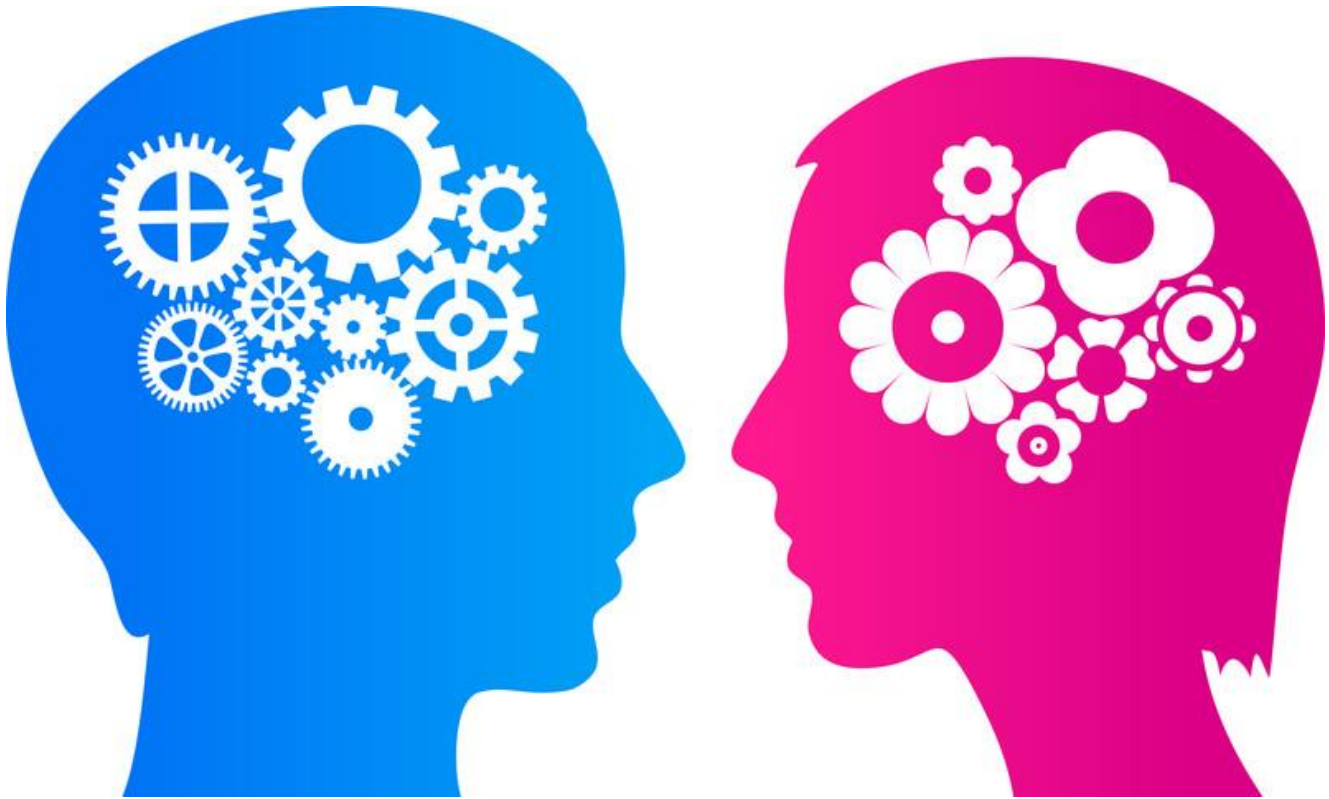


# Moraliteit, taal en geslacht

De verschillen in moreel redeneren door mannelijke en vrouwelijke adolescenten in het Nederlands en in het Engels



Bachelorscriptie moderne taalkunde

Door: Nina Brouwers

Begeleidster: Dr. Susanne Brouwer

Tweede beoordelaar: Dr. Stefan Frank

20 januari 2017

## **Inhoudsopgave**

Abstract	3
Inleiding	3
Dual process theory	3
Foreign language effect	4
Geslacht	6
Het huidige onderzoek	7
Methode	8
Participanten	8
Materiaal	9
Procedure	10
Design	11
Data analyse	11
Resultaten	12
Conclusie en discussie	14
Moedertaal versus vreemde taal	14
Mannen versus vrouwen	16
Mogelijkheden voor vervolgonderzoek	17
Tot slot	18
Referenties	18
Bijlagen	21
Bijlage A: Nederlandse versie materiaal	21
Bijlage B: Engelse versie materiaal	24
Bijlage C: Controlevragen Engelse versie	27
Bijlage D: Aanvullende vragen voor participanten	28
Bijlage E: Uitkomsten normaliteit per taal, geslacht en versie	29
Bijlage F: Uitkomsten normaliteit na log per taal, geslacht en versie	30

## **Abstract**

What would you do? Push one man off a bridge in order to save five others? Or would you do nothing? There are many aspects that could influence your decision, like language in which the dilemma is presented, gender and emotionality of the dilemma. In our current research we examined the influence of gender and language on moral decision making in adolescents. By moral decision making we mean the extent to which you choose an utilitarian or a deterministically answer. Do you choose for the most survivors or do you refuse to deliberately kill another person? From our participants, 127 fourth year high school students, the first half got three Dutch dilemmas and the other half got three English dilemmas. It turns out that the language in which the dilemmas are presented, doesn't make any difference in the answers. Gender, on the other hand, did show an effect. In one of the three dilemmas we used, boys scored significantly more utilitarian than girls. This shows that not only gender, but also the kind of dilemma has effect on the results.

## **Inleiding**

Iedereen heeft ideeën over wat goed is en wat slecht. De meeste mensen zijn het er bijvoorbeeld over eens dat men niet zomaar een ander mag vermoorden. Het kan dat deze mening verandert als er andere aspecten bij komen kijken. Als iemand bijvoorbeeld vijf mensen kan redden door één iemand te doden, dan wordt het al een stuk lastiger om een keuze te maken. Kinderen vinden het nog heel lastig om te bepalen wat wel en niet mag. Hun moraliteit is nog volop in de groei. De ontwikkeling van moraliteit verloopt niet bij alle mensen hetzelfde. Kohlberg (1969) heeft veel invloed gehad op de vraag hoe moraliteit zich ontwikkelt. Hij ontwierp een model bestaande uit zes stappen. Morele groei kwam volgens Kohlberg niet zozeer door het ontwikkelen van de hersenen, maar meer door ervaringen in het aannemen van verschillende rollen en het bekijken van een probleem vanuit verschillende perspectieven. Zijn werk heeft veel invloed gehad op de ontwikkeling van theorieën over moraliteit. Toch kreeg Kohlberg later ook veel kritiek. Bijvoorbeeld Gilligan (1982, 1982, 1987, 1988), Gilligan en Attanucci (1988) en Gilligan, Ward en Taylor (1988) hadden kritiek op Kohlberg's ideeën. Zij waren van mening dat het belang van morele emoties werd genegeerd en dat de verschillen tussen mannen en vrouwen niet goed uiteen werden gezet. Het doel van dit onderzoek is om meer inzicht te krijgen in de invloed van taal en geslacht op morele keuzes.

### *Dual process theory*

Volgens de *dual process theory* speelt naast cognitie ook emotie een rol bij moreel redeneren. Deze twee processen zijn aan de ene kant het komen tot een keuze via de cognitieve weg en aan de andere kant een beslissing maken via de emotionele weg. Volgens Greene (2009, p. 581) worden utilitaristische keuzes vaak gedreven door gecontroleerde cognitieve processen en worden deontologische keuzes vaak gedreven door automatische emotionele reacties.

Volgens de utilitaristische visie moet moraliteit altijd uitgaan van ‘het grote goed’. Hoe meer mensen overleven, hoe beter. Het gaat om een zo goed mogelijke uitkomst, niet om de weg daar naartoe. Een utilitaristische keuze wordt dan ook vaak met de ratio gemaakt. De deontologische visie stelt daarentegen dat er niet over bepaalde grenzen heen gegaan mag worden, ook al zou dit wel ten goede komen aan ‘het grote goed’.

Vooraf bij morele dilemma’s moet er een keuze gemaakt worden tussen de utilitaristische en de deontologische visie. Bij een moreel dilemma moet een keuze gemaakt worden uit twee opties die beide (on)aantrekkelijk zijn. Volgens Young en Koenigs (2007, p. 76-77) is er sprake van meer afkeer van de utilitaristische keuze bij een dilemma dat sterkere emoties met zich meebrengt. Dit betekent dat men bij minder emotie eerder voor een utilitaristische keuze gaat. De mate van emotie kan worden bepaald door het onderwerp, maar ook door andere zaken. Men zal bijvoorbeeld ook minder emotie voelen als men minder goed begrijpt waar het over gaat. Het is dus mogelijk dat een dilemma in een andere taal dan de moedertaal een minder emotionele reactie met zich meebrengt.

### *Foreign language effect*

Uit onderzoek van onder andere Costa, Foucart, Hayakawa, Apesteguia, Heafner en Keysar (2014), Costa, Foucart, Arnon, Aparici en Apesteguia (2014), Alter, Oppenheimer, Epley en Eyre (2007) en Geipel, Hadjichristidis en Surian (2015) blijkt dat het presenteren van een moreel probleem in een tweede taal (L2) in plaats van in de eerste taal (L1) invloed heeft op de keuze die gemaakt wordt. Dit wordt ook wel het *foreign language effect* (FLE) genoemd. Het FLE treedt onder andere op bij het gebruik van morele dilemma’s. Er zijn verschillende versies van dit soort dilemma’s die variëren in de mate van emotionaliteit. Een van de bekendste morele dilemma’s is het trolleydilemma, opgezet door Thomson (1985). Dit dilemma heeft twee versies. In de zogenaamde brugversie (*‘footbridge’*) moeten de participanten kiezen of ze één persoon van de brug af zouden duwen om vijf anderen te redden, of dat ze niets zouden doen en de vijf personen dood zouden gaan. In de hendelversie (*‘switch’*) hebben de participanten de keuze om vijf mensen te laten sterven, of om een hendel over te halen waardoor de trein naar een ander spoor zou gaan waar maar één persoon lag. Dit is dus een minder confronterende, minder emotionele versie.

Costa, Foucart, Hayakawa, Apesteguia, Heafner en Keysar (2014, p. 2-8) hebben deze twee morele dilemma’s aangeboden aan volwassenen. Zij gebruikten in hun onderzoek meerdere moedertalen en vreemde talen, namelijk Engels/Spaans, Spaans/Engels, Koreaans/Engels, Engels/Frans en Spaans of Engels/Hebreeuws. De participanten moesten

aangeven wat zij zouden doen bij het brugdilemma en in een tweede experiment moesten ze ook aangeven wat ze zouden doen bij het hendeldilemma. De resultaten van de brugversie lieten zien dat de keuzes van de participanten meer utilitaristisch waren in een vreemde taal dan in hun moedertaal. Men koos met de utilitaristische keuze dus voor het duwen van één persoon, zodat vijf andere mensen gered zouden worden. Dit FLE kwam bij de versie met de hendel niet naar voren. De mogelijke reden die hiervoor wordt gegeven is dat deze keuze minder emotioneel is. Daardoor wordt in de moedertaal ook vaker voor de utilitaristische oplossing gekozen.

Costa, Foucart, Hayakawa et al. (2014, p. 2-8) gaven de participanten een keuze tussen twee opties: wel of niet de man duwen/de hendel overhalen. Geipel, Hadjichristidis en Surian (2015, p. 3-9) hebben drie experimenten gedaan en hebben bij hun tweede en derde experiment naast een keuze uit twee opties ook een zevenpuntsschaal (1 = forbidden, 4= permissible, 7= obligatory) gebruikt. Uit experiment 1, met Italiaans als moedertaal en Duits of Engels als vreemde taal, bleek dat het gebruik van een vreemde taal het aantal utilitaristische reacties vergrootte bij het brugdilemma, maar niet bij het trolleydilemma. Uit experiment 2, met Chinees als moedertaal en Engels als vreemde taal, bleek hetzelfde. Experiment 3, met Duits als moedertaal en Engels als tweede taal, werd uitgevoerd om te kijken of het van invloed is hoe persoonlijk het dilemma is. Experiment 3 laat geen effect van dilemma zien, maar geeft wel een significant verschil tussen mannen en vrouwen, waarbij mannen de utilitaristische keuze meer goedkeuren. Een zevenpuntsschaal geeft meer genuanceerde keuzes en dat heeft invloed op de resultaten. Om een genuanceerder beeld van de data te krijgen, wordt in het huidige onderzoek ook een zevenpuntsschaal gebruikt.

Costa, Foucart, Arnon, Aparici en Apestequia (2014, p. 250-252) zijn verder ingegaan op het FLE en wilden weten of dit effect optreedt bij verschillende contexten. Zij concluderen dat dit inderdaad het geval is. Zowel bij winst-verlies dilemma's, als bij psychologische dilemma's, dilemma's waaraan risico's en onzekere uitkomsten verbonden zijn en logische dilemma's trad het FLE op. Uit de resultaten is ook gebleken dat op de dilemma's die gepresenteerd werden in een vreemde taal minder intuïtief/vanuit de emotie werd gereageerd. Deze verminderde emotie wordt door zowel Costa, Foucart en Hayakawa et al. (2014, p. 252) als Keysar, Hayakawa en Gyu An (2012, p. 666) als mogelijke verklaring gegeven voor het effect van een vreemde taal. Het is volgens hen het meest waarschijnlijk dat het effect op het maken van beslissingen komt doordat er meer psychologische afstand is bij een vreemde taal.

Een alternatieve verklaring wordt gegeven door Alter, Oppenheimer, Epley en Eyre (2007, p. 575). Zij suggereren dat de hersenen bij het nadenken in een vreemde taal meer

aandacht geven aan het dilemma. De hersenen geven een moeilijker probleem meer aandacht en doordat het nadenken in een vreemde taal lastiger is dan in de moedertaal, denken de hersenen dat het lastigere probleem (in de vreemde taal) meer aandacht nodig heeft. De consequentie is volgens Alter en collega's dat het probleem in de vreemde taal meer vanuit de analytische kant dan vanuit de emotie wordt bekeken (p. 575).

### *Geslacht*

Naast de invloed van de taal (L1 of L2) waarin een moreel dilemma wordt aangeboden, kan ook het geslacht van de participant bepalend zijn voor de morele keuzes die gemaakt worden. Op dit gebied zijn uiteenlopende resultaten gevonden. Aan de ene kant zijn er onderzoekers die geen (of weinig) effect van geslacht hebben gevonden. Bijvoorbeeld Walker (1984, p. 677-691) reageerde met zijn meta-analyse op de kritiek die werd gegeven op Kohlbergs theorie en op de beweringen dat er sekseverschillen bestaan bij moreel redeneren. De onderzoeken die Walker meenam in zijn meta-analyse, werden in de eerste taal van de participanten uitgevoerd. Walker deelde alle onderzoeken in op leeftijd en kwam zo op drie groepen. De eerste groep bestond uit jongeren (5-17 jaar), de tweede groep besprak adolescenten (middelbare school en universiteit) en in de derde groep werd het onderzoek bij volwassenen besproken (21-65 jaar). De conclusie van Walkers meta-analyse is dat er meer gelijkenissen bestaan dan verschillen bij moreel redeneren door mannen en vrouwen.

Aan de andere kant zijn er onderzoeken die wel verschillen laten zien tussen moreel redeneren bij mannen en vrouwen. Stimpson, Neff en Jensen (1991, p. 412-413) lieten mannen en vrouwen adjectieven beoordelen op 'goedheid' en vrouwen scoorden significant hoger bij adjectieven gerelateerd aan 'zorg'. Bij andere adjectieven scoorden vrouwen niet significant hoger. Elm, Kennedy en Lawton (2001, p. 252-259) gebruikten een verkorte vorm van de Defining Issues Test (manier om niveau van moreel redeneren te meten) en vrouwen van verschillende opleidingsniveaus, schooltypes en academische majors, lieten een hoger niveau van moreel redeneren zien dan mannen. In het onderzoek van Fumagalli, Ferrucci, Mameli, Marceglia, Mrakic-Sposta, Zago, Lucchiari, Consonni, Nordio, Ravettoni, Cappa en Priori (2010, p. 222-223) gaven mannen ook significant meer utilitaristische antwoorden op persoonlijke morele dilemma's dan vrouwen. You, Yukiko en Bebeau (2011, p. 273) hebben een meta-analyse uitgevoerd over negentien eerdere onderzoeken naar sekse en moraliteit. Hun resultaten suggereren bovendien een hogere score voor vrouwen op morele gevoeligheid.

Ryan, David en Reynolds (2004, p. 253-254) denken dat het niet zozeer het geslacht is dat zorgt voor verschillen in moraliteit, maar eerder de context en de relatie van de 'ik' tot de

persoon die bij het moreel dilemma betrokken is. Jaffee en Hyde (2000, p. 719) hebben een meta-analyse uitgevoerd over 113 studies om kwantitatief het werk over morele oriëntatie en de invloed van geslacht te bekijken. Jaffee en Hyde (2000) onderzochten de theorie van Gilligan (die kritiek had op Kohlberg) dat vrouwen vooral georiënteerd zijn op de zorg en mannen vooral georiënteerd zijn op gerechtigheid. De conclusie was dat deze oriëntaties niet sterk gerelateerd zijn aan geslacht.

Naast onderzoek naar de daadwerkelijke keuzes van mannen en vrouwen met betrekking tot morele dilemma's, is er ook onderzoek gedaan naar de hersengebieden die geactiveerd worden bij morele dilemma's. Harenski, Antonenko, Shane en Kiehl (2008, p. 319) hebben de hersenactiviteit van mannen en vrouwen bekeken. De resultaten lieten zien dat mannen en vrouwen verschillende neurale systemen gebruiken, ook al reageren ze ongeveer hetzelfde op morele stimuli. Dus naast onderzoek naar verschillen in (gedrags)reactie geeft dit ook bevestiging van de verschillen tussen mannen en vrouwen bij moreel redeneren vanuit neurobiologisch opzicht.

### *Het huidige onderzoek*

De conclusie die getrokken kan worden uit het vorige onderzoek is dat er een drietal aspecten zijn die een rol spelen bij het maken van morele keuzes. Emotionaliteit van het dilemma, de taal waarin het dilemma wordt gepresenteerd en het geslacht van degene die de morele keuze maakt. Er is op dit moment nog geen onderzoek gedaan naar een combinatie van die factoren. Kohlberg heeft zijn theorie gebaseerd op onderzoek onder alleen mannelijke participanten. Alter et al. (2007) geven in hun artikel geen informatie over de verhouding mannen en vrouwen die deelnamen aan het experiment. Er is zowel onderzoek gedaan naar de invloed van taal (bijvoorbeeld door Costa, Foucart, Hayakawa et al. 2014) als naar de invloed van geslacht (bijvoorbeeld door Fumagalli et al. 2010), maar deze twee variabelen zijn nog nooit gezamenlijk bekeken. Daarom wordt in het huidige onderzoek gekeken naar zowel de rol van de taal waarin het dilemma wordt gepresenteerd als naar het geslacht van de participant.

Eerder onderzoek werd uitgevoerd onder volwassenen en omdat adolescenten een doelgroep vormen die nog niet goed is onderzocht, wordt er in ons onderzoek naar adolescenten gekeken. Stey et al. (2013, p. 209) waren de eersten die morele keuzes bij vroege en late adolescenten onderzochten. Zij vonden geen verschillen tussen deze twee groepen. Het is wel mogelijk dat de groep adolescenten als geheel afwijkt van volwassenen, omdat adolescenten nog in ontwikkeling zijn wat betreft moreel redeneren.

De eerste vraag die in het huidige onderzoek gesteld wordt is:

‘In hoeverre verschilt het oplossen van morele dilemma’s door adolescenten in hun moedertaal van het oplossen van deze dilemma’s in een vreemde taal?’

Uit vorig onderzoek met volwassenen is gebleken dat in een vreemde taal minder emotioneel en meer utilitaristisch wordt gereageerd (Costa, Foucart en Havakawa et al. 2014, p. 2-8, Costa, Foucart en Arnon et al. 2014, p. 250-252, Keysar, Hayakawa en Gyu An 2012, p. 666 en Alter et al. 2007, p. 575). Daarom is de volgende hypothese opgesteld:

‘Adolescenten geven in een vreemde taal significant vaker een utilitaristisch antwoord dan in hun moedertaal.’

De tweede vraag die in het huidige onderzoek wordt gesteld is:

‘Wat is de invloed van sekse op moreel redeneren door adolescenten in de eerste en de vreemde taal?’

Bij onderzoek naar geslacht en moreel redeneren bij volwassenen zijn uiteenlopende resultaten gevonden. Aan de ene kant zijn er onderzoeken die geen verschillen tussen mannen en vrouwen hebben laten zien (Walker 1984, Ryan, David en Reynolds 2004, Jaffee en Hyde 2000). Aan de andere kant zijn er onderzoeken waarbij er wel verschil is gevonden (Stimpson, Neff en Jensen 1991, Elm, Kennedy en Lawton 2001, Fumagalli et al. 2010 en You, Yukiko en Bebeau 2011). Bij onderzoek van Fumagalli et al. (2010, p. 222-223) is specifiek terug te vinden dat mannelijke participanten vaker een utilitaristische keuze maken. De tweede hypothese luidt als volgt:

‘Mannelijke adolescenten kiezen (moedertaal en vreemde taal samengenomen) significant vaker voor een utilitaristisch antwoord dan vrouwelijke adolescenten.’

## **Methode**

### *Participanten*

Het onderzoek werd afgenomen bij leerlingen van de Katholieke Scholengemeenschap Etten-Leur die de Nederlandse taal als moedertaal hebben. Deze leerlingen zaten in het vierde jaar



van hun havo-opleiding. ( $N = 127$ , 58 vrouwen,  $M_{leeftijd} = 16;1$  jaar,  $SD = 7.154$ ). De ouders hebben een toestemmingsformulier getekend en de leerlingen deden vrijwillig mee aan het onderzoek. In tabel 1 is een overzicht te zien van de participanten.

Tabel 1: Aantal en gemiddelde leeftijd mannen en vrouwen per aangeboden taal en de Engelse taalvaardigheid

	Nederlands		Engels		Engelse taalvaardigheid
	N	$M(SD)$	N	$M(SD)$	
Man	29	16.08 (6.724)	40	15.99 (7.221)	6.93 (1.20)
Vrouw	37	16.09 (6.013)	21	16.15 (9.443)	6.69 (1.37)

### *Materiaal*

De participanten ontvingen een pakket met een voorblad met instructie, waarna drie morele dilemma's (brug, boot, vliegtuig) volgden, die beoordeeld moesten worden. De instructie en de dilemma's werden in het Nederlands of in het Engels aangeboden (zie Bijlage A voor het Nederlands en Bijlage B voor het Engels). Er waren twaalf versies: zes Engelse en zes Nederlandse. Het enige verschil tussen de zes versies in dezelfde taal was de volgorde van het aanbod van de dilemma's. Lanteri, Cheline en Rizzello (2008, p. 797) gebruikten het trolleydilemma om te onderzoeken of de volgorde waarin de dilemma's werden aangeboden van invloed was op de antwoorden en dit bleek inderdaad het geval te zijn. Daarom is er in ons onderzoek voor gekozen om alle mogelijke volgordes van de dilemma's te gebruiken, zodat er geen sprake kan zijn van volgorde-effecten. Bij de Engelse versies kreeg men een extra blad met controlevragen om te controleren of de participanten de Engelse dilemma's hadden begrepen (zie Bijlage C).

Uit het onderzoek van Costa, Foucart, Hayakawa et al. (2014) is gebleken dat bij het trolleydilemma met de brug een vreemde taal zorgde voor meer utilitaristische keuzes. Dit gold niet voor de hendelversie. Daarom is er in het huidige onderzoek voor de brugversie gekozen. Daarnaast zijn er twee soortgelijke dilemma's gekozen. Het dilemma met de brug en het dilemma met de boot zijn ontleend aan het onderzoek van Stey, Lapsley en McKeever (2013). Het dilemma met het vliegtuig is uit de actualiteit gehaald om bij de belevingswereld van de participanten aan te sluiten. In hun discussie geven Stey et al. (p. 219) aan dat het mogelijk is dat de gekunsteldheid van de dilemma's die zij gebruikten invloed had op hun

resultaten. Daarom is er voor een dilemma uit de actualiteit gekozen. Dit dilemma is omgevormd tot eenzelfde opbouw als de andere twee dilemma's. De hoofdpersonen binnen de dilemma's waren alleen mannen (Frank, Joe, Tom), zodat geslacht geen invloed zou hebben.

Aan de hand van de dilemma's kon de keuze (utilitaristisch of deontologisch) die gemaakt werd, gemeten worden op een zevenpuntsschaal. Er is geen gebruik gemaakt van afbeeldingen, waardoor de invloed van de taal gemeten kon worden, zonder dat dit werd beïnvloed door een afbeelding. Er is niet gevarieerd in formulering van de vraagstelling. Daarnaast is er ook niet gevarieerd in de richting van de antwoordcategorie. De meest utilitaristische keuze was altijd 1 op de zevenpuntsschaal en de meest deontologische keuze was altijd 7. Doordat de opbouw van de vragen hetzelfde is gehouden, kan dit niet van invloed zijn geweest op de resultaten. Uit de resultaten van Greene et al. (2008, p. 1148) is gebleken dat de richting van de antwoordcategorie (1 = zeer slecht of 7 = zeer goed) geen effect had op de reactietijd. Als de richting zelfs op de reactietijd geen effect heeft, dan is het onwaarschijnlijk dat onze resultaten hierdoor worden beïnvloed.

Vervolgens moesten de participanten een korte vragenlijst invullen met zeven vragen over leeftijd, geslacht en talenkennis (Bijlage D). Van de gegevens die met de vragenlijst verzameld werden, was niet alleen geslacht belangrijk voor het huidige onderzoek. Als blijkt dat jongens of meisjes over het algemeen beter zijn in hun vreemde taal, dan kan het zijn dat niet geslacht, maar taalvaardigheid de oorzaak is van de resultaten.

### *Procedure*

De ouders van de participanten kregen een toestemmingsformulier en deze leverden de participanten voor de start van het onderzoek in. Het onderzoek werd uitgevoerd op een middelbare school in Etten-Leur. De participanten zaten per klas samen in een lokaal en kregen een zo beknopt mogelijke uitleg in de taal waarin ze getest werden. Hierbij werd niet aangegeven wat het daadwerkelijke doel van het onderzoek was. Het onderzoek werd in acht klassen uitgevoerd. Vijf klassen kregen de dilemma's en bijbehorende vragen in het Engels en vijf klassen kregen de dilemma's en bijbehorende vragen in het Nederlands aangeboden. Per klas werden er zes versies uitgedeeld, met als verschil de volgorde van de dilemma's. De versies werden zo uitgedeeld, dat twee leerlingen die naast elkaar zaten beiden een andere versie kregen. Zo kon er niet worden afgekeken. De leerlingen mochten niet praten en overleggen en konden door de uitleg op het voorblad zelf aan de slag. Van tevoren werd aangegeven dat ze tien minuten de tijd kregen. De participanten die langer over het onderzoek

deden werden wel meegenomen in het onderzoek, maar de langere duur werd wel bijgehouden door de onderzoekers.

De participanten lazen de uitleg en de drie dilemma's en moesten invullen wat ze ervan vonden als de hoofdpersoon een bepaalde keuze maakte. Daarna moesten ze beknopt noteren waarom ze dit vonden. De participanten noteerden hoe goed of slecht ze de utilitaristische keuze vonden aan de hand van een schaal van 1 (zeer slecht) tot en met 7 (zeer goed). Vervolgens moesten ze de achtergrondvragenlijst invullen. Nadat iedereen klaar was, werden de formulieren opgehaald. Het onderzoek werd uitgevoerd door drie onderzoekers en onder toezicht van een docent. De totale testsessie duurde ongeveer 10 minuten.

### *Design*

In dit onderzoek is de afhankelijke variabele de mate waarin de participanten de utilitaristische keuze goedkeurden of afkeurden. Geslacht (man vs. vrouw) en Taal (Nederlands vs. Engels) zijn de *between-subjects* factoren en Dilemma (brug vs. boot vs. vliegtuig) is de *within-subjects* factor. Er is gekozen om Taal *between* aan te bieden zodat iedere participant de dilemma's maar één keer, in één taal aangeboden krijgt. Een *within-subjects* design is niet geschikt voor het onderzoek omdat de kans groot is dat men de tweede keer dat de vragen worden gesteld (in de andere taal) nog weet wat de eerste keer was ingevuld. Hierdoor zouden de resultaten niet betrouwbaar zijn.

### *Data analyse*

Aan de hand van skewness, kurtosis en de Q-Q-plots is gecheckt of de groepen normaal verdeeld zijn. Omdat de groep mannen met de Nederlandse versie en de groep vrouwen met de Engelse versie minder dan 30 participanten bevatten, is hierbij de normaliteit getest met de Kolmogorov-Smirnov en de Shapiro-Wilk toets. Een overzicht hiervan is te vinden in Tabel 2, in Bijlage E. Omdat het grootste gedeelte van de data niet normaal verdeeld was, is er een log-transformatie uitgevoerd, op de scores van de dilemma's. Vervolgens is de normaliteit opnieuw gecheckt. De resultaten hiervan zijn te zien in Tabel 3 in Bijlage F. Uit deze resultaten bleek dat de data niet normaler verdeeld raakten na de transformatie. Daarom zijn de toetsen gedaan met de niet getransformeerde data.

Het geslacht van de participant en de taal waarin de dilemma's werden aangeboden zijn allebei nominale data. De reactie op de dilemma's is ordinaal, namelijk een zevenpuntsschaal. Alle variabelen zijn kwantitatief. Omdat er sprake is van ordinale data kan er officieel geen ANOVA worden uitgevoerd, want het meetniveau moet minimaal interval

zijn. Toch is er na het checken van de assumpties en het transformeren van de data gekozen voor een 3 (dilemma) x 2 (geslacht) x 2 (taal) –weg mixed ANOVA op de mate waarin de participanten de utilitaristische keuze goedkeurden of afkeurden. Daarna is er een onafhankelijke t-toets gedaan tussen mannen en vrouwen, op de verschillende dilemma's en het gemiddelde van de drie dilemma's.

Vervolgens zijn er Pearson Chi-kwadraat toetsen gedaan met de scores (utilitaristisch of deontologisch) in combinatie met Taal (Nederlands of Engels) en Geslacht (man of vrouw). De participanten hebben onder hun keuze op de zevenpuntsschaal een verantwoording van deze keuze genoteerd. Deze is ingedeeld in vijf categorieën, namelijk 1) utilitaristisch antwoord, 2) deontologisch antwoord, 3) mix van utilitaristisch en deontologisch antwoord, 4) lot/plicht/opgedragen, 5) overig/onvolledig of geen antwoord. De categorieën 1) utilitaristisch en 2) deontologisch zijn gebruikt in een Pearson Chi-kwadraat toets in combinatie met Taal (Nederlands of Engels) en Geslacht (man of vrouw). Om te controleren of geslacht echt de oorzaak is van de resultaten en niet de taalvaardigheid, is er ten slotte nog een Pearson Chi-kwadraat toets gedaan tussen Geslacht (man of vrouw) en Taalvaardigheid (laag of hoog). De taalvaardigheid is gemeten door *self-rated proficiency*. De participanten konden hun Engelse taalvaardigheid scoren op een tienpuntsschaal. Er is voor een tienpuntsschaal gekozen omdat de participanten allemaal scholieren waren en deze schaal het cijfermodel is in Nederland. Bovendien kon er bij de data-analyse hierdoor een duidelijke scheiding gemaakt worden in twee groepen (hoge en lage taalvaardigheid), aan de hand van de mediaan.

## Resultaten

Tabel 4 laat de gemiddelde utilitaire scores van mannen en vrouwen zien per dilemma. Tussen haakjes staat de standaarddeviatie. Hoe hoger de score (maximum is 7), hoe meer utilitaristisch het antwoord.

Tabel 4: Gemiddelde utilitaire scores van mannen en vrouwen per aangeboden taal en dilemma

	Nederlandse versie				Engelse versie			
	Brug <i>M(SD)</i>	Boot <i>M(SD)</i>	Vliegtuig <i>M(SD)</i>	Alle <i>M(SD)</i>	Brug <i>M(SD)</i>	Boot <i>M(SD)</i>	Vliegtuig <i>M(SD)</i>	Alle <i>M(SD)</i>
Man	3.52 (1.60)	5.03 (1.35)	5.10 (1.37)	4.55 (1.13)	3.57 (1.62)	4.80 (1.04)	5.35 (1.42)	4.58 (1.02)
Vrouw	2.70	4.78	4.81	4.10	3.05	4.71	4.90	4.22

Als de gemiddelde scores per persoon van zowel het brugdilemma, het bootdilemma en het vliegtuigdilemma worden genomen, dan is te zien dat de mannen bij zowel de Nederlandse als de Engelse versie hoger scoren dan de vrouwen. Hoger scoren wil in dit geval zeggen dat de utilitaristische keuze van de hoofdpersoon als minder ‘fout’ wordt ervaren. Daarnaast is uit de tabel ook af te lezen dat zowel mannen als vrouwen op de Engelse versie gemiddeld hoger scoren dan op de Nederlandse versie. Dus zij scoren op de Engelse versie gemiddeld meer utilitaristisch.

Uit de 3 (dilemma) x 2 (geslacht) x 2 (taal) –drieweg mixed ANOVA op de mate waarin de participanten de utilitaristische keuze goedkeurden of afkeurden, is gebleken dat er significant hoofdeffect is van Geslacht ( $F(1,123) = 5.27, p < .05$ ). Mannen scoorden significant meer utilitaristisch ( $M = 4.56, SE = .12$ ) dan vrouwen ( $M = 4.16, SE = .13$ ). Uit de toets blijkt wel dat er een hoofdeffect van Dilemma is ( $F(2,246) = 100.36, p < .05$ ). Dit wil zeggen dat er significante verschillen bestaan tussen de scores op de drie dilemma’s. Een posthoc toets laat zien dat er een significant verschil is in resultaten tussen het brugdilemma en het bootdilemma ( $p < .05$ ) en tussen het brugdilemma en het vliegtuigdilemma ( $p < .05$ ). Er is daarentegen geen significant verschil in resultaten tussen het bootdilemma en het vliegtuigdilemma ( $p > .05$ ). Op het brugdilemma scoorden mannen significant meer utilitaristisch, terwijl dit bij de andere twee dilemma’s niet het geval was. Er is echter geen significant hoofdeffect van Taal (en geen interactie-effect tussen Taal en Geslacht - alle  $ps > .05$ ).

Vervolgens is er een 3 (Dilemma) bij 2 (Geslacht) –tweeweg ANOVA uitgevoerd. Bij het bootdilemma en het vliegtuigdilemma is er geen significant effect van Geslacht op de resultaten ( $p > .05$ ), terwijl er bij het brugdilemma wel een significant effect van Geslacht werd gevonden ( $F(1, 123) = 548.60, p < .05$ ). Mannen scoorden significant meer utilitaristisch ( $M = 3.55, SE = .18$ ) dan vrouwen ( $M = 2.88, SE = .21$ ) op het brugdilemma. Deze resultaten werden bevestigd door een onafhankelijke t-toets tussen geslacht. Dat de resultaten bij een gemiddelde van de drie dilemma’s wel significant waren, komt doordat er bij de brugversie gemiddeld door mannen meer utilitaristisch werd gekozen ( $M = 3.55, SD = 1.60$ ) dan door vrouwen ( $M = 2.83, SD = 1.35$ ). Dit verschil, .723, BCa 95% CI [.198, 1.248], is significant,  $t(125) = 2.725, p < .05$ .

Om onze resultaten direct te kunnen vergelijken met die van Costa et al. (2014) hebben we onze zevenpuntsschaal opgedeeld in een utilitaristische keuze (score van 5-7)

versus een deontologische keuze (score van 1-3). De participanten die een score van 4 hebben gegeven, werden niet meegenomen (brugdilemma:  $N = 30$ , bootdilemma:  $N = 37$ , vliegtuigdilemma:  $N = 20$ ). Een eerste Pearson Chi-kwadraat toets tussen Dilemma en Taal laat zien dat er bij geen van de drie dilemma's significante effecten bestaan van Taal op de keuze voor wel of niet utilitaristisch ( $p > .05$ ). Een tweede Pearson Chi-kwadraat toets tussen Dilemma en Geslacht laat zien dat er bij geen van de drie dilemma's significante effecten bestaan van Geslacht op de keuze voor wel of niet utilitaristisch ( $p > .05$ ).

De participanten gaven uitleg bij de keuze die ze maakten op de zevenpuntsschaal. Deze uitleg is opgedeeld in de eerder genoemde vijf categorieën. Van deze vijf categorieën zijn de categorieën *utilitaristische verantwoording* versus *deontologische verantwoording* gebruikt om mee te toetsen. Er zijn twee Pearson Chi-kwadraat toetsen uitgevoerd per dilemma. Eerst tussen Verantwoording (utilitaristisch of deontologisch) en Taal (Nederlands of Engels) en vervolgens tussen Verantwoording (utilitaristisch of deontologisch) en Geslacht (man of vrouw). De toets tussen Verantwoording (utilitaristisch of deontologisch) en Taal (Nederlands of Engels) laat zien dat bij geen enkel dilemma significante effecten bestaan ( $p > .05$ ). De toets tussen Dilemma en Geslacht laat geen significante effecten zien bij het bootdilemma en het vliegtuigdilemma ( $p > .05$ ), maar wel bij het brugdilemma ( $\chi^2(1) = 4.896$ ,  $p < .05$ ). Tot slot bleek uit een Pearson Chi-kwadraat toets tussen Geslacht (man en vrouw) en Taalvaardigheid (hoog of laag) dat er geen significante effecten bestaan van Geslacht op Taalvaardigheid ( $p > .05$ ). Dus het is niet zo dat de meer utilitaristische score van mannen kwam door een verschil in taalvaardigheid tussen mannen en vrouwen.

## **Conclusie en discussie**

### *Moedertaal versus vreemde taal*

In dit onderzoek werden twee vragen gesteld. Ten eerste: 'In hoeverre verschilt het oplossen van morele dilemma's door adolescenten in hun moedertaal van het oplossen van deze dilemma's in een vreemde taal?' De hypothese die hierbij werd opgesteld was dat adolescenten in een vreemde taal significant vaker een utilitaristisch antwoord geven dan in hun moedertaal. Deze hypothese is niet uitgekomen. Ondanks dat zowel mannen als vrouwen gemiddeld meer utilitaristisch scoren in hun vreemde taal dan in hun moedertaal, waren deze resultaten niet significant. Opvallend is dat zowel bij de mannen als bij de vrouwen de scores op het brugdilemma en het vliegtuigdilemma in de vreemde taal meer utilitaristisch waren. Het bootdilemma is het enige dilemma waarbij zowel door mannen als door vrouwen in de vreemde taal minder utilitaristisch gescoord is. Dit suggereert dat het soort dilemma dat wordt

aangeboden zeer belangrijk is voor de uitkomsten. Mogelijk zouden de resultaten met andere dilemma's wel significant zijn.

Dat het aangeboden dilemma een grote rol speelt, is ook terug te zien in de overige resultaten. De scores op het brugdilemma verschillen significant van de scores op de andere twee dilemma's. Mogelijk komt dit doordat het brugdilemma het meest emotioneel is, net zoals het geval was bij Costa et al. (2014). Bij hun onderzoek was het dilemma met de brug emotioneler dan het dilemma met de hendel en het FLE dat men bij het brugdilemma zag, was niet te vinden bij de versie met de hendel. In ons huidig onderzoek is het brugdilemma ook het meest emotioneel. Bij het brugdilemma moet je bewust een onschuldige man doden om vijf anderen te redden. Bij het bootdilemma red je vijf mensen door niets te doen (waardoor er één verdrinkt). Je vermoordt die ene persoon dan niet actief. Bij het vliegtuigdilemma moet je bewust een vliegtuig met onschuldige passagiers uit de lucht schieten, maar hierdoor blijven veel meer mensen leven. Bij de uitleg rechtvaardigden veel participanten dat ze het vliegtuig neerschoten met het feit dat er veel meer mensen zouden blijven leven.

Een mogelijke verklaring voor het niet repliceren van het FLE is dat de participanten die de Engelse versie kregen, niet genoeg in de Engelse modus zaten. Bij vervolgonderzoek zouden de participanten voordat ze starten met de dilemma's een taak in het Engels uit kunnen voeren, waardoor ze meer in de modus van de tweede taal zitten. Omdat de participanten de dilemma's in het Engels moeten lezen en hun antwoorden moeten noteren, lijkt het de beste optie om te kiezen voor een Engelse leestaak waarbij vragen beantwoord moeten worden. Een korte tekst over een algemeen onderwerp, met begripsvragen kan ervoor zorgen dat de participanten meer in het Engels gaan nadenken.

Een andere mogelijke verklaring voor het ontbreken van het FLE is de leeftijd van de participanten. Uit onderzoek van Stey et al. (2013, p. 218) is gebleken dat adolescenten oordelen kunnen vellen over hypothetische dilemma's, ook al zijn ze cognitief gezien nog niet helemaal ontwikkeld. Ze hebben dus een betere morele competentie dan zou worden verwacht, bijvoorbeeld vanuit de Kohlbergtraditie. Daar staat tegenover dat verschillen tussen jongere en oudere adolescenten deels verklaard konden worden door verschillen in taalvaardigheid. Het is dus mogelijk dat er geen FLE is opgetreden omdat adolescenten in het algemeen nog minder ontwikkeld zijn en meer moeite hebben met het begrijpen van teksten.

Ook proficiency kan hebben gezorgd voor het ontbreken van het FLE. We hebben met behulp van controlevragen gecontroleerd of de participanten de dilemma's in het Engels hadden begrepen. Dit was het geval. Het kan zijn dat de participanten juist te goed waren in Engels. Misschien waren ze zo goed in Engels, dat er bijna geen verschil bestaat tussen de

vreemde taal (Engels) en de moedertaal (Nederlands). Adolescenten zijn opgegroeid met Engels op tv en op internet en misschien zou het FLE wel zichtbaar zijn als er een vreemde taal gekozen zou zijn die minder in het dagelijks leven van adolescenten verweven zou zijn, zoals Frans. Daarnaast zijn Nederlands en Engels beide Germaanse talen. Daardoor liggen ze dicht bij elkaar. Frans, als Romaanse taal, ligt verder af van de moedertaal Nederlands.

Het FLE kan ook weg zijn gebleven, doordat men in een vreemde taal door stress minder utilitaristisch reageert. Zoals Costa, Foucart, Hayakawa et al. (2014) schrijven, is het goed mogelijk dat men meer stress ervaart door het lezen van een vreemde taal en dat dit zorgt voor een minder gecontroleerd cognitief proces.

### *Mannen versus vrouwen*

De tweede vraag die gesteld werd binnen dit onderzoek is: ‘Wat is de invloed van sekse op moreel redeneren door adolescenten in de eerste en de vreemde taal?’ De hypothese die hierbij werd opgesteld was dat mannelijke tweetalige adolescenten significant vaker kiezen voor een utilitaristisch antwoord dan vrouwelijke tweetalige adolescenten. Deze hypothese is uitgekomen, want er werd een effect gevonden van Geslacht. Onze resultaten lieten zien dat mannen significant hoger (meer utilitaristisch) scoorden dan vrouwen. Toch betekent dit niet dat de uitkomsten voor de drie dilemma’s apart ook significant waren. Dit bleek echter alleen het geval voor het brugdilemma.

Er is getest of er daadwerkelijk een effect van Geslacht was en of het verschil tussen mannen en vrouwen niet werd veroorzaakt doordat een van de twee groepen in het algemeen een hogere taalvaardigheid had. Uit de testen die gedaan zijn, blijkt dat geslacht en taalvaardigheid niet samenhangen. De resultaten sluiten aan bij de onderzoeken die een verschil tussen mannen en vrouwen vonden. Bij onderzoek waarin geen verschillen naar voren kwamen, kunnen oorzaken aangewezen worden. Walker (1984) vond in zijn meta-analyse meer gelijkenissen dan verschillen. Dit betekent dat er dus wel verschillen gevonden werden. Een tweede punt is dat aan het onderzoek van Walker veel jongeren deel hebben genomen. Bij jongere participanten kunnen de dilemma’s niet te heftig/emotioneel worden gemaakt en dat zou kunnen verklaren waarom er geen verschil is gevonden. Uit ons onderzoek blijkt dat er verschillen zichtbaar zijn bij het heftige dilemma en dat er bij de minder heftige dilemma’s geen verschillen zichtbaar zijn. Bij Jaffee en Hyde (2000) zijn ook geen significante verschillen gevonden. Daar staat tegenover dat zij met hun meta-analyse geen uitspraken kunnen doen over de utilitaristische versus de deontologische keuze.



Vrouwen scoorden binnen onderzoek van Stimpson, Neff en Jensen (1991) significant hoger bij adjectieven gerelateerd aan 'zorg'. De dilemma's in ons onderzoek gingen niet over zorg voor anderen, maar kunnen daar wel aan gekoppeld worden. Als vrouwen het zorgen voor anderen als meer goed zien dan mannen, dan is het mogelijk dat iemand doden meer tegen de natuur van vrouwen in gaat dan van mannen. Mannen zijn van oorsprong ook jagers en verzamelaars. In ons onderzoek stonden de hoofdpersonen in de dilemma's los van de participanten. Er was geen persoonlijke relatie met de hoofdpersonen. Onderzoek van Fumagalli et al. (2010) en Ryan, David en Reynolds (2004) laat zien dat de relatie van de participant met de hoofdpersoon bepalend is voor de resultaten. Wij hebben bewust gekozen voor onpersoonlijke dilemma's om dit geen invloed te laten hebben.

Een kanttekening bij de resultaten is dat het effect van Geslacht alleen geldt voor het brugdilemma. Zowel wanneer er op de zevenpuntsschaal als op de ja/nee-schaal als op de Verantwoording wordt gekeken, worden deze resultaten bevestigd. De resultaten gelden dus niet voor alle morele dilemma's. Dit kan te maken hebben met de emotionaliteit van het dilemma. Zoals Young en Koenigs (2007) aangaven, is er sprake van meer afkeer van de utilitaristische keuze bij een dilemma dat sterkere emoties met zich meebrengt. Het kan dus zijn dat de minder emotionele dilemma's (boot en vliegtuig) hebben gezorgd voor minder afkeer van de utilitaristische keuze en dat hierdoor de verschillen minder goed naar voren kwamen.

Een tweede kanttekening die geplaatst kan worden, is dat de scores van de mannen op het brugdilemma wel significant meer utilitaristisch waren, maar nog steeds niet echt utilitaristisch genoemd kunnen worden, namelijk 3.52 op een zevenpuntsschaal. Dat is nog onder de helft van de schaal. Uit de resultaten kan dus niet geconcludeerd worden dat vrouwen meer deontologisch zijn en mannen meer utilitaristisch, want de mannen scoren nog steeds meer in de deontologische richting dan in de utilitaristische richting.

### *Mogelijkheden voor vervolgonderzoek*

De resultaten van dit onderzoek geven veel mogelijkheden voor vervolgonderzoek. Ten eerste kunnen de participanten meer in de Engelse taalmodus gebracht worden door ze van tevoren een tekst in het Engels te laten lezen en daar vragen over te laten beantwoorden. Ten tweede kunnen er naast het brugdilemma meer emotionelere dilemma's gebruikt worden, zodat er bij andere dilemma's ook effect gevonden wordt. Het zou goed zijn om meer dan drie dilemma's te testen, zodat er grotere uitspraken gedaan kunnen worden en zodat het onderzoek betrouwbaarder wordt.

Een derde mogelijkheid voor vervolgonderzoek is het gebruiken van een andere schaal. Misschien is men bij een vijfpuntsschaal minder voorzichtig met het maken van een keuze en wordt er eerder voor de uiteinden van de schaal (1 of 5) gekozen. Een laatste optie voor vervolgonderzoek is het variëren in de vraagstelling. Nu zijn alle vragen op dezelfde manier opgebouwd, namelijk wat de hoofdpersoon doet is het meest utilitaristisch en de vraag aan de participant is wat ze daarvan vinden. Misschien verschillen de resultaten als de vraag is wat de participant vindt van de meest deontologische keuze.

Alter et al. (2007) gaven als verklaring van het FLE dat de hersenen bij een moeilijker probleem (dus ook een probleem in een vreemde taal) meer aandacht geven en daardoor cognitiever het probleem benaderen. Dat lijkt naar aanleiding van de huidige resultaten niet de juiste verklaring. Na het analyseren van de verantwoording die de participanten gaven bij hun keuze, werd zichtbaar dat bij de keuzes die omschreven werden als ‘moeilijk’ vaak de neutrale keuze werd omcirkeld. Het lijkt erop dat men bij een moeilijk dilemma niet meer utilitaristisch kiest, maar meer neutraal. Er kan dus ook een verklaring voor het FLE zijn.

#### *Tot slot*

We kunnen concluderen dat er nog veel onderzoek gedaan kan worden naar dit onderwerp, maar dat we met ons werk meer inzicht hebben verkregen in de invloed van taal en geslacht op morele dilemma's bij adolescenten. De taal waarin een dilemma gepresenteerd wordt had in het huidige onderzoek geen effect, maar zou na aanpassingen (taak vooraf, andere vreemde taal, vijfpuntsschaal) wel mogelijk effect kunnen hebben. Het geslacht had wel effect op de antwoorden die gegeven werden, maar dit was bij één van de drie dilemma's. Een groter effect zou zichtbaar kunnen worden door aanpassing van het booddilemma en het vliegtuigdilemma. De conclusie die we in ieder geval kunnen trekken, is dat adolescenten een beredeneerde keuze kunnen maken als het gaat om emotionele morele dilemma's en dat mannen en vrouwen verschillen in de antwoorden die zij geven.

#### **Referenties**

- Alter, A.L., Oppenheimer, D.M., Epley, N. en Eyre, R.N. (2007). Overcoming Intuition: Metacognitive Difficulty Activates Analytic Reasoning. *Journal of Experimental Psychology: General*, 136(4), 569-576.
- Berkum, J.J.A. van, Holleman, B., Nieuwland, M. Otten, M. en Murre, J. (2009). Right or Wrong? The Brain's Fast Response to Morally Objectionable Statements. *Psychological Science*, 20(9), 1092-1099.

- Costa, A., Foucart, A., Arnon, I., Aparici, M. en Apesteguia, J. (2014). “Piensa” twice: On the foreign language effect in decision making. *Cognition*, 130, 236-254.
- Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J. en Keysar, B. (2014). Your morals depend on language. *PLoS ONE*, 9(4), e94842.
- Elm, D.R., Kennedy, E.J. en Lawton, L. (2001). Determinants of Moral Reasoning: Sex Role Orientation, Gender, and Academic Factors. *Business Society*, 40(3), 241-265.
- Fumagalli, M., Ferrucci, R., Mameli, F., Marceglia, S., Mrakic-Sposta, S., Zago, S., Lucchiari, C., Consonni, D., Nordio, F., Ravettoni, G., Cappa, S. en Priori, A. (2010). Gender-related differences in moral judgments. *Cognitive Process* 11, 219-226.
- Green, J.D. (2009). Dual-process morality and the personal/ impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Social Psychology*, 45, 581-584.
- Greene, J. en Haidt, J. (2002). How (and where) does moral judgement work? *TRENDS in Cognitive Sciences*, 6(12), 517-523.
- Harenski, C.L., Antonenko, O., Shane, M.S. en Kiehl, K.A. (2008). Gender differences in neural mechanisms underlying moral sensitivity. *Social Cognitive and Affective Neuroscience* 3(4), 313-321.
- Jaffee, S. en Hyde, J.S. (2000). Gender Differences in Moral Orientation: A Meta-Analysis. *Psychological Bulletin*, 126(5), 703-726.
- Keysar, B., Hayakawa, S.L. en GyuAn, S. (2012). The Foreign-Language Effect: Thinking in a Foreign Tongue Reduces Decision Biases. *Psychological Science*, 23(6), 661-668.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Chicago, IL: Rand McNally.
- Lanteri, A., Chelini, C., Rizzello, S. (2008). An Experimental Investigation of Emotions and Reasoning in the Trolley Problem. *Journal of Business Ethics*, 83, 789-804.
- Ryan, M.K., David, B. en Reynolds, K.J. (2004). Who cares? The effect of gender and context on the self and moral reasoning. *Psychology of Women Quarterly*, 28, 246-255.

- Stey, C., Lapsley, D. en McKeever, M., O. (2013). Moral judgement in adolescents: Age differences in applying and justifying three principles of harm. *European Journal of Developmental Psychology*, 10(2), 206-220.
- Stimpson, D., Neff, W. en Jensen, L.C. (1991). The caring morality and gender differences. *Psychological Reports*, 69, 407-414.
- Thomson, J. (1985). The trolley problem. *Yale Law, J* 94, 1395–1415.
- Walker, L.J. (1984). Seks Differences in the Development of Moral Reasoning: A Critical Review. *Child Development*, 55(3), 677-691.
- You, D., Yukiko, M. en Bebeau, M.J. (2011). Gender Differences in Moral Sensitivity: A Meta-Analysis. *Ethics en Behavior*, 21(4), 263-282.
- Young, L. en Koenigs, M. (2007). Investigating emotion in moral cognition: a review of evidence from functional neuroimaging and neuropsychology. *British Medical Bulletin*, 84, 69-79.

## **Bijlage A: Nederlandse versie materiaal**

VERSIE 1

### **Bacheloronderzoek Anke, Fenne en Nina (Radboud Universiteit Nijmegen)**

Heel erg fijn dat je ons wilt helpen door mee te doen aan ons onderzoek. Je gaat hierna drie korte verhaaltjes lezen en daar enkele vragen over beantwoorden. Let goed op de volgende vier punten:

1. Je hebt maximaal **10 minuten** de tijd om alles volledig in te vullen.
2. Let goed op dat je alles zo volledig mogelijk invult.
3. Sla de pagina pas om als je **helemaal** klaar bent met die pagina. Je mag **niet** terugbladeren.
4. Blader **niet** vooruit.

**Door hieronder je handtekening te zetten, geef je ons toestemming om jouw gegevens anoniem te gebruiken.**

**Handtekening:** \_\_\_\_\_

**INSTRUCTIE: Lees onderstaande dilemma's en beantwoord de vragen.****Dilemma Frank**

Frank staat op een brug met daaronder een spoorlijn. Hij ziet een op hol geslagen trein aankomen. Verderop staan vijf mensen op het spoor die de trein niet horen aankomen en die niet op tijd van het spoor af kunnen komen. De enige manier om de op hol geslagen trein te stoppen, is om iets zwaars op het spoor te laten vallen. Maar het enige zware wat er is, is een heel zware man die ook op de brug staat. Frank kan de man op het spoor duwen, waardoor hij hem doodt en de vijf mensen verderop op het spoor redt of hij kan de man niet duwen, maar dan zullen de andere vijf mensen dood gaan.

Hoe goed of slecht is het als Frank de man op het spoor duwt?

Zeer slecht	Slecht	Enigszins slecht	Noch goed, noch slecht	Enigszins goed	Goed	Zeer goed
1	2	3	4	5	6	7

Leg hieronder kort uit waarom je dit denkt.

.....

.....

.....

**Dilemma Joe**

Joe zit in zijn motorboot aan de monding van een rivier als hij een haai aan ziet komen. Verderop ziet hij vijf zwemmers in de rivier. Als Joe blijft waar hij is, zal zijn boot de rivier blokkeren en voorkomen dat de haai de vijf zwemmers aanvalt. Joe ziet ook een andere zwemmer in de verte verdrinken. Joe kan die ene zwemmer redden, maar dan zal de haai de rivier in kunnen zwemmen. Als Joe naar die ene zwemmer vaart met zijn boot, dan zal die zwemmer het overleven, maar zullen de vijf zwemmers opgegeten worden door de haai. Als Joe blijft waar hij is, dan zal die ene zwemmer verdrinken maar zullen de andere vijf zwemmers veilig zijn.

Hoe goed of slecht is het als Joe blijft waar hij is?

Zeer slecht	Slecht	Enigszins slecht	Noch goed, noch slecht	Enigszins goed	Goed	Zeer goed
1	2	3	4	5	6	7

Leg hieronder kort uit waarom je dit denkt.

.....

.....

.....

**Dilemma Tom**

Tom is een gevechtspiloot die op weg is naar een gekaapt vliegtuig. De kapers zijn op weg naar een voetbalstadion om daar neer te storten op 70.000 mensen. De enige manier om de kapers te stoppen is om het gekaapte vliegtuig uit de lucht te schieten, maar dan gaan de 164 onschuldige passagiers aan boord dood. Als Tom het vliegtuig niet uit de lucht schiet, dan stort het vliegtuig neer op het stadion en gaan er 70.000 mensen dood.

Hoe goed of slecht is het als Tom het vliegtuig uit de lucht schiet?

Zeer slecht	Slecht	Enigszins slecht	Noch goed, noch slecht	Enigszins goed	Goed	Zeer goed
1	2	3	4	5	6	7

Leg hieronder kort uit waarom je dit denkt.

.....

.....

.....

## **Bijlage B: Engelse versie materiaal**

VERSION 1

### **Bachelorthesis Anke, Fenne and Nina (Radboud University Nijmegen)**

Thank you for helping us out by participating in our experiment. You will read three short stories and you'll have to answer questions about them. Please keep the following four points in mind:

1. You have a maximum of **10 minutes** to complete everything.
2. Make sure you fill in all the questions as accurately as possible.
3. Only turn to the next page when you are **completely** finished. Please **do not** go back to previous pages.
4. **Do not** go to the next page when you're not finished yet.

**By signing here, you give us your permission to use your data anonymously.**

**Signature:** \_\_\_\_\_



Participant code: .....

VERSION 1

**INSTRUCTION: Please read the following dilemmas and answer the questions.**

**Dilemma Frank**

Frank is on a bridge that passes over train tracks. He sees an out-of-control train coming down the track. Further down the track there are five people who will not hear the train and will not be able to get off the track in time. The only way to stop this out-of-control train is to drop a very heavy weight into its path. But the only available, heavy weight is a very heavy man who is also on the bridge. Frank can push the man onto the track in the path of the train, killing him; or he cannot push the man, letting the five people die.

How right or wrong would it be for Frank to push the man onto the track?

Very wrong	Wrong	Slightly wrong	Neither wrong nor right	Slightly right	Right	Very right
1	2	3	4	5	6	7

Please explain briefly why you think this.

.....

.....

.....

**Dilemma Joe**

Joe is in his motorboat at the mouth of a river when he notices a shark approaching. Further down the river he sees five swimmers. If Joe stays where he is, his boat will block the river and prevent the shark from attacking the five swimmers. Joe also sees another swimmer drowning in the distance. Joe can save the one swimmer, but that will leave the river open to the shark. If Joe moves his motorboat towards the one swimmer, the drowning swimmer will live but the five other swimmers will be eaten by the shark. If Joe stays where he is, the one swimmer will drown but the five swimmers will be safe.

How right or wrong would it be for Joe to stay where he is?

Very wrong	Wrong	Slightly wrong	Neither wrong nor right	Slightly right	Right	Very right
1	2	3	4	5	6	7

Please explain briefly why you think this.

.....  
.....  
.....

**Dilemma Tom**

Tom is a fighter pilot who is on his way to a hijacked airplane. The hijackers are on their way to a soccer stadium. They will crash on 70.000 people. The only way to stop the hijackers is to shoot the hijacked airplane out of the air, but that means that the 164 innocent passengers will die. If Tom doesn't shoot the airplane out of the air, then it will crash on the stadium and 70.000 people will die.

How right or wrong would it be for Tom to shoot the airplane out of the air?

Very wrong	Wrong	Slightly wrong	Neither wrong nor right	Slightly right	Right	Very right
1	2	3	4	5	6	7

Please explain briefly why you think this.

.....  
.....  
.....

**Bijlage C: Controlevragen Engelse versie**

Leg kort in je eigen woorden uit wat er gebeurde in het eerste dilemma van Frank en de brug. Gebruik maximaal drie zinnen.

.....  
.....  
.....  
.....  
.....  
.....

Leg kort in je eigen woorden uit wat er gebeurde in het tweede dilemma van Joe en de boot. Gebruik maximaal drie zinnen.

.....  
.....  
.....  
.....  
.....  
.....

Leg kort in je eigen woorden uit wat er gebeurde in het derde dilemma van Tom en het vliegtuig. Gebruik maximaal drie zinnen.

.....  
.....  
.....  
.....  
.....  
.....

## Bijlage D: Aanvullende vragen voor participanten

**INSTRUCTIE: Vul onderstaande vragenlijst zo nauwkeurig mogelijk in.**

Geboortedatum: ...../...../.....  
 Geslacht: man / vrouw  
 In welk land ben je geboren? .....

Wat is je moedertaal (of talen)?  
 1. ....  
 2. ....  
 3. ....

Welke taal (of talen) spreek je thuis?  
 1. ....  
 2. ....  
 3. ....

### Kennis van talen

Welke talen kun je goed en gebruik je? Nederlands en Engels staan al ingevuld. Andere talen mag je zelf invullen. Geef jezelf een cijfer en noteer dat in de tabel. Gebruik hiervoor de volgende schaal:

Niet: 1    2    3    4    5    6    7    8    9    10    Vloeiend

Taal	Spreken	Luisteren	Schrijven	Lezen
Nederlands				
Engels				

Hoe oud was je toen je begon met het leren van deze talen? Vul per taal in hoe oud je was toen je begon met het leren van de taal.

Taal	Leeftijd
Nederlands	
Engels	

## Bijlage E: Uitkomsten normaliteit per taal, geslacht en versie

Tabel 2: Uitkomsten normaliteit per taal, geslacht en versie

Taal	Geslacht	Versie	Kolmogorov-Smirnov	Shapiro-wilk	Skewness	Kurtosis
NL	Man	Brug	.030	.057	.307	-.127
NL	Man	Boot	.002	.013	-.723	.137
NL	Man	Vliegtuig	.007	.020	-.556	.163
NL	Vrouw	Brug	-	-	.438	-.688
NL	Vrouw	Boot	-	-	-.162	1.014
NL	Vrouw	Vliegtuig	-	-	-.332	.415
ENG	Man	Brug	-	-	-.104	-.953
ENG	Man	Boot	-	-	-1.149	3.009
ENG	Man	Vliegtuig	-	-	-1.166	1.436
ENG	Vrouw	Brug	.013	.051	-.100	-1.020
ENG	Vrouw	Boot	.007	.015	-.771	.745
ENG	Vrouw	Vliegtuig	.000	.000	-2.059	6.379

In tabel 1 is te zien dat de Nederlandse versies ingevuld door mannen alle drie niet normaal verdeeld zijn. Bij de versie met de brug geeft de Shapiro-Wilk toets aan dat het wel net verdeeld is, maar de Kolmogorov-Smirnov geeft aan van niet. Van de Nederlandse versies ingevuld door vrouwen, is alleen de versie met de boot niet normaal verdeeld, de andere twee wel. Van de Engelse versie ingevuld door mannen zijn de versie met de boot en met het vliegtuig niet normaal verdeeld. Van de Engelse versies ingevuld door vrouwen, zijn de versies met de brug en met het vliegtuig niet normaal verdeeld.

## Bijlage F: Uitkomsten normaliteit na log per taal, geslacht en versie

Tabel 3: Uitkomsten normaliteit na log per taal, geslacht en versie

Taal	Geslacht	Versie	Kolmogorov- Smirnov	Shapiro-wilk	Skewness	Kurtosis
NL	Man	Brug	.000	.000	-.862	.163
NL	Man	Boot	.000	.000	-2.030	4.532
NL	Man	Vliegtuig	.000	.000	-1.931	4.196
NL	Vrouw	Brug	-	-	-.736	-.734
NL	Vrouw	Boot	-	-	-1.272	1.261
NL	Vrouw	Vliegtuig	-	-	-1.715	3.466
ENG	Man	Brug	-	-	-.929	.025
ENG	Man	Boot	-	-	-3.105	15.416
ENG	Man	Vliegtuig	-	-	-2.033	4.381
ENG	Vrouw	Brug	.000	.000	-.909	.142
ENG	Vrouw	Boot	.000	.000	-1.000	.303
ENG	Vrouw	Vliegtuig	.000	.000	-.997	.124