alternatieve modellen en aannames.

De resultaten van de Seemingly Unrelated Regression (SUR), vertoond in tabel B5.4 en B5.5, tonen geen wezenlijke afwijkingen ten opzichte van de schattingen in tabel B5.3. Tabel B5.4 laat de drie verschillende regressies voor elke taalvariabele zien en de coëfficiënten waarmee ze vanuit de SUR test de mechanismes beïnvloeden. Tabel B5.5 laat daarentegen in maar één regressie zien hoe de mechanismes met de SUR test de gezondheidsvariabelen beïnvloeden. De afwezigheid van afwijkingen ten opzichte van tabel B5.3 houdt in dat er geen systematische fouten aanwezig zijn als gevolg van het negeren van de onderlinge samenhang tussen vergelijkbare afhankelijke variabelen, zoals in de oorspronkelijke regressie- en SEM-modellen het geval is. Het gebruik van de SUR analyse toont daarmee aanvullend bewijs voor de robuustheid van de eerdere effecten gevonden in tabel B5.3 doordat het aantoont dat deze bevindingen niet voortkomen uit modelruis of specificatiefouten door gebruik van de SEM modellering.

De tweede robuustheidstoets onderzoekt het effect van het opnemen van de schattingsmethode `method(mlmv)` in het SEM-model. De bevindingen van deze robuustheidstest zijn te zien in tabel B5.6 en is te lezen zoals tabel B5.3. In het bijzonder in modellen waarin SES wordt gemodelleerd als mechanisme, blijken enkele verbanden minder sterk of geheel afwezig. Zo neemt de significantie van het effect van de variabelen *Laaggeletterd* en *Anderstalig* op *Educatie* af, hoewel *Educatie* nog steeds op het 99%-betrouwbaarheidsniveau samenhangt met *Tevredenheid taalvermogen* als taalvariabele.

Opvallend is dat het eerder waargenomen verband tussen *Educatie* en *Hoeveelheid ziek* niet meer aanwezig is. Deze bevindingen vormen bewijs tegen de eerder besproken tegenstrijdige resultaten uit het hoofdmodel, waarin een negatief verband werd gevonden tussen educatie en gezondheid, in tegenstelling tot wat op basis van de theorie verwacht werd.. Daarnaast verliest het verband tussen *Sport* en *Gezondheid* haar significantie voor alle modellen die de SES variabelen gebruiken, en is deze slechts significant op het lagere 90% betrouwbaarheidsniveau in modellen zonder SES.

Om deze schommelingen te duiden is onderzocht welke respondenten zijn uitgesloten door het ontbreken van SES-gegevens bij gebruik van method(mlmv) en de variabelen in de aangetaste verbanden samen te vatten binnen deze groep. De uitkomsten van deze samenvatting is te zien in tabel B5.7. Hieruit is bevonden dat deze respondenten verschillen van de oorspronkelijke steekproef, zoals te zien is in de beschrijvende statistieken van tabel B5.1, doordat ze minder vaak laaggeletterd maar vaker anderstalig zijn. Daarnaast ligt hun opleidingsniveau gemiddeld bijna een volledig niveau lager, sporten zij vaker, en vertonen zij vergelijkbare scores op de variabele *Hoeveelheid ziek* in vergelijking met de oorspronkelijke groep.

Hieruit volgt het vermoeden dat het gebruik van method(mlmv) leidt tot een selectie-effect dat de steekproef beïnvloedt. De uitgesloten respondenten vertonen juist karakteristieken die in lijn zijn met de hypothesen van dit onderzoek, wat impliceert dat het uiteindelijke model mogelijk conservatief is in de schatting van verbanden. Tegelijkertijd toont deze bevinding aan hoe gevoelig de resultaten zijn voor kleine veranderingen in de onderzoeksgroep, aangezien het wegvallen van acht respondenten aanzienlijke verschuivingen veroorzaakt.

De derde robuustheidsanalyse betreft het opnemen van directe paden tussen de taalvariabelen en de gezondheidsuitkomsten binnen het SEM-model. Deze test is opgenomen in tabel B5.8. Hierbij komt naar voren dat de verschillende taalvariabelen uiteenlopende effecten hebben. De *Tevredenheid taalvermogen* variabele laat in deze test bijvoorbeeld geen direct significant verband zien met de gezondheidsvariabelen, wat conform is aan de regressies van tabel B5.2 waarin geen enkel verband tussen *Tevredenheid taalvermogen* en de gezondheidsvariabelen significant was. De toevoeging van de taalvariabele zorgt wel voor een minder significant verband van *Dokter* op *Hoeveelheid ziek* voor de analyse met én zonder de SES variabelen, waarin het nu een betrouwbaarheidsniveau van 90% heeft in plaats van 95%.

De taalvariabele *Anderstalig* heeft, in tegenstelling tot *Tevredenheid taalvermogen*, wel een direct verband met een van de gezondheidsvariabelen doordat het direct *Hoeveelheid ziek* beïnvloedt. Dit is onder een positief verband dat significant is op het 90%-niveau in modellen

met SES en op het 95%-niveau in modellen zonder SES. Dit wijst erop dat anderstaligheid leidt tot een hogere zelf gerapporteerde ziektelast.

De opname van deze directe paden leidt verder tot het verdwijnen en ontstaan van nieuwe verbanden in vergelijking tot die uit tabel B5.3. Zo ontstaat er een positief significant effect van *Dokter* op *Gezondheid* in de analyse met en zonder SES, terwijl er in de analyse met SES het eerdere verband tussen *Anderstalig* en *Educatie* niet meer significant is. Verder laat deze robuustheidstest bij de verbanden van de taalvariabelen *Laaggeletterd* het vergelijkbare effect zien dat de significantie wegvalt in het verband tussen *Dokter* en *Hoeveelheid ziek*, maar zonder het bij *Anderstalig* geobserveerde verband tussen *Laaggeletterd* en *Hoeveelheid ziek*. Verder is bij het model met SES weer een met 90% significantieniveau verband ontstaan tussen *Dokter* en *Gezondheid* wat niet wordt gedeeld in het model zonder SES.

In deze robuustheidstests valt de significantie van de variabele *Dokter* op *Hoeveelheid ziek* deels tot volledig weg, wat betekent dat het directe verband tussen de taal- en gezondheidsvariabelen vermoedelijk de statistische kracht van dat verband overneemt. Dit vermoeden kenmerkt het eventuele bestaan van multicollineariteit tussen de taalvariabelen en *Dokter* in het verband met *Hoeveelheid ziek* en *Gezondheid*. Als dit zo zou zijn zou de eerder gemaakte presumptie, dat de codering van de variabele *Dokter* het interpreteren van gezondheids outputs afkomstig van medische professionals bevat, niet kloppen omdat het dan waarschijnlijk afhangt van taal als er teveel multicollineariteit aanwezig is. Hierdoor is een test naar de multicollineariteit genoodzaakt, waardoor een Variance Inflation Factor (VIF) test is uitgevoerd en afgebeeld in tabel B5.9. Deze test is alleen gedaan met *Hoeveelheid ziek* en *Gezondheid*, omdat in *Geluk* geen veranderingen waren bij de toevoeging van het directe verband tussen de taal- en gezondheidsvariabelen.

In de VIF test is een waarde van 1.37 tot 1.62 te zien, wat betekent dat de toevoeging van het directe verband tussen taal en gezondheid zorgt voor een groei van variantie in de *Dokter* variabele van 1.37 tot 1.62 keer de variabiliteit zonder die toevoeging. Dit komt overeen met een beperkte tot matige multicollineariteit tussen de taalvariabelen en *Dokter* in het verband met *Hoeveelheid ziek* en *Gezondheid*, wat betekent dat overlap van verklaarde variantie in de twee variabelen bij de schatting van de gezondheidsvariabelen beperkt zijn. Dit geeft daarbij ook weer robuustheid voor de redenatie achter de variabele *Dokter* dat het niet een taal deficiëntie aantoont maar een slechter interpretatievermogen van de gezondheidsoutputs van medische professionals. Wel betekent het ontbreken van multicollineariteit bij de geobserveerde effecten op *Dokter* bij de toevoeging van een direct verband tussen taal- en gezondheidsvariabelen dat er een inhoudelijke reden achter zit waardoor het directe effect

mogelijk een weggelaten verband in het oorspronkelijke model was en dus een vorm van omitted variable bias is.

In de laatste robuustheidstest wordt de variabele Arabisch eerst als vervangende X-variabele gebruikt en vervolgens als mechanisme, wat wordt afgebeeld in tabel B5.10. Tijdens de analyse waarin Arabisch als X-variabele wordt opgenomen zonder de SES-variabelen, blijken de uitkomsten vrijwel identiek aan de modellen waarin *Anderstalig* en *Laaggeletterd* als taalvariabelen zijn opgenomen. Dit is niet per se verrassend, aangezien *Anderstalig* mede gebaseerd is op het spreken van een andere moedertaal dan het Nederlands, terwijl *Arabisch* specifiek verwijst naar het spreken van Arabisch als moedertaal. In het model met SES-variabelen blijkt *Arabisch* een significant verband te hebben met *Rondkomen*, in plaats van het in het standaardmodel gevonden verband tussen *Anderstalig* en *Educatie*.

Bij het opnemen van *Arabisch* als mechanisme zijn er daarentegen meerdere interessante effecten waarneembaar in de modellen met en zonder SES. Zo heeft het als mechanisme een statistisch significante relatie met de oorspronkelijke taalvariabelen en heeft het een significant positief effect op *Geluk* op het 95%-betrouwbaarheidsniveau. Bovendien leidt de toevoeging van *Arabisch* tot een significant verband op het 90%-niveau tussen *Dokter* en *Gezondheid* in het model met SES-variabelen, zonder dat de significantie van andere mechanismen daardoor afneemt. Dit versterkt het vermoeden dat er in het oorspronkelijke model, met name in de schakels tussen de mechanismen, sprake is van omitted variable bias die verder zullen worden toegelicht in de discussie.

Dat vermoeden wordt verder onderbouwd door de verschuivingen die optreden bij het toevoegen van een extra variabele, zoals ook zichtbaar was in robuustheidstest drie. De oorspronkelijke aanleiding voor deze robuustheidstest, het nagaan of het invullen van de enquête in het Arabisch voor vertekening van de uitkomsten heeft gezorgd, kan met de beschikbare data niet direct worden geverifieerd. Toch wijzen de uitkomsten sterk op het bestaan van omitted variable bias, en vormt dit dus een plausibele verklaring.

5.4 Discussie

Deze analyse toont meerdere uitkomsten en complicaties van het verband tussen taal en gezondheid, gemedieerd door mechanismes op basis van beschikbare middelen, bij

participanten aan het Instapeconomie project. De grootste bevindingen zijn de verbanden die zijn gesignaleerd omtrent *Dokter* en *Stress*. *Dokter* toont zichzelf als een mechanisme van het verband tussen taal en gezondheid omdat het ondersteuning levert voor de redentatie dat mensen met een lagere Nederlandse taalbeheersing door verminderde interpretatie van gezondheids outputs afkomstig van medische professionals jaarlijks vaker ziek zijn, en levert daarmee bewijs voor de waarheid van hypothese zeven van dit onderzoek.

Deze bevinding komt overeen met de redentatie van Clarke en Isphording (2017) dat onbegrijpelijk taalgebruik zorgt voor distantiëring en zo een slechtere behandeling, wat daarna leidt tot het meer ziek zijn. Dit verband wordt echter verstoord door het directe effect vanuit de taalvariabelen op de gezondheidsvariabelen. Alleen is er door het testen voor multicollineariteit wel bewijs geleverd voor *Dokter* tegen het zijnde van een andere formulering van de taalvariabelen, waardoor de validiteit van de variabele op zich standhoudt. Wel is er door het wegvallen van het verband door het directe X->Y effect enige twijfel over de validiteit van de bevinding waardoor hypothese niet zal worden aangenomen noch verworpen. Dit betekent dat de bevonden verbanden louter als bewijs voor een mogelijk verband zullen worden gezien.

Het verband rond *Stress* daarentegen is niet volledig bewezen, maar toont wel potentie om te werken als mechanisme doordat het in elk uitgevoerde model een significant effect heeft gehad op alle gezondheidsvariabelen. Hierdoor kan vrij zeker worden vastgesteld dat *Stress* een belangrijke determinant van gezondheid is. Maar doordat alleen *Laaggeletterd* enig verband heeft met *Stress*, wat daarbij statistisch significant is op een 10%-niveau wat door de robuustheidstesten heen stand houdt waardoor het wel enig bewijs levert voor een effect van laaggeletterdheid op de stressniveaus en zo voor hypothese vier.

Het bewijs voor deze twee verbanden betekent daarbij ook dat de data ondersteuning biedt voor de waarheid van hypothesen acht en negen. Want deze twee verbanden laten zien dat er meerdere mechanismes zijn waardoorheen taalbeheersing de gezondheidsverschillen verklaren doordat de twee mechanismes potentie laten zien voor waarheid. Daarnaast is er enig bewijs dat taalvaardigheid samenhangt met meerdere gezondheidsuitkomsten door de samenhang van *Laaggeletterd* met *Stress* die daarna sterk samenhangt met de frequentie van ziekte, geluk en zo mentale gezondheid, en algemene zelf gerapporteerde gezondheid. Dit produceert verdere ondersteuning voor de veronderstellingen van het FCT raamwerk van Link en Phelan (1995), en zo voor het in dit onderzoek gehanteerde theoretisch raamwerk.

Verder laten de besproken verschillen in de effecten van de modellen met *Anderstalig* en *Laaggeletterd* ook zien dat de twee variabelen andere verbanden vertonen van elkaar doordat ze afzonderlijk van elkaar significante verband hebben met meerdere losse mechanismes, wat

bewijs levert voor hypothese een en zo de veronderstelling dat sociale en literaire taalvermogens andere effecten teweeg brengen op mensen hun gezondheid. Verder zijn er ook meerdere verbanden bevonden tussen taal en gezondheid, alleen zijn deze uiteenlopend in richting en wisselend per gebruikte model waardoor er geen echt bewijs voor hypothese twee is vastgesteld.

Het blijft alleen wel bij bewijs voor de hypothesen omdat de uitkomsten niet genoeg uitsluitend bieden en de robuustheidstesten zorgen voor extra onzekerheid. De robuustheidstesten hebben laten zien hoe gevoelig de uitkomsten zijn doordat het wegnemen van method(mlmv) zorgt voor veel veranderingen in de modellen. Dit geeft aan dat de dataset te klein is voor de complexiteit van dit onderzoek en dat het enorm volatiel is door dit gegeven.

Daarnaast is er bij het toevoegen van de extra variabelen in de mechanisme laag door een direct verband van X op Y en de variabele *Arabisch* zodanig veel verstoring opgemerkt dat er een sterk vermoeden van het omitted variable bias is ontstaan. Enerzijds is de enquête voor het tot stand komen van dit onderzoek geproduceerd, waardoorde vragen niet zijn afgestemd op de gewenste verbanden voor dit onderzoek. Dit zorgt ervoor dat controle voor bijvoorbeeld het hebben van een chronische ziekte, of het hebben van een familielid die de tijd en vaardigheden heeft om constant te vertolken niet mogelijk was. Verder heeft dit onderzoek niets kunnen doen met het mogelijke effect dat het taalniveau van Engels op de bevindingen heeft. Want Nederland staat al jaren in de top van best Engels sprekende landen waarin het Engels geen moedertaal is, dus dat heeft waarschijnlijk een significant effect op de uitkomsten van dit onderzoek.

Anderzijds zijn mede doordat er geen controlevariabelen aanwezig zijn om de haalbaarheid van het al gecompliceerde rekenmodel te behouden, veel mogelijke alternatieve verklaarders vanuit persoonlijke en sociale kenmerken. Voorbeelden hiervan zijn persoonskenmerken als geslacht, leeftijd en huwelijksstatus, of eventuele open vragen die vragen wie men het meeste helpt. Door het ontbreken hiervan zal er zeker storing en onverklaarde variantie overblijven in het model, wat kan worden opgelost in de context van dit model als er meer enquêtes zijn ingevuld.

Deze complicaties leveren wel kansen voor vervolgonderzoek, want het gehanteerde SEM model levert in veel opzichten bewijs voor een mogelijk verband tussen Nederlands taalvermogen en gezondheidsuitkomsten dat overeenkomstig is met de FCT literatuur. Bij reproductie van dit onderzoek, als er meer onderzoekslichtingen zijn toegevoegd aan de dataset, wordt de statistische kracht flink versterkt waardoor eventuele verbanden die nu nog mogelijk kans zijn dan kunnen worden bevestigd. Verder kan er bij een grotere dataset ook meer nadruk

worden gelegd op controlevariabelen, waardoor validiteit wordt vergroot en er meer diepte aan het besproken verband kan worden toegevoegd. Zo biedt dit onderzoek een waardevolle basis waarop in de toekomst onderzoek kan voortbouwen om het inzicht in de relatie tussen taalvaardigheid en gezondheidsuitkomsten verder te verdiepen en te robuuster te maken.

6. Conclusie

In deze scriptie is onderzocht wat het effect is dat door mechanismes loopt van taal naar gezondheid, met als hoofdvraag: “*Wat is het effect van een lagere beheersing van de Nederlandse taal op de gezondheid van participanten aan de Instapeconomie in Den Haag regio Zuidwest en omstreken?*”. Kijkend naar de analyse zijn er twee mechanismes naar voren gekomen die een antwoord bieden op deze vraag. Conform aan het werk van Clarke en Isphording (2017), schetst de variabele *Dokter* bewijs voor het verband dat het kunnen interpreteren van gezondheid inputs zorgt voor minder ziekte. Dit brengt bewijs voor de theorie dat iemand met een verminderd taalvermogen minder goed met een dokter kan converseren, waardoor ze mogelijke adviezen minder goed kunnen uitvoeren en daardoor vaker ziek zijn. Verder laat de variabele *Stress* bewijs zien voor de negatieve effecten op de gezondheid bij het ervaren van stress en is er conform de theorie enig bewijs gevonden voor het verband tussen taal en het verkrijgen van stress door het verband met *Laaggeletterd*. Dit bevestigt een verband tussen stress en gezondheid, en brengt enig bewijs voor het verband tussen taal en ervaren stress.

Deze bevindingen blijven alleen bewijs in de context van dit onderzoek door de gebreken aan de datakwaliteit en lagere onderzoekspopulatie van dit onderzoek. Doordat er veel onzekerheden blijven onder de variabeleselectie, die zorgen voor meerdere alternatieve verklaarders zoals het effect van het Engelse taalvermogen onder de participanten en enkele variabelen die onzekerheden hebben of ze het beoogde verband vastleggen, wordt daarin de mogelijkheid tot het vaststellen van causaliteit gecompromitteerd. Daarnaast is het model zeer volatiel gebleken, wat aangeeft dat de onderzoekspopulatie te klein was voor het ingebrachte model.

Desondanks kan er worden gesteld dat er bewijs is geleverd voor de bevindingen dat taal een effect heeft op de gezondheid van participanten aan de Instapeconomie doordat de bevindingen wijzen naar een verband dat er onder participanten met een lager taalvermogen

meer stress ervaren wordt en minder goed het advies van de dokter wordt geïnterpreteerd. Dit heeft veel potentie voor het begrip van de situatie van de participanten. Hieruit is bewijs geleverd dat wijst naar de extra moeilijkheden van taal onder langdurig werklozen. Dit bewijs is waardevol voor het begrip in de benodigdheden van deze groepen, omdat taal in meerdere aspecten van de meervoudige problematiek een versterkende rol kan spelen, doordat communicatie door een verminderd taalvermogen slechter is. Daarnaast laat dit onderzoek ook de potentie zien voor het zijn van een fundamentele veroorzaker van gezondheidsverschillen van het taalvermogen. Er is bewijs voor de aanwezigheid van meerdere mechanismes die ervoor zorgen dat meerdere aspecten van gezondheid worden aangetast.

Voor bevestiging van de potentie moet er wel verder onderzoek worden uitgevoerd met een grotere onderzoekspopulatie en meer flexibiliteit in de gestelde vragen. Hierom zijn de adviezen voor vervolgonderzoek dat er mogelijke herhalingen van dit onderzoek waarbij meerdere lichten aan het Instapeconomie project zijn onderzocht een beter beeld geven over de gepresenteerde verbanden van dit onderzoek. Door de grotere onderzoekspopulatie is de verwachting dat er minder volatiliteit ontstaat in het model bij modificaties. Daarnaast zou een grotere onderzoeksgroep ook de mogelijkheid tot longitudinaal onderzoek of toevoeging van controlevariabelen bieden, waardoor er meer validiteit aan de bevindingen kan worden toegevoegd en eventuele onderliggende verbanden beter kunnen worden blootgelegd.

Afsluitend is er het advies dat, door de potentie die in dit onderzoek is blootgelegd, verder onderzoek naar het zijn van taalvermogen als fundamentele veroorzaker van gezondheidsverschillen gegrond is. Verder onderzoek naar dit verband zal zorgen voor een bredere context en meer begrip in de gezondheidsproblematiek van taalbarrières, wat maatschappelijk gewenst is. Daarnaast zal een onderzoek naar het zijn van een fundamentele veroorzaker van gezondheidsverschillen bijdragen aan de waarheidsbevinding van de FCT als theorie. Dit verdere onderzoek kan dan het beste ingericht worden door middel van een toegespitste vragenlijst, waardoor meer alternatieve verklaarders kunnen worden geïmplementeerd die in dit onderzoek zijn benoemd.

7. Literatuurlijst

- Ballen, C. J., & Salehi, S. (2021). Mediation analysis in discipline-based education research using structural equation modeling: beyond “what works” to understand how it works, and for whom. *Journal of microbiology & biology education*, 22(2), 10-1128.
- Buisman, M., Bollen, I., Jacobs, B., Huijts, T., Cornelisse, R., Van Guilik, N., Elshof, D., & Van Griensven, L. (2024). *PIAAC 2023: Kernvaardigheden van volwassenen. Resultaten van de Nederlandse survey 2023*. Amstelveen: Kohnstamm Instituut. Rapport 1143.
- Centraal Bureau voor de Statistiek (CBS). (z.d.). Bevolking; onderwijsniveau; geslacht, leeftijd En migratieachtergrond. StatLine.
<https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37296ned/table>
- English, F. W. (Ed.). (2006). *Encyclopedia of educational leadership and administration*. Sage publications.
- Centraal Bureau voor de Statistiek (CBS). (2020). Jaarrapport Integratie 2020. CBS.
<https://www.cbs.nl/nl-nl/publicatie/2020/46/jaarrapport-integratie-2020>
- Centraal Bureau voor de Statistiek (CBS). (2024). Rapportage Integratie en Samenleven 2024. CBS.
<https://www.cbs.nl/nl-nl/publicatie/2024/47/rapportage-integratie-en-samenleven-2024>
- Chiswick, B. R., & Miller, P. W. (2001). A model of destination-language acquisition: Application to male immigrants in Canada. *Demography*, 38, 391-409.
- Chiswick, B. R., & Wang, Z. (2019). *Social contacts, Dutch language proficiency and immigrant economic performance in the Netherlands* (No. 419). GLO Discussion Paper.
- Christoffels, I., Baay, P., Bijlsma, I., & Levels, M. (2016). Over de relatie tussen laaggeletterdheid en armoede. Den Haag: Stichting Lezen & Schrijven i.s.m. ECBO.
- Clarke, A., & Ispording, I. E. (2017). Language barriers and immigrant health. *Health economics*, 26(6), 765-778.
- Clouston, S. A., & Link, B. G. (2021). A retrospective on fundamental cause theory: State of the literature and goals for the future. *Annual Review of Sociology*, 47(1), 131-156.
- Dekker, S. V., Duarte, J., & Loerts, H. (2021). ‘Who really speaks like that?’ –Children’s implicit and explicit attitudes towards multilingual speakers of Dutch. *International journal of multilingualism*, 18(4), 551-569.

- Dhabhar, F. S. (2014). Effects of stress on immune function: the good, the bad, and the beautiful. *Immunologic research*, 58, 193-210.
- Easton, P., Entwistle, V. A., & Williams, B. (2013). How the stigma of low literacy can impair patient-professional spoken interactions and affect health: insights from a qualitative investigation. *BMC health services research*, 13, 1-12.
- Fransen, M. P., Leenaars, K. E., Rowlands, G., Weiss, B. D., Maat, H. P., & Essink-Bot, M. L. (2014). International application of health literacy measures: adaptation and validation of the newest vital sign in The Netherlands. *Patient education and counseling*, 97(3), 403-409.
- Freese, J., & Lutfey, K. (2010). Fundamental causality: challenges of an animating concept for medical sociology. In *Handbook of the sociology of health, illness, and healing: A blueprint for the 21st century* (pp. 67-81). New York, NY: Springer New York.
- Frost, D. M., & Meyer, I. H. (2023). Minority stress theory: Application, critique, and continued relevance. *Current opinion in psychology*, 51, 101579.
- Gemeente Den Haag. (2022, februari). *Evaluatie pilot Instapeconomie*. Gemeente Den Haag. Geraadpleegd via het Raadsinformatiesysteem van de gemeente Den Haag op denhaag.raadsinformatie.nl/document/11197928/1/RIS311530_Bijlage_1_Evaluatie_pilot_Instapeconomie
- Gemeente Den Haag. (2023, maart). *Beleidsplan Werk en Participatie Den Haag 2023–2026*. Gemeente Den Haag. Geraadpleegd via denhaag.raadsinformatie.nl/document/12618395/1#page=6.99
- Het Kenniscentrum Arbeidsmigranten. (2023). *De Nederlandse taalvaardigheid onder arbeidsmigranten: De resultaten van het 6e arbeidsmigrantenpanel*. [https://hetkenniscentrumarbeidsmigranten.nl/wp-content/uploads/2023/10/Eindrapport-De-Nederlandse-taalvaardigheid-onder-arbeidsmigranten.pdf?#8203::contentReference\[oaicite:5\]{index=5}](https://hetkenniscentrumarbeidsmigranten.nl/wp-content/uploads/2023/10/Eindrapport-De-Nederlandse-taalvaardigheid-onder-arbeidsmigranten.pdf?#8203::contentReference[oaicite:5]{index=5})
- Van der Heide, I., & Rademakers, J. (2015). *Laaggeletterdheid en gezondheid: stand van zaken*. Nivel. <https://www.nivel.nl/nl/publicatie/laaggeletterdheid-en-gezondheid-stand-van-zaken>
- Inspectie van het Onderwijs. (2023). *Peil.Taal en rekenen einde basisonderwijs en speciaal*

(basis)onderwijs 2021-2022.

<https://www.onderwijsinspectie.nl/documenten/rapporten/2023/05/10/peil-taal-en-rekenen-2021-2022>

- Kadushin, C. (1964). Social Class and the Experience of Ill Health. *Sociological Inquiry*, 34(1).
- Kickbusch, I., & Maag, D. (2008). Health literacy. In K. Heggenhougen & S. Quah (Eds.), *International encyclopedia of public health* (pp. 204–211). San Diego, CA: Academic Press. <https://ilonakickbusch.com/kickbusch-wAssets/docs/kickbusch-maag.pdf>
- Kjøllestad, M. K., Gerwing, J., & Indseth, T. (2023). Proficiency in the Norwegian language and self-reported health among 12 immigrant groups in Norway: A cross-sectional study. *Scandinavian Journal of Public Health*, 51(3), 363-370.
- Lee, H., & Singh, G. K. (2021). Monthly trends in self-reported health status and depression by race/ethnicity and socioeconomic status during the COVID-19 Pandemic, United States, April 2020–May 2021. *Annals of Epidemiology*, 63, 52-62.
- van de Looij-Jansen, P., Tamerus, S., Baerveldt, C., & de Waart, F. (2014). De samenhang tussen de kwaliteit van de wijk en psychosociale problemen bij jongeren in Rotterdam. *Tijdschrift voor gezondheidswetenschappen*, 92, 163-168.
- Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of health and social behavior*, 80-94.
- Lutfey, K., & Freese, J. (2005). Toward some fundamentals of fundamental causality: Socioeconomic status and health in the routine clinic visit for diabetes. *American Journal of Sociology*, 110(5), 1326-1372.
- Ministerie van Onderwijs, Cultuur en Wetenschap. (z.d.). *Behaald onderwijsniveau*. OCW in cijfers. Geraadpleegd op 23/06/2025. <https://www.ocwincijfers.nl/sectoren/onderwijs-algemeen/behaald-onderwijsniveau>
- Motivaction. (2024, 9 september). *Betrek laaggeletterden in onderzoek voor betere overheidscommunicatie*. <https://www.motivaction.nl/actualiteiten/nieuwsberichten/betrek-laaggeletterden-in-onderzoek-voor-betere-overheidscommunicatie>
- Nationaal Programma Den Haag Zuidwest. (2023, juli). *Uitvoeringsplan Nationaal Programma Den Haag Zuidwest 2023–2027* (p. 3.99). NPZW. Geraadpleegd van <https://npzw.nl/wp-content/uploads/2023/09/Uitvoeringsplan-Zuidwest.pdf#page=3.99>
- Nifadkar, S., Tsui, A. S., & Ashforth, B. E. (2012). The way you make me feel and behave: Supervisor-triggered newcomer affect and approach-avoidance behavior. *Academy of Management Journal*, 55(5), 1146-1168.

- Nivel. (2024). *Kennissynthese 2024: Gezondheidsvaardigheden in Nederland – actuele kennis en inzichten*. Nivel. <https://www.nivel.nl/nl/publicatie/kennissynthese-2024-gezondheidsvaardigheden-nederland-actuele-kennis-en-inzichten>
- Non, M., Dinkova, M., & Dahmen, B. (2021). *Skill up or get left behind? Digital skills and labor market outcomes in the Netherlands* (CPB Discussion Paper No. 419). CPB Netherlands Bureau for Economic Policy Analysis. <https://www.cpb.nl/en/skill-up-or-get-left-behind-digital-skills-and-labor-market-outcomes-in-the-netherlands>
- Parikh, N. S., Parker, R. M., Nurss, J. R., Baker, D. W., & Williams, M. V. (1996). Shame and health literacy: the unspoken connection. *Patient education and counseling*, 27(1), 33-39.
- Pew Research Center. (2024, January 18). *Language and traditions are considered central to national identity*. <https://www.pewresearch.org/global/2024/01/18/language-and-traditions-are-considered-central-to-national-identity/>
- Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: theory, evidence, and policy implications. *Journal of health and social behavior*, 51(1_suppl), S28-S40.
- Phelan, J. C., & Link, B. G. (2015). Is racism a fundamental cause of inequalities in health?. *Annual Review of Sociology*, 41(1), 311-330.
- PwC (2018). *Maatschappelijke kosten laaggeletterdheid*. Amsterdam: PwC.
- Rademakers, J. (2014). *Niet voor iedereen vanzelfsprekend*. Utrecht: Nivel.
- Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. *Social science & medicine*, 90, 24-31.
- Rijk, M., & Lubbers, M. (2020). *Hebben recente Bulgaarse, Poolse, Spaanse en Turkse migranten hun draai gevonden in Nederland?* (NIDI Working Paper No. 2020/03). Nederlands Interdisciplinair Demografisch Instituut (NIDI). <https://nidi.nl/wp-content/uploads/2020/11/nidi-wp-2020-03.pdf>
- Schachter, A., Kimbro, R. T., & Gorman, B. K. (2012). Language proficiency and health status: are bilingual immigrants healthier?. *Journal of health and social behavior*, 53(1), 124-145.
- Schmeets, H., & Cornips, L. (2023). Het belang van taalkeuze voor het sociaal kapitaal in Nederland. *Dutch Journal of Applied Linguistics*, 12. <https://doi.org/10.51751/dujal13267>

- StataCorp LLC. (2025). Example 26 — Fitting a model with data missing at random (Stata Manual). College Station, TX: StataCorp LLC. Geraadpleegd op 21 juli 2025, van <https://www.stata.com/manuals/semexample26.pdf>
- Steunpunt Uitkeringen. (z.d.). *Bijstand en bijverdiene*n. Geraadpleegd op 9 juli 2025, van <https://www.steunpuntuitkeringen.nl/bijstand-en-bijverdiene/n/>
- Stichting Lezen en Schrijven. (z.d.). *Kennisblad Laaggeletterdheid in Nederland*. <https://www.lezenenschrijven.nl/over-laaggeletterdheid/informatie-over-laaggeletterdheid-nederland>
- Stichting Lezen & Schrijven, De Greef, M., Segers, M., & Nijhuis, J. (2014). *Feiten & cijfers geletterdheid: Overzicht van de gevolgen van laaggeletterdheid en opbrengsten van investeringen voor samenleving en individu*. Stichting Lezen & Schrijven
- Torres, Z., Tomás, J. M., Sentandreu-Mañó, T., Fernández, I., & Pla-Sanz, N. (2024). Social participation, loneliness, and physical inactivity over time: evidence from SHARE. *International psychogeriatrics*, 36(9), 799-807.
- UCLA Institute for Digital Research and Education. (z.d.). *What is seemingly unrelated regression and how can I perform it in Stata?* UCLA: Statistical Consulting Group. Geraadpleegd op 21 juli 2025, van <https://stats.oarc.ucla.edu/stata/faq/what-is-seemingly-unrelated-regression-and-how-can-i-perform-it-in-stata/#:~:text=A%20single%20model%20may%20contain,seemingly%20unrelated%20regression%20is%20sureg>
- de Vries, J., & Lautenbach, H. (2023). *Nederlandse taalvaardigheid van migranten en hun positie op de arbeidsmarkt*. Centraal Bureau voor de Statistiek. <https://www.cbs.nl/nl-nl/longread/statistische-trends/2023/nederlandse-taalvaardigheid-van-migranten-en-hun-positie-op-de-arbeidsmarkt>
- Wang, Z., De Graaff, T., & Nijkamp, P. (2018). Barriers of culture, networks, and language in international migration: A review. *Region*, 5(1), 73-89.
- Wieland, M. L., Tiedje, K., Meiers, S. J., Mohamed, A. A., Formea, C. M., Ridgeway, J. L., ... & Sia, I. G. (2015). Perspectives on physical activity among immigrants and refugees to a small urban community in Minnesota. *Journal of Immigrant and Minority Health*, 17, 263-275.
- Willems, A. E. M., Heijmans, M., Brabers, A. E. M., & Rademakers, J. (2022).

Gezondheidsvaardigheden in Nederland: factsheet cijfers 2021. Nivel.
<https://www.nivel.nl/nl/publicatie/gezondheidsvaardigheden-nederland-factsheet-cijfers-2021>

World Population Review. (2025). *English proficiency by country 2025.*

<https://worldpopulationreview.com/country-rankings/english-proficiency-by-country>

Zhang, Y. B., & Giles, H. (2018). Communication accommodation theory. *The international encyclopedia of intercultural communication, 1*, 95-108.

8. Bijlage

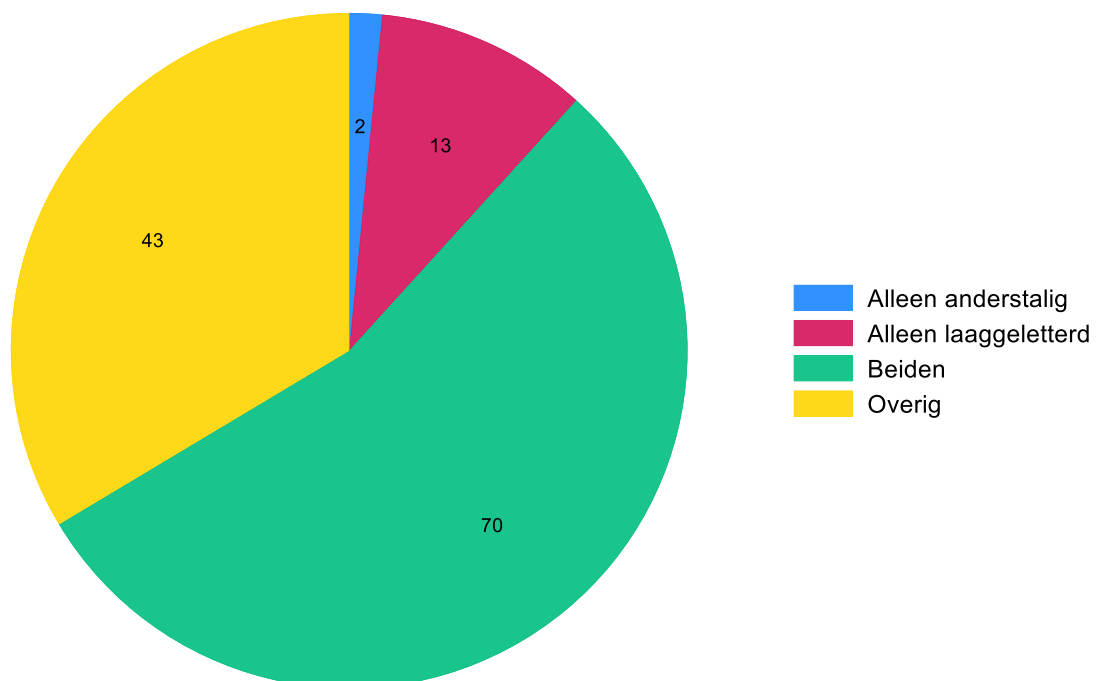
Tabel B4.1 Naam en codering mechanismevariabelen.

| Naam variabele | Gebruikte vraag | Antwoordopties & codering |
|----------------|----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Educatie | Vraag 38 "Wat is je hoogst afgeronde opleiding?" | 1="Geen" 2="Basisschool" 3=VMBO op de middelbare school" 4="HAVO op de middelbare school" 5="VWO op de middelbare school" 6="Middelbaar beroepsonderwijs (MBO)" 7="Hoger beroepsonderwijs (HBO)" 8="Wetenschappelijk onderwijs (WO)" 9="Doctoraal". |
| Besteden | Vraag 28 "Hoeveel kan je per week ongeveer besteden, naast je vaste lasten?" | 1= "0 tot 50 euro" 2= "50 tot 100 euro" 3="100 tot 150 euro" 4="150 tot 200 euro" 5= "200 euro of meer" |
| Rondkomen | Vraag 29 "Ervaar je uitdagingen met rondkomen?" | 9 punts schaal zoals in afbeelding X.X. 1="Ja, heel veel" & 9= "Nee, helemaal niet" |
| Stress | Vraag 25 "Hoeveel stress ervaar je in een week?" | 9 punts schaal zoals in afbeelding X.X. 1="Altijd" 9="Nooit" |
| Marginalisatie | Vraag 47 "Zou jij graag contacten willen met mensen in jouw omgeving?" | 1= "Ja" 0= "Nee" |
| Sport | Vraag 20 "Doe je aan sport activiteiten?" | 1= "Nooit" 2="1x in de maand" 3= "1x in de week" 4= "3-5x per week" 5="Elke dag" |
| Dokter | Vraag 10 "Kan je goed uitleggen aan de dokter als er iets is in het Nederlands?" | 0= "Nee, dat lukt niet" 1="Ja, maar zonder hele zinnen" 2= "Ja, makkelijk" |

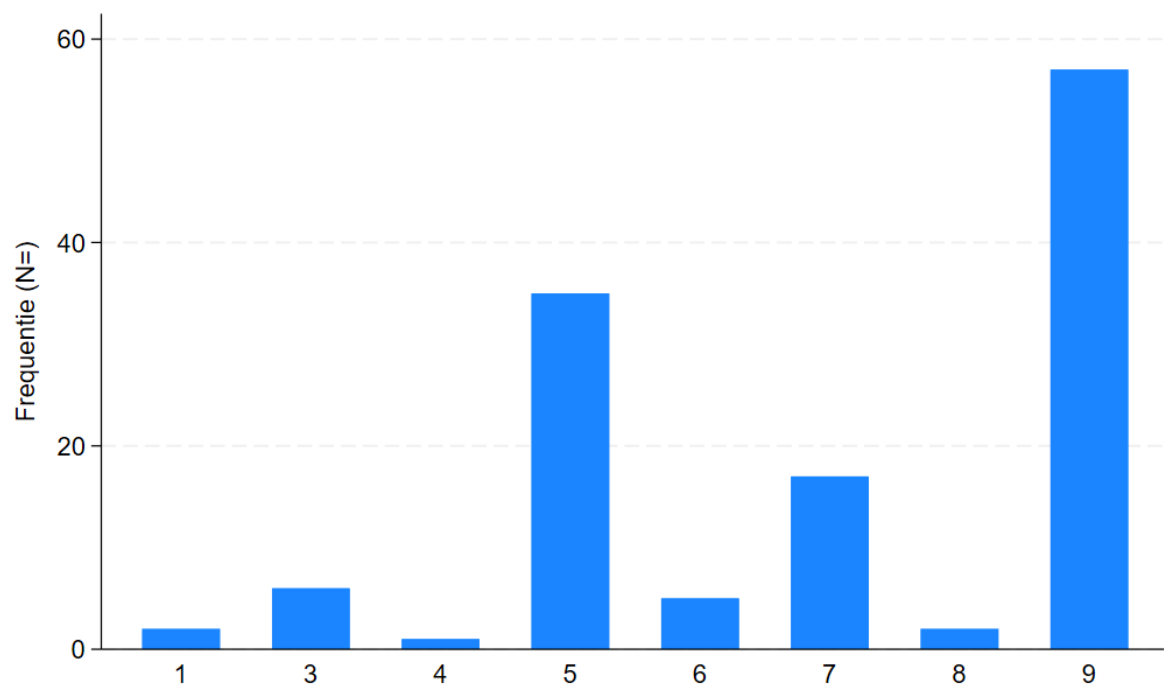
Tabel B5.1: Beschrijvende statistiek

| Variabele | Observaties | Gemiddelde | Standaarddeviatie | Minimum | Maximum |
|---------------------------|-------------|------------|-------------------|---------|---------|
| Laaggeletterd | 126 | 0.65873 | 0.476029 | 0 | 1 |
| Anderstalig | 128 | 0.5625 | 0.498028 | 0 | 1 |
| Tevredenheid taalvermogen | 125 | 7.016 | 2.109823 | 1 | 9 |
| Stress | 127 | 5.047244 | 2.232012 | 1 | 9 |
| Marginalisatie | 120 | 0.516667 | 0.501817 | 0 | 1 |
| Sport | 128 | 2.734375 | 1.41134 | 1 | 5 |
| Dokter | 127 | 1.417323 | 0.750036 | 0 | 2 |
| Educatie | 126 | 3.706349 | 2.194785 | 1 | 8 |
| Besteden | 123 | 2.691057 | 1.331399 | 1 | 5 |
| Rondkomen | 124 | 4.58871 | 2.288224 | 1 | 9 |
| Hoeveelheid ziek | 128 | 3.03125 | 1.309791 | 1 | 5 |
| Geluk | 128 | 6.601563 | 2.196417 | 1 | 9 |
| Gezondheid | 128 | 5.328125 | 2.578254 | 1 | 9 |

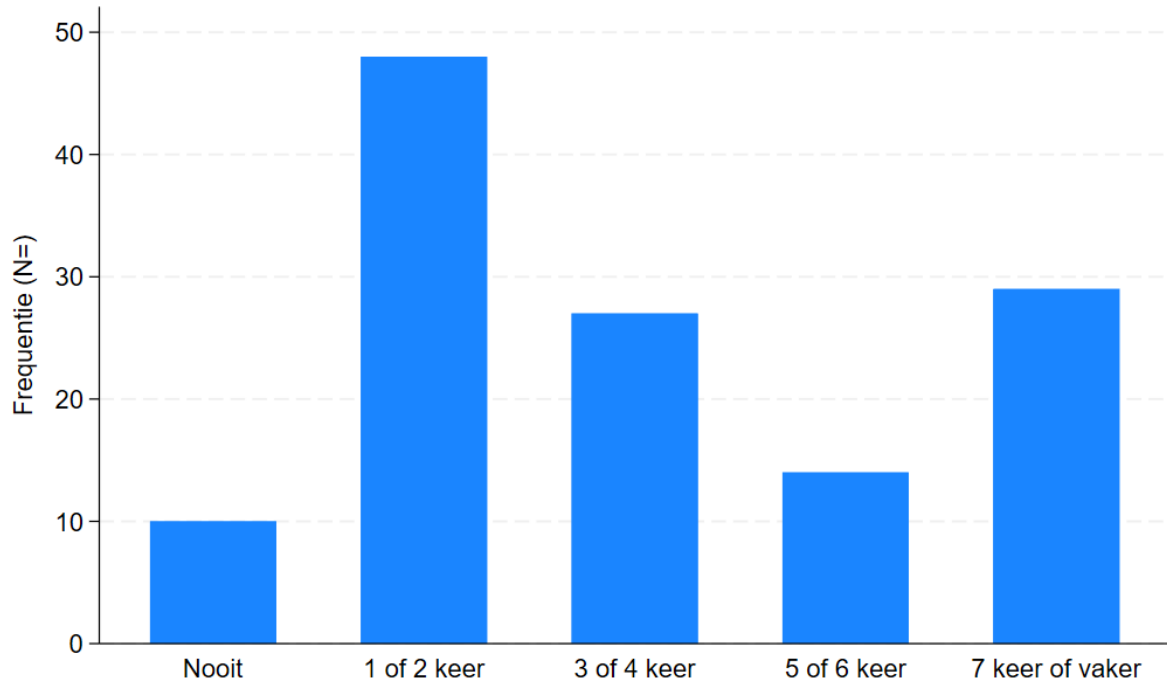
Figuur B5.1: Pie chart verdeling anderstalig en laaggeletterd



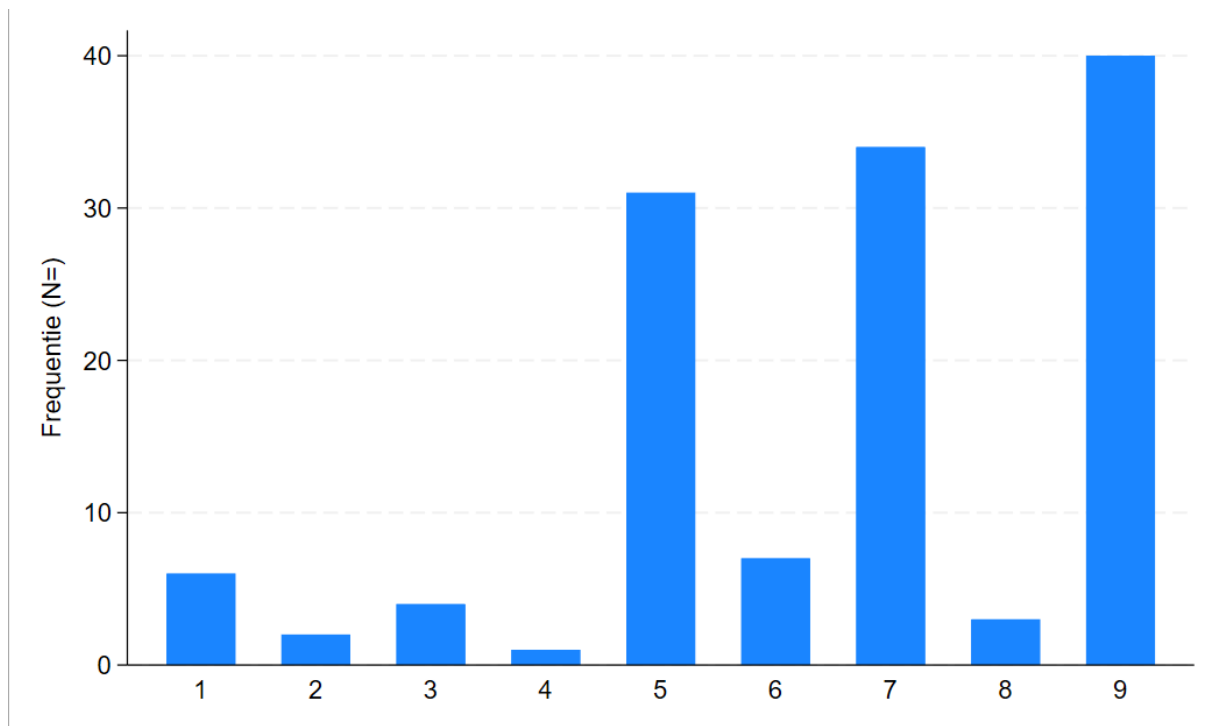
Figuur B5.2: Histogram verdeling antwoorden variabele Tevredenheid taalvermogen



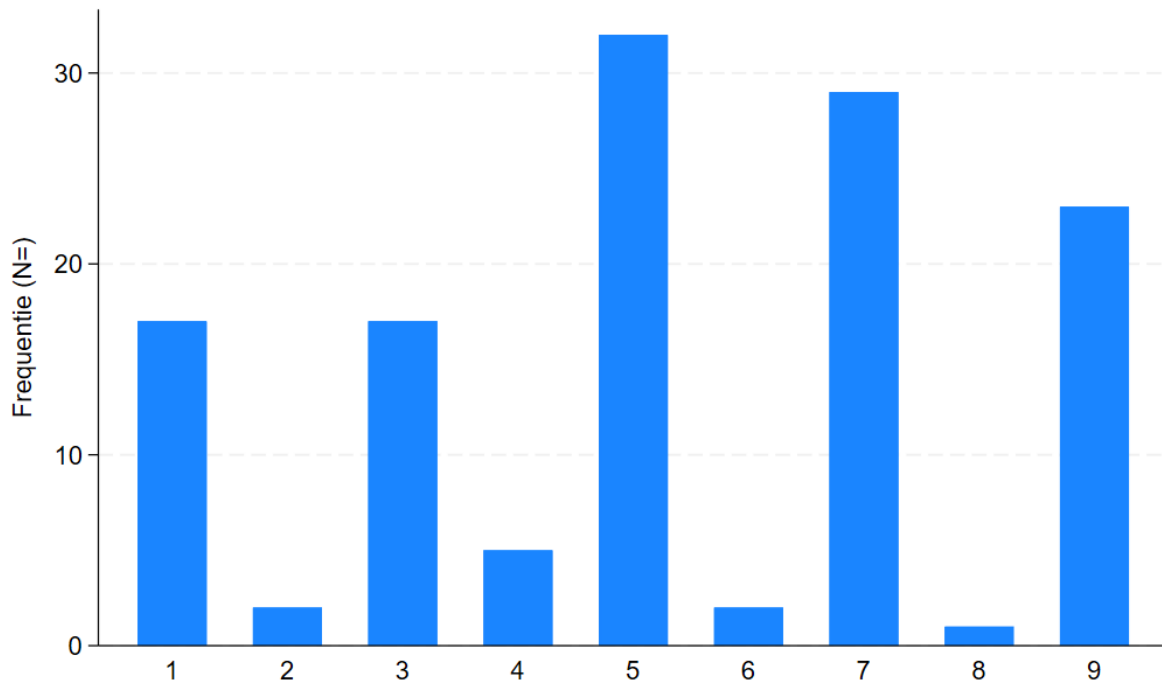
Figuur B5.3: Histogram verdeling antwoorden variabele Hoeveelheid ziek



Figuur B5.4: Histogram verdeling antwoorden variabele Geluk



Figuur B5.5: Histogram verdeling antwoorden variabele Gezondheid



Tabel B5.2: OLS regressies van alle X_n op alle Y_j .

| | Frequentie ziek | Geluk | Gezondheid |
|---------------------------|-----------------|---------|------------|
| Laaggeletterd | 0.518** | 0.856** | 0.411 |
| stdv | 0.242 | 0.408 | 0.488 |
| Anderstalig | 0.627*** | 0.625 | 0.298 |
| stdv | 0.228 | 0.389 | 0.460 |
| Tevredenheid taalvermogen | -0.073 | -0.118 | 0.111 |
| stdv | 0.056 | 0.093 | 0.110 |

Tabel B5.3: Bevindingen SEM basismodel.

| Effect X_n op mechanisme | Effect mechanisme op Y_j | Laaggeletterd | | Anderstalig | | Tevredenheid taalvermogen | |
|----------------------------|----------------------------|---------------|------------|-------------|------------|---------------------------|------------|
| | | Met SES | Zonder SES | Met SES | Zonder SES | Met SES | Zonder SES |
| Stress | | 0.710* | 0.737* | 0.502 | 0.506 | -0.053 | -0.053 |
| | | stdv | 0.414 | 0.412 | 0.395 | 0.395 | 0.094 |
| | Hoeveelheid ziek | -0.134** | -0.116** | -0.136** | -0.117** | -0.134** | -0.116** |

| | | | | | | | |
|----------------|------------------|-----------|-----------|-----------|-----------|----------|----------|
| | stdv | 0.055 | 0.051 | 0.055 | 0.051 | 0.055 | 0.051 |
| | Geluk | 0.473*** | 0.473*** | 0.473*** | 0.473*** | 0.474*** | 0.473*** |
| | stdv | 0.082 | 0.076 | 0.082 | 0.076 | 0.082 | 0.076 |
| | Gezondheid | 0.556*** | 0.547*** | 0.558*** | 0.549*** | 0.555*** | 0.547*** |
| | stdv | 0.098 | 0.091 | 0.098 | 0.091 | 0.098 | 0.090 |
| Dokter | | -0.820*** | -0.820*** | -0.932*** | -0.932*** | 0.215*** | 0.215*** |
| | stdv | 0.119 | 0.120 | 0.105 | 0.105 | 0.025 | 0.025 |
| | Hoeveelheid ziek | -0.389** | -0.323** | -0.387** | -0.323** | -0.392** | -0.323** |
| | stdv | 0.166 | 0.151 | 0.166 | 0.151 | 0.167 | 0.151 |
| | Geluk | -0.304 | -0.255 | -0.288 | -0.251 | -0.301 | -0.259 |
| | stdv | 0.251 | 0.226 | 0.250 | 0.226 | 0.251 | 0.226 |
| | Gezondheid | 0.455 | 0.327 | 0.475 | 0.338 | 0.450 | 0.320 |
| | stdv | 0.300 | 0.270 | 0.298 | 0.269 | 0.301 | 0.270 |
| Marginalisatie | | 0.135 | 0.138 | 0.201** | 0.203** | -0.018 | -0.018 |
| | stdv | 0.096 | 0.096 | 0.090 | 0.090 | 0.022 | 0.022 |
| | Hoeveelheid ziek | 0.176 | 0.223 | 0.188 | 0.242 | 0.173 | 0.220 |
| | stdv | 0.230 | 0.235 | 0.230 | 0.234 | 0.230 | 0.235 |
| | Geluk | -0.252 | -0.204 | -0.248 | -0.194 | -0.269 | -0.221 |
| | stdv | 0.350 | 0.351 | 0.350 | 0.351 | 0.349 | 0.350 |
| | Gezondheid | -0.302 | -0.326 | -0.306 | -0.328 | -0.312 | -0.335 |
| | stdv | 0.408 | 0.407 | 0.408 | 0.406 | 0.408 | 0.406 |
| Sport | | -0.128 | -0.149 | 0.036 | 0.036 | 0.061 | 0.062 |
| | stdv | 0.265 | 0.264 | 0.250 | 0.250 | 0.059 | 0.059 |
| | Hoeveelheid ziek | -0.047 | -0.027 | -0.046 | -0.027 | -0.046 | -0.028 |
| | stdv | 0.079 | 0.080 | 0.079 | 0.080 | 0.080 | 0.080 |
| | Geluk | -0.135 | -0.123 | -0.133 | -0.123 | -0.134 | -0.123 |
| | stdv | 0.120 | 0.120 | 0.120 | 0.120 | 0.120 | 0.120 |
| | Gezondheid | 0.306** | 0.301** | 0.305** | 0.300** | 0.306** | 0.301** |
| | stdv | 0.143 | 0.142 | 0.143 | 0.142 | 0.143 | 0.142 |
| Educatie | | -0.817** | | -0.517** | | 0.317*** | |
| | stdv | 0.404 | | 0.389 | | 0.089 | |
| | Hoeveelheid ziek | 0.099* | | 0.098* | | 0.098* | |
| | stdv | 0.055 | | 0.055 | | 0.055 | |
| | Geluk | 0.093 | | 0.092 | | 0.094 | |
| | stdv | 0.082 | | 0.082 | | 0.082 | |
| | Gezondheid | -0.019 | | -0.020 | | -0.017 | |
| | stdv | 0.098 | | 0.097 | | 0.098 | |
| Rondkomen | | 0.648 | | 0.246 | | -0.042 | |
| | stdv | 0.429 | | 0.411 | | 0.100 | |

| | | | | | | |
|----------|------------------|---------|----------|-----------|-----|-----|
| | Hoeveelheid ziek | 0.034 | 0.035 | 0.034 | | |
| | stdv | 0.055 | 0.055 | 0.055 | | |
| | Geluk | -0.013 | -0.014 | -0.015 | | |
| | stdv | 0.082 | 0.082 | 0.082 | | |
| | Gezondheid | -0.024 | -0.024 | -0.022 | | |
| | stdv | 0.097 | 0.097 | 0.098 | | |
| Besteden | | 0.565** | 0.974*** | -0.231*** | | |
| stdv | | 0.249 | 0.225 | 0.054 | | |
| | Hoeveelheid ziek | 0.048 | 0.051 | 0.041 | | |
| | stdv | 0.091 | 0.091 | 0.091 | | |
| | Geluk | 0.091 | 0.100 | 0.098 | | |
| | stdv | 0.136 | 0.136 | 0.136 | | |
| | Gezondheid | 0.217 | 0.220 | 0.205 | | |
| | stdv | 0.166 | 0.166 | 0.167 | | |
| N | | 128 | 128 | 128 | 128 | 128 |

*= p<0.1, **= p<0.05, ***= p<0,01

Tabel B5.4: Bevindingen SUR test van afzonderlijke taalvariabelen X_n op mechanismes.

| | Laaggeletterd | Anderstalig | Tevredenheid taalvermogen |
|----------------|---------------|-------------|---------------------------|
| Stress | 0.801* | 0.411 | -0.053 |
| stdv | 0.440 | 0.422 | 0.102 |
| Dokter | -0.863*** | -0.948*** | 0.229*** |
| stdv | 0.121 | 0.107 | 0.025 |
| Marginalisatie | 0.128 | 0.193** | -0.012 |
| stdv | 0.099 | 0.094 | 0.023 |
| Sport | -0.146 | 0.073 | 0.064 |
| stdv | 0.285 | 0.271 | 0.065 |
| Educatie | -0.787* | -0.693* | 0.355*** |
| stdv | 0.412 | 0.395 | 0.090 |
| Rondkomen | 0.506 | 0.125 | 0.006 |
| stdv | 0.458 | 0.439 | 0.104 |

| | | | |
|----------|---------|----------|-----------|
| Besteden | 0.508** | 0.891*** | -0.222*** |
| stdv | 0.252 | 0.230 | 0.055 |

*= p<0.1, **= p<0.05, ***= p<0,01

Tabel B5.5: Bevindingen SUR test van mechanismes op gezondheidsvariabelen Y_j .

| | Hoeveelheid ziek | Geluk | Gezondheid |
|----------------|---------------------|----------|------------|
| Stress | -0.135** | 0.453*** | 0.593*** |
| stdv | 0.059 | 0.090 | 0.106 |
| Dokter | -0.403** | -0.366 | 0.290 |
| stdv | 0.183 | 0.276 | 0.325 |
| Marginalisatie | 0.128 | -0.276 | -0.304 |
| stdv | 0.243 | 0.368 | 0.432 |
| Sport | -0.038 | -0.161 | 0.210 |
| stdv | 0.086 | 0.130 | 0.153 |
| Educatie | 0.065 | 0.094 | 0.029 |
| stdv | 0.061 | 0.093 | 0.109 |
| Rondkomen | 0.037 | -0.026 | -0.037 |
| stdv | 0.058 | 0.088 | 0.103 |
| Besteden | -0.021 | 0.030 | 0.208 |
| stdv | 0.100 | 0.152 | 0.179 |

*= p<0.1, **= p<0.05, ***= p<0,01

Tabel B5.6: Bevindingen SEM basismodel zonder method(mlmv).

| Effect Xn op mechanisme | Effect mechanisme op Y_j | Laaggeletterd | | Anderstalig | | Tevredenheid taalvermogen | |
|-------------------------|----------------------------|---------------|------------|-------------|------------|---------------------------|------------|
| | | Met SES | Zonder SES | Met SES | Zonder SES | Met SES | Zonder SES |
| Stress | | 0.801* | 0.729* | 0.411 | 0.397 | -0.053 | -0.055 |
| stdv | | 0.440 | 0.426 | 0.422 | 0.410 | 0.102 | 0.097 |
| | Hoeveelheid ziek | -0.139** | -0.128** | -0.135** | -0.127** | -0.139** | -0.127** |
| | stdv | 0.059 | 0.052 | 0.059 | 0.052 | 0.059 | 0.053 |
| | Geluk | 0.449*** | 0.455*** | 0.453*** | 0.456*** | 0.455*** | 0.475*** |
| | stdv | 0.090 | 0.078 | 0.090 | 0.078 | 0.089 | 0.078 |
| | Gezondheid | 0.593*** | 0.548*** | 0.593*** | 0.548*** | 0.588*** | 0.536*** |

| | | | | | | | |
|----------------|------------------|-----------|-----------|-----------|-----------|-----------|----------|
| | stdv | 0.106 | 0.096 | 0.106 | 0.095 | 0.107 | 0.097 |
| Dokter | | -0.863*** | -0.896*** | -0.948*** | -0.980*** | 0.229*** | 0.227*** |
| | stdv | 0.121 | 0.120 | 0.107 | 0.106 | 0.025 | 0.025 |
| | Hoeveelheid ziek | -0.410** | -0.357** | -0.403** | -0.336** | -0.406** | -0.356** |
| | stdv | 0.181 | 0.153 | 0.183 | 0.155 | 0.184 | 0.156 |
| | Geluk | -0.373 | -0.372 | -0.366 | -0.350 | -0.308 | -0.319 |
| | stdv | 0.276 | 0.231 | 0.276 | 0.231 | 0.275 | 0.229 |
| | Gezondheid | 0.292 | 0.235 | 0.290 | 0.232 | 0.242 | 0.201 |
| | stdv | 0.326 | 0.283 | 0.325 | 0.282 | 0.330 | 0.287 |
| Marginalisatie | | 0.128 | 0.128 | 0.132 | 0.193** | 0.191** | -0.012 |
| | stdv | 0.099 | 0.096 | 0.094 | 0.091 | 0.023 | 0.022 |
| | Hoeveelheid ziek | 0.176 | 0.252 | 0.128 | 0.212 | 0.160 | 0.237 |
| | stdv | 0.242 | 0.231 | 0.243 | 0.232 | 0.247 | 0.235 |
| | Geluk | -0.227 | -0.161 | -0.276 | -0.203 | -0.298 | -0.248 |
| | stdv | 0.370 | 0.347 | 0.368 | 0.347 | 0.369 | 0.346 |
| | Gezondheid | -0.315 | -0.295 | -0.304 | -0.289 | -0.294 | -0.271 |
| | stdv | 0.437 | 0.426 | 0.432 | 0.422 | 0.443 | 0.433 |
| Sport | | -0.146 | -0.130 | 0.073 | 0.083 | 0.064 | 0.053 |
| | stdv | 0.285 | 0.273 | 0.271 | 0.261 | 0.065 | 0.061 |
| | Hoeveelheid ziek | -0.032 | -0.010 | -0.038 | -0.015 | -0.026 | -0.005 |
| | stdv | 0.085 | 0.081 | 0.086 | 0.082 | 0.087 | 0.083 |
| | Geluk | -0.155 | -0.129 | -0.161 | -0.134 | -0.127 | -0.103 |
| | stdv | 0.130 | 0.123 | 0.130 | 0.123 | 0.130 | 0.122 |
| | Gezondheid | 0.209 | 0.277* | 0.210 | 0.278* | 0.200 | 0.272* |
| | stdv | 0.154 | 0.150 | 0.153 | 0.150 | 0.156 | 0.153 |
| Educatie | | -0.787* | | -0.693* | | 0.355*** | |
| | stdv | 0.412 | | 0.395 | | 0.090 | |
| | Hoeveelheid ziek | 0.054 | | 0.065 | | 0.052 | |
| | stdv | 0.061 | | 0.061 | | 0.062 | |
| | Geluk | 0.082 | | 0.094 | | 0.057 | |
| | stdv | 0.093 | | 0.093 | | 0.093 | |
| | Gezondheid | 0.031 | | 0.029 | | 0.048 | |
| | stdv | 0.110 | | 0.109 | | 0.111 | |
| Rondkomen | | 0.506 | | 0.125 | | 0.006 | |
| | stdv | 0.458 | | 0.439 | | 0.104 | |
| | Hoeveelheid ziek | 0.044 | | 0.037 | | 0.046 | |
| | stdv | 0.058 | | 0.058 | | 0.059 | |
| | Geluk | -0.019 | | -0.026 | | 0.014 | |
| | stdv | 0.088 | | 0.088 | | 0.089 | |
| | Gezondheid | -0.039 | | -0.037 | | -0.063 | |
| | stdv | 0.104 | | 0.103 | | 0.106 | |
| Besteden | | 0.508** | | 0.891*** | | -0.222*** | |

| | | | | | | | |
|------|------------------|--------|-----|--------|-----|--------|-----|
| stdv | | 0.252 | | 0.230 | | 0.055 | |
| | Hoeveelheid ziek | -0.018 | | -0.021 | | -0.019 | |
| | stdv | 0.100 | | 0.100 | | 0.101 | |
| | Geluk | 0.033 | | 0.030 | | 0.010 | |
| | stdv | 0.152 | | 0.152 | | 0.151 | |
| | Gezondheid | 0.207 | | 0.208 | | 0.225 | |
| | stdv | 0.180 | | 0.179 | | 0.181 | |
| N | | 111 | 118 | 112 | 119 | 109 | 116 |

*= p<0.1, **= p<0.05, ***= p<0,01

Tabel B5.7: Samenvatting van waarden van waargenomen significantieveranderingen voor participanten die missende waarden hebben voor de SES variabelen.

| | N | Gemiddeld | Stdv | Minimum | Maximum |
|---------------------|---|-----------|----------|---------|---------|
| Anderstalig | 8 | 0.625 | 0.517549 | 0 | 1 |
| Laaggeletterd | 8 | 0.625 | 0.517549 | 0 | 1 |
| Educatie | 6 | 3 | 2.75681 | 1 | 7 |
| Hoeveelheid ziek | 8 | 2.875 | 1.356203 | 1 | 5 |
| Sport | 8 | 3.125 | 1.125992 | 1 | 5 |

Tabel B5.8: Bevindingen SEM analyse waarbij taalvariabelen X_n een direct effect hebben op gezondheidsvariabelen Y_j .

| Effect X_n op mechanisme | Effect mechanisme op Y_j | Laaggeletterd | | Anderstalig | | Tevredenheid taalvermogen | |
|----------------------------|----------------------------|---------------|------------|-------------|------------|---------------------------|------------|
| | | Met SES | Zonder SES | Met SES | Zonder SES | Met SES | Zonder SES |
| Stress | | 0.707* | 0.732* | 0.501 | 0.505 | -0.052 | -0.052 |
| stdv | | 0.414 | 0.413 | 0.395 | 0.395 | 0.094 | 0.094 |
| | Hoeveelheid ziek | -0.138** | -0.122** | -0.139*** | -0.118** | -0.134** | -0.115** |

| | | | | | | | |
|----------------|------------------|-----------|-----------|-----------|-----------|----------|----------|
| | stdv | 0.054 | 0.051 | 0.054 | 0.050 | 0.055 | 0.051 |
| | Geluk | 0.459*** | 0.466*** | 0.472*** | 0.472*** | 0.475*** | 0.476*** |
| | stdv | 0.082 | 0.076 | 0.082 | 0.076 | 0.082 | 0.076 |
| | Gezondheid | 0.551*** | 0.539*** | 0.554*** | 0.547*** | 0.553*** | 0.543*** |
| | stdv | 0.098 | 0.090 | 0.098 | 0.090 | 0.097 | 0.090 |
| Dokter | | -0.819*** | -0.819*** | -0.933*** | -0.933*** | 0.215*** | 0.215*** |
| | stdv | 0.120 | 0.120 | 0.105 | 0.105 | 0.025 | 0.025 |
| | Hoeveelheid ziek | -0.266 | -0.187 | -0.180 | -0.083 | -0.379* | -0.319* |
| | stdv | 0.186 | 0.150 | 0.199 | 0.190 | 0.196 | 0.193 |
| | Geluk | -0.173 | -0.118 | -0.196 | -0.109 | -0.219 | -0.154 |
| | stdv | 0.282 | 0.263 | 0.304 | 0.289 | 0.296 | 0.290 |
| | Gezondheid | 0.610* | 0.491 | 0.682* | 0.617** | 0.247 | 0.162 |
| | stdv | 0.336 | 0.314 | 0.361 | 0.343 | 0.350 | 0.344 |
| Marginalisatie | | 0.132 | 0.132 | 0.134 | 0.199** | 0.199** | -0.018 |
| | stdv | 0.096 | 0.096 | 0.090 | 0.090 | 0.022 | 0.022 |
| | Hoeveelheid ziek | 0.124 | 0.166 | 0.080 | 0.118 | 0.169 | 0.218 |
| | stdv | 0.230 | 0.234 | 0.233 | 0.235 | 0.233 | 0.237 |
| | Geluk | 0.321 | -0.265 | -0.302 | -0.276 | -0.296 | -0.245 |
| | stdv | 0.143 | 0.351 | 0.359 | 0.359 | 0.352 | 0.351 |
| | Gezondheid | -0.367 | -0.392 | -0.414 | -0.466 | -0.247 | -0.301 |
| | stdv | 0.410 | 0.409 | 0.418 | 0.415 | 0.411 | 0.409 |
| Sport | | -0.128 | -0.148 | 0.036 | 0.036 | 0.061 | 0.063 |
| | stdv | 0.265 | 0.265 | 0.250 | 0.250 | 0.059 | 0.059 |
| | Hoeveelheid ziek | -0.035 | -0.015 | -0.038 | -0.019 | -0.045 | -0.027 |
| | stdv | 0.079 | 0.080 | 0.078 | 0.079 | 0.081 | 0.081 |
| | Geluk | -0.122 | -0.110 | -0.129 | -0.117 | -0.124 | -0.110 |
| | stdv | 0.120 | 0.120 | 0.120 | 0.119 | 0.121 | 0.122 |
| | Gezondheid | 0.321** | 0.316** | 0.314** | 0.310** | 0.279* | 0.282* |
| | stdv | 0.143 | 0.143 | 0.142 | 0.141 | 0.144 | 0.145 |
| Direct Xn->Y | | | | | | | |
| | Hoeveelheid ziek | 0.394 | 0.417 | 0.545* | 0.586** | -0.009 | -0.003 |
| | stdv | 0.279 | 0.282 | 0.298 | 0.291 | 0.071 | 0.069 |
| | Geluk | 0.419 | 0.421 | 0.240 | 0.345 | -0.057 | -0.060 |
| | stdv | 0.419 | 0.421 | 0.454 | 0.441 | 0.108 | 0.105 |
| | Gezondheid | 0.498 | 0.503 | 0.550 | 0.679 | 0.140 | 0.090 |
| | stdv | 0.498 | 0.501 | 0.541 | 0.523 | 0.126 | 0.121 |
| Educatie | | -0.813** | | -0.519 | | 0.315*** | |
| | stdv | 0.404 | | 0.389 | | 0.089 | |
| | Hoeveelheid ziek | 0.101* | | 0.091* | | 0.099* | |
| | stdv | 0.054 | | 0.054 | | 0.055 | |
| | Geluk | 0.095 | | 0.089 | | 0.101 | |
| | stdv | 0.082 | | 0.082 | | 0.083 | |

| | | | | | |
|-----------|------------------|---------|----------|-----------|-----|
| | Gezondheid | -0.017 | -0.027 | -0.033 | |
| | stdv | 0.097 | 0.097 | 0.098 | |
| Rondkomen | | 0.646 | 0.246 | -0.038 | |
| | stdv | 0.430 | 0.411 | 0.100 | |
| | Hoeveelheid ziek | 0.030 | 0.042 | 0.034 | |
| | stdv | 0.054 | 0.054 | 0.055 | |
| | Geluk | -0.018 | -0.010 | -0.012 | |
| | stdv | 0.081 | 0.082 | 0.082 | |
| | Gezondheid | -0.030 | -0.017 | -0.030 | |
| | stdv | 0.097 | 0.097 | 0.097 | |
| Besteden | | 0.564** | 0.971*** | -0.231*** | |
| | stdv | 0.249 | 0.225 | 0.054 | |
| | Hoeveelheid ziek | 0.042 | 0.007 | 0.039 | |
| | stdv | 0.090 | 0.092 | 0.094 | |
| | Geluk | 0.085 | 0.081 | 0.079 | |
| | stdv | 0.135 | 0.140 | 0.141 | |
| | Gezondheid | 0.209 | 0.172 | 0.255 | |
| | stdv | 0.166 | 0.172 | 0.172 | |
| N | | 128 | 128 | 128 | 128 |

*= p<0.1, **= p<0.05, ***= p<0,01

Tabel B5.9: VIF test naar aanleiding van de waargenomen variabiliteit afkomstig van het directe effect van X_n op Y_j

| Gezondheidsvariabele Taalvariabele | Hoeveelheid ziek | | | Gezondheid | | |
|---------------------------------------|------------------|-------------|---------------------------|---------------|-------------|---------------------------|
| | Laaggeletterd | Anderstalig | Tevredenheid taalvermogen | Laaggeletterd | Anderstalig | Tevredenheid taalvermogen |
| VIF taalvariabele | 1.37 | 1.62 | 1.58 | 1.37 | 1.62 | 1.58 |
| VIF Dokter | 1.37 | 1.62 | 1.58 | 1.37 | 1.62 | 1.58 |
| Gemiddelde VIF | 1.37 | 1.62 | 1.58 | 1.37 | 1.62 | 1.58 |

Tabel B5.10: Bevindingen SEM analyse waarbij variabele Arabisch wordt geïntroduceerd als X_n én mechanisme.

| Effect X_n op mechanisme | Effect mechanisme op Y_j | Als mechanisme met X tevredenheid taalvermogen | | | |
|-------------------------------|-------------------------------|------------------------------------------------------|------------|---------------|---------------|
| | | Als X | | Zonder SES | |
| | | Met SES | Zonder SES | Met SES | Zonder SES |
| Stress | | 0.350 | 0.347 | -0.053 | -0.052 |
| stdv | | 0.421 | 0.421 | 0.094 | 0.094 |
| | Hoeveelheid ziek | -0.135** | -0.116** | -0.135** | -0.115** |

| | | | | | |
|----------------|------------------|----------|-----------|----------|----------|
| | stdv | 0.055 | 0.051 | 0.055 | 0.051 |
| | Geluk | 0.474*** | 0.473*** | 0.480*** | 0.471*** |
| | stdv | 0.082 | 0.076 | 0.081 | 0.075 |
| | Gezondheid | 0.557*** | 0.548*** | 0.558*** | 0.546*** |
| | stdv | 0.098 | 0.090 | 0.098 | 0.090 |
| | | - | | | |
| Dokter | | 0.617*** | -0.617*** | 0.215*** | 0.215*** |
| | stdv | 0.131 | 0.131 | 0.025 | 0.025 |
| | Hoeveelheid ziek | -0.395** | -0.323** | -0.416** | -0.317* |
| | stdv | 0.166 | 0.151 | 0.177 | 0.164 |
| | Geluk | -0.299 | -0.256 | -0.121 | -0.053 |
| | stdv | 0.250 | 0.226 | 0.263 | 0.240 |
| | Gezondheid | -0.299 | 0.327 | 0.541* | 0.441 |
| | stdv | 0.251 | 0.270 | 0.319 | 0.291 |
| Marginalisatie | | 0.108 | 0.108 | 0.109 | -0.017 |
| | stdv | 0.097 | 0.097 | 0.022 | 0.022 |
| | Hoeveelheid ziek | 0.167 | 0.218 | 0.181 | 0.216 |
| | stdv | 0.230 | 0.235 | 0.232 | 0.236 |
| | Geluk | -0.266 | -0.216 | -0.342 | -0.306 |
| | stdv | 0.350 | 0.351 | 0.344 | 0.343 |
| | Gezondheid | -0.304 | -0.329 | -0.350 | -0.384 |
| | stdv | 0.408 | 0.406 | 0.409 | 0.406 |
| Sport | | -0.076 | -0.076 | 0.062 | 0.063 |
| | stdv | 0.266 | 0.266 | 0.059 | 0.059 |
| | Hoeveelheid ziek | -0.047 | -0.028 | -0.049 | -0.027 |
| | stdv | 0.079 | 0.080 | 0.080 | 0.080 |
| | Geluk | -0.134 | -0.211 | -0.113 | -0.106 |
| | stdv | 0.120 | 0.120 | 0.118 | 0.118 |
| | Gezondheid | 0.307** | 0.301** | 0.317** | 0.311** |
| | stdv | 0.143 | 0.142 | 0.143 | 0.142 |
| Educatie | | 0.036 | | 0.313*** | |
| | stdv | 0.418 | | 0.089 | |
| | Hoeveelheid ziek | 0.101* | | 0.101* | |
| | stdv | 0.055 | | 0.055 | |

| | | | | |
|-----------|------------------|----------|-----------|-----------|
| | Geluk | 0.093 | 0.070 | |
| | stdv | 0.082 | 0.081 | |
| | Gezondheid | -0.021 | -0.029 | |
| | stdv | 0.097 | 0.098 | |
| Rondkomen | | 0.824* | -0.036 | |
| stdv | | 0.432 | 0.100 | |
| | Hoeveelheid ziek | 0.033 | 0.036 | |
| | stdv | 0.055 | 0.055 | |
| | Geluk | 0.013 | -0.031 | |
| | stdv | 0.082 | 0.081 | |
| | Gezondheid | -0.024 | -0.030 | |
| | stdv | 0.098 | 0.098 | |
| Besteden | | 0.940*** | -0.231*** | |
| stdv | | 0.243 | 0.055 | |
| | Hoeveelheid ziek | 0.042 | 0.049 | |
| | stdv | 0.091 | 0.093 | |
| | Geluk | 0.095 | 0.035 | |
| | stdv | 0.136 | 0.138 | |
| | Gezondheid | 0.211 | 0.035 | |
| | stdv | 0.167 | 0.138 | |
| Arabisch | | | -0.094*** | -0.095*** |
| stdv | | | 0.018 | 0.018 |
| | Hoeveelheid ziek | | -0.105 | 0.025 |
| | stdv | | 0.270 | 0.261 |
| | Geluk | | 0.793** | 0.846** |
| | stdv | | 0.400 | 0.383 |
| | Gezondheid | | 0.405 | 0.490 |
| | stdv | | 0.485 | 0.463 |
| N | | 129 | 129 | 129 |

*= p<0.1, **= p<0.05, ***= p<0,01