

De invloed van genderidentificatie op de relatie tussen stereotiepe bedreiging en wiskundeprestaties van vrouwelijke studenten

Student: R.J Gerbens (s4108051)

Begeleider en eerste beoordelaar: dr. A. H. Stoevenbelt

Tweede beoordelaar: prof. dr. D.D. van Bergen

Rijksuniversiteit Groningen

Faculteit der Gedrags- en Maatschappijwetenschappen

Bachelorwerkstuk Pedagogische Wetenschappen

Juni 2025

Aantal woorden: 4979

Abstract

This study examined whether gender identification moderates the effect of stereotype threat on math performance among female university students. Previous research showed that women underperform on math tasks when gender stereotypes are activated, particularly when gender is an important part of their identity. The current study features that data of Stoevenbelt et al. (2025), who replicated Johns et al. (2005), using a cross-sectional, experimental and quantitative approach.

During the study, 1,275 female students from five countries were randomly assigned to one of three conditions: Stereotype threat, or one of two control conditions: Teaching intervention, or problem-solving. Math performance was measured by a standardized test, gender and gender identification by self-report.

Indicating an effect in the opposite direction of what was hypothesized, participants in the stereotype threat condition performed, on average, slightly better than those in the control conditions. Women who identified more strongly with their gender tended to perform slightly worse ($B = -.048$, $p = .009$). However, the interaction was not statistically significant ($B = .011$, $SE = .062$, $t = .183$, $p = .855$), indicating that gender identification did not moderate the effect of stereotype threat. These findings suggest that stereotype threat may not consistently affect math performance, but gender identification does.

Keywords: stereotype threat, genderidentification, math performance, regression analysis

Wist je dat vrouwen die geconfronteerd worden met negatieve wiskunde stereotypen gemiddeld slechter presteren op toetsen dan mannen? Het fenomeen waarbij vrouwen ondermaats presteren in wiskunde vanwege de angst om bevestigd te worden in negatieve stereotypen, staat bekend als “stereotiepe bedreiging”. Volgens onderzoek van Steele en Aronson (1995) betekent stereotiepe bedreiging, de angst om negatieve stereotypen over een specifieke groep te bevestigen. Negatieve stereotypen kunnen bij sommige vrouwen leiden tot het geloof dat zij niet goed zijn in wiskunde (Eccles et al., 1999).

Uit eerder onderzoek kwam naar voren dat vrouwen slechter presteerden op moeilijke wiskunde toetsen wanneer hen verteld werd dat de test genderverschillen meet, terwijl dit niet zo was wanneer dergelijke informatie niet werd verstrekt (Spencer et al., 1999). Daarbij geeft eerder onderzoek ook aan dat stereotiepe bedreiging het gevoel van erbij horen kan verminderen, het gevoel van uitsluiting en isolatie kan vergroten en dit kan vervolgens mogelijk leiden tot ontkoppeling van het domein. Stereotiepe bedreiging heeft ook een significant positief effect op de kans dat vrouwen een bèta-, technologie-, technische of wiskundige studie verlaten (Beasley & Fischer, 2012; Casad et al., 2017; Cheryan et al., 2009; Woodcock et al., 2012). Vrouwen die een laag zelfbeeld hebben wat betreft wiskunde zijn dus ook minder geneigd om een studie te volgen op het gebied van wetenschap, technologie, ingenieurswetenschappen en wiskunde (STEM-studie, Jiang et al., 2020). Dus de maatschappelijke relevantie van deze studie is: vrouwen zijn ondervertegenwoordigd in technische en bèta-opleidingen, terwijl carrières in deze sector vaak beter betaald worden en meer professioneel prestige bieden (Brown & Leaper, 2010). Het is essentieel om begrip van de relatie tussen stereotiepe bedreiging en wiskundeprestaties te krijgen, om belemmeringen voor vrouwen in STEM te doorbreken.

Stereotiepe bedreiging blijkt ook samen te hangen met een grotere kans dat vrouwen afhaken in STEM-richtingen, vooral bij een laag wiskundig zelfbeeld (Jiang et al., 2020). De maatschappelijke relevantie van dit onderwerp is: vrouwen zijn ondervertegenwoordigd in STEM-richtingen, terwijl carrières in deze sectoren vaak beter betaald worden en meer professioneel prestige bieden. Begrip van de relatie tussen stereotiepe bedreiging en wiskundeprestaties is daarom essentieel om barrières voor vrouwen in STEM te doorbreken.

Deze meta-analyse laat zien dat er verschillende moderatoren zijn. Een die is meegenomen in het onderzoek van Stoevenbelt et al. (2025) is genderidentiteit. Genderidentiteit is ook wel de mate waarin een persoon zichzelf ziet als een typische vertegenwoordiger van zijn gender en dit belangrijk vindt (Selimbegovic et al., 2019). Uit onderzoek van Schmader (2002) kwam naar voren dat vrouwen die zich sterk identificeerden met hun gender, kwetsbaarder waren voor de negatieve effecten van stereotiepe bedreiging. Vermoedelijk doordat zij het belangrijker vonden om negatieve stereotypen over hun gendergroep te weerleggen, waardoor de druk om te presteren groter werd. Dit verhoogde niveau van stress belastte cognitieve hulpbronnen, zoals het werkgeheugen, wat leidde tot verminderde

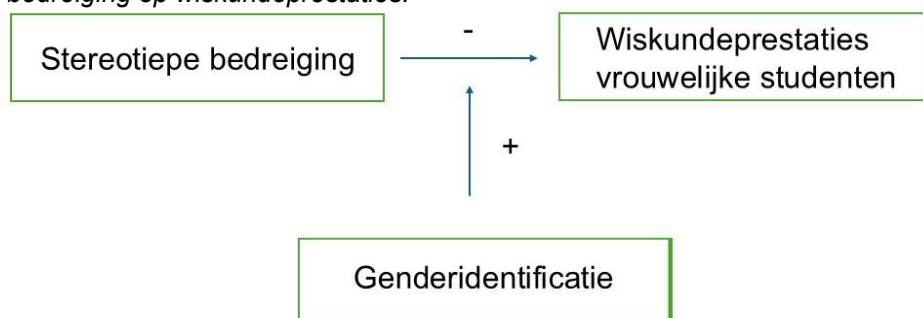
prestaties op taken die relevant zijn voor het stereotiepe, zoals wiskunde. Schmader (2002) toonde aan dat vrouwen voor wie vrouw-zijn belangrijk was, meer stereotiepe bedreiging ervoeren op een wiskundetest dan vrouwen voor wie gender minder centraal stond, althans wanneer hen werd verteld dat de vaardigheden van vrouwen in het algemeen zouden worden beoordeeld op basis van de testcores van deelnemers. Verder concludeerden Shapiro en Neuberg (2007) dat vrouwen die zich sterker met hun gender identificeerden: meer persoonlijke dreiging voelden, meer groepsverantwoordelijkheid ervoeren en het stereotiepe sneller activeerden. Dit leidde tot hetzelfde als wat Schmader (2002) benoemde.

Aangezien, naast de genoemde studies, nog weinig onderzoek is gedaan naar het effect van de moderator genderidentificatie op de relatie tussen stereotiepe bedreigingen wiskundeprestaties van vrouwen is het van wetenschappelijk belang dat hiernaar wordt gekeken. Het kan mogelijk bijdragen aan het begrijpen van hoe sociale identiteiten en stereotypen cognitieve prestaties beïnvloeden. Momenteel duiden meerdere meta-analyses (Nguyen & Ryan, 2008; Doyle & Voyer, 2016) dus aan dat vrouwelijke studenten lager presteren op wiskunde vanwege stereotiepe bedreiging. Echter is er minder bekend over de invloed van genderidentificatie op dit fenomeen. Het doel hiervan is dan ook meer inzicht krijgen in dit probleem door onder andere het onderzoeken van de moderator genderidentificatie en de invloed hiervan.

Het doel van deze studie is onderzoeken of de genderidentificatie invloed heeft op de relatie tussen stereotiepe bedreiging en de wiskundeprestaties van vrouwen. De mate waarin iemand zich een vrouw voelt kan dus mogelijk invloed hebben op de prestaties in wiskunde wanneer gekeken wordt naar de stereotiepe bedreiging. Dit wordt onderzocht met de volgende onderzoeksvraag: “In hoeverre beïnvloedt genderidentificatie de mate waarin stereotiepe bedreiging effect heeft op wiskundeprestaties van vrouwelijke studenten?”. Zoals weergegeven in Figuur 1 en op basis van eerder onderzoek wordt verwacht dat genderidentificatie de negatieve relatie tussen stereotiepe bedreiging en wiskundeprestaties van vrouwelijke studenten versterkt. Met andere woorden, de hypothese luidt: hoe sterker een vrouwelijke student zich identificeert met haar gender, hoe groter de negatieve impact van stereotiepe bedreiging op haar wiskundeprestaties.

Figuur 1

Conceptueel model van de verwachte modererende relatie tussen genderidentificatie en stereotiepe bedreiging op wiskundeprestaties.



Methode

Onderzoeksdesign

De data van deze studie kwam uit een replicatieonderzoek van Stoevenbelt et al. (2025), waarbij de studie van Johns et al. (2005) gerepliceerd werd. Het had een experimenteel, cross-sectioneel onderzoeksdesign. Dit onderzoek werd goedgekeurd door de Ethische Toetsingscommissie van Tilburg University (referentienummer: EC-2017.104).

Populatie en steekproef

De doelpopulatie van dit onderzoek waren vrouwelijke studenten. De deelnemers werden geworven via universiteiten en konden zich vrijwillig aanmelden in ruil voor studiepunten of een financiële vergoeding. Vanwege deze manier van werven, op basis van beschikbaarheid en vrijwillige aanmelding, betrof het een niet-willekeurige gelegenheidssteekproef. Deelnemers moesten tussen de 18 en 25 jaar oud zijn en tenminste één cursus in de wiskunde, statistiek of onderzoeksmethodologie hebben gevolgd. De steekproef bestond uit 1275 vrouwelijke deelnemers, afkomstig uit de Verenigde Staten, Slowakije, Duitsland, Oostenrijk en Nederland.

Procedure

De deelnemers voor het onderzoek werden geworven uit deelnemersgroepen van bachelorstudenten van acht universiteiten. Zij werden in kleine groepen in een klaslokaal getest waarbij telkens minimaal één man aanwezig was. Deze opzet was bedoeld om de aanwezigheid van een beoordelende omgeving te benadrukken, wat relevant was voor het onderzoeken van stereotiepe bedreiging. Er werden uitsluitend vrouwelijke deelnemers geïnccludeerd. De deelnemersgroepen werden vervolgens willekeurig toegewezen aan één van de drie experimentele condities. De ene helft aan de experimentele stereotiepe bedreiging conditie en de andere helft werd gelijk verdeeld over de twee controlecondities: onderwijsinterventie- en de probleemoplossende conditie.

Experimentele manipulatie

Binnen de groepen onderging elke deelnemer dezelfde manipulatie. De stereotiepe bedreiging conditie hield in dat deelnemers te horen kregen dat ze een wiskundetoets zouden maken die bedoeld was om hun capaciteiten te meten. Deze instructie was bedoeld om het negatieve stereotiepe over vrouwen en wiskunde te activeren, waardoor de deelnemers zich bewust werden van mogelijke beoordeling op basis van hun gender. Dit bewustzijn kon leiden tot verhoogde angst en verminderde prestaties. Bij de onderwijsinterventie conditie werd de wiskundetoets op dezelfde manier gepresenteerd als in de stereotiepe bedreiging conditie. Echter, deelnemers kregen daarnaast

informatie over het fenomeen stereotiepe bedreiging en hoe dit hun prestaties kon beïnvloeden. Door deze bewustwording konden deelnemers hun angst toeschrijven aan externe factoren in plaats van aan een gebrek aan bekwaamheid, wat mogelijk voorkwam dat hun prestaties verslechterden. Als laatste werden de deelnemers bij de probleemoplossende conditie verteld dat ze een probleemoplossende taak zouden uitvoeren, zonder vermelding van wiskunde of beoordeling. Deze benadering was bedoeld om de activering van negatieve stereotypen te vermijden, waardoor deelnemers mogelijk minder angst ervoeren en daardoor niet slechter presteerden.

De experimentleider gaf een korte introductie. Vervolgens vroeg hij naar geïnformeerde toestemming en voerde de experimentele manipulatie uit door een audio-opname van een mannenstem af te spelen. Deze opname bevatte de toetsbeschrijving die varieerde afhankelijk van de conditie: stereotiepe bedreiging, onderwijsinterventie of probleemoplossende taak. Deze opnames speelde in op de bovengenoemde inhoudelijke verschillen. Na het afspelen van de audio herhaalde de experimentleider de instructies. De deelnemers in de stereotiepe bedreiging conditie en onderwijsinterventie conditie verstrekten na de manipulatie demografische informatie (zoals gender) en gaven hun etniciteit aan nadat zij de wiskundetoets hadden voltooid. Deelnemers in de probleemoplossende conditie verstrekten deze informatie pas na het voltooien van de wiskundetoets, om mogelijke effecten van stereotiepe bedreiging te minimaliseren (Steele & Aronson, 1995). De wiskundetoets duurde 20 minuten, waarbij deelnemers kladpapier mochten gebruiken, maar geen rekenmachine. Na afloop van de toets vulden de deelnemers een vragenlijst in over genderidentificatie (Luhtanen & Crocker, 1992; Schmader, 2002), een meting van stereotiepe bewustzijn en algemene vragen over het experiment. Tot slot werden de deelnemers gecompenseerd en kregen ze instructie om niet met medestudenten over de aard van het onderzoek te spreken.

Variabelen en instrumenten

In dit onderzoek werd gefocust op een aantal variabelen, namelijk wiskundeprestatie, genderidentificatie, stereotiepe bedreiging en de interactie tussen genderidentificatie en stereotiepe bedreiging.

Wiskundeprestatie

De variabele *wiskundeprestatie* werd gemeten aan de hand van de proportionele totaalscore op de wiskundetoets, berekend als het aantal correct beantwoorde vragen gedeeld door het totaal aantal gemaakte vragen. De scores varieerden van 0 tot 1, waarbij een hogere score duidde op betere prestaties. Een item werd als beantwoord beschouwd als de deelnemer ten minste één antwoordoptie had aangekruist. Wanneer een item niet gemaakt was, werd deze fout gerekend.

Om de wiskundeprestaties te beoordelen waren er twee wiskundetoetsen (een makkelijkere en

een reguliere toets). Deelnemers maakten een van de twee testen, afhankelijk van de onderzoekslocatie. De twee toetsen hadden beide 30 Graduate Records Examination-items (een gestandaardiseerde Amerikaanse wiskundetest; eenvoudigere versie: <https://osf.io/9y73x>; reguliere versie: <https://osf.io/x952c>), om ervoor te zorgen dat de toetsen op elke onderzoekslocatie moeilijk genoeg waren om een effect van stereotiepe bedreiging op te roepen (Keller, 2007; Spencer et al., 1999). Beide toetsen bestonden uit 25 items en hadden vijf unieke items. Om scores tussen verschillende labs en testversies vergelijkbaar te maken, werden de ruwe totaalscores gelijkgesteld met behulp van de chained linear equating methode, zoals beschreven door Kolen en Brennan (2014, pp. 122-123). Deze methode stelde scores van verschillende testversies op een vergelijkbare schaal, waardoor eerlijke vergelijkingen mogelijk waren.

Alle items waren meerkeuzevragen, met vijf antwoordopties en één juist antwoord. De items bestonden uit vragen over onderwerpen uit algebra, meetkunde en statistiek. De steekproefbetrouwbaarheid van de eenvoudigere en reguliere versies varieerde van Cronbachs $\alpha = .83$ ($N = 824$, gebaseerd op de 30 eenvoudigere items) tot $\alpha = .88$ ($N = 678$, gebaseerd op de 30 reguliere items). Deze werd vastgesteld op de volledige hoeveelheid deelnemers, dus mannen inbegrepen.

Gender

Gender werd gemeten als zelf-geïdentificeerde genderidentiteit, met drie opties man/vrouw/overig. In het originele onderzoek werden deelnemers die zich niet als man of vrouw identificeerden of geen antwoord op de gendervraag hebben gegeven al geëxcludeerd.

Genderidentiteit

Om de moderator genderidentificatie te beoordelen werd de genderidentificatie schaal gebruikt (Luhtanen & Crocker, 1992; Schmader, 2002). De schaal bestond uit vier items (bijv. 'Een vrouw/man zijn is een belangrijk onderdeel van mijn zelfbeeld') en werd gemeten op een 5-punts Likertschaal (variërend van 1 = *helemaal oneens* tot 5 = *helemaal eens*, Cronbach's $\alpha = .77$, gebaseerd op $N = 1275$). De variabele *genderidentiteit* vertegenwoordigde de gemiddelde score van de vier schaalvragen die de mate van identificatie met het eigen gender mat. Voor de regressieanalyse werd genderidentificatie gecentreerd (door het gemiddelde af te trekken van iedere score), zodat de hoofdeffecten in het model geïnterpreteerd konden worden bij een gemiddeld niveau van genderidentificatie. Dit maakte de interpretatie van het effect van stereotiepe bedreiging zuiverder, omdat deze dan niet vertekend werd door een arbitrair nulpunt van de modererende variabele. Bovendien verminderde centreren de multicollineariteit tussen de hoofdeffecten en de interactieterm.

Experimentele conditie

De variabele conditie was een dummyvariabele waarbij 0 stond voor deelname aan één van de twee controlecondities (probleemoplossende taak of onderwijsinterventie) en 1 voor deelname aan de stereotiepe bedreiging conditie. Zoals hierboven beschreven in de procedure, kregen de deelnemers in de stereotiepe bedreiging conditie te horen dat de wiskundetaak bedoeld was om hun cognitieve capaciteiten te meten. Deze instructie was bedoeld om het stereotiepe bewustzijn te activeren.

Interactie tussen stereotiepe bedreiging en genderidentificatie

Er werd een interactieterm (stereotiepe bedreiging x genderidentificatie) toegevoegd aan het regressiemodel. Deze werd gecreëerd door het product te nemen van de experimentele conditie en de gecentreerde score op genderidentificatie.

Analyseplan

Deze studie onderzocht alleen de vrouwelijke studenten ($N = 1275$). Dit werd gedaan met een multiple lineaire regressieanalyse, met behulp van SPSS (versie 28). Er werd gekozen voor een multiple lineaire regressieanalyse omdat op deze manier onderzocht kon worden hoe verschillende factoren, zoals stereotiepe bedreiging en genderidentificatie, wiskundeprestaties zouden kunnen beïnvloeden bij vrouwelijke studenten. Om te beginnen werd per groep de beschrijvende statistieken berekend, met daarbij de gemiddelden en de standaarddeviaties. Daarnaast werden ontbrekende waarden ($N = 4$) via listwise deletion behandeld en dus buiten de analyse gelaten. Na uitsluiting van ontbrekende waarden werden de analyses gebaseerd op $N = 1271$. Ook werden uitbijters geïnterpreteerd via een scatterplot, maar werden er geen correcties of uitsluitingen toegepast. De afhankelijke variabele was wiskundeprestatie; de onafhankelijke variabele was stereotiepe bedreiging, gecodeerd als dummyvariabele (0 = controlecondities, 1 = stereotiepe bedreiging conditie); en de moderator was genderidentificatie. De multiple lineaire regressieanalyse werd in twee stappen uitgevoerd. In de eerste stap werden stereotiepe bedreiging en genderidentificatie als hoofdeffecten opgenomen in het model. In de tweede stap werd de interactieterm tussen stereotiepe bedreiging en genderidentificatie toegevoegd om te kijken of genderidentificatie invloed uitoefende op de relatie tussen stereotiepe bedreiging en wiskundeprestaties. Hierna werden de R^2 -waarden vergeleken om te beoordelen hoeveel extra variantie in wiskundeprestaties werd verklaard door het toevoegen van de interactieterm tussen stereotiepe bedreiging en genderidentificatie. Om de significantie van de interactieterm te testen werd deze beoordeeld op een significantieniveau van $\alpha = 0,05$.

Er werd gecontroleerd of aan de volgende assumpties werd voldaan: lineariteit, onafhankelijkheid van residuen, normale verdeling van residuen, homoscedasticiteit en geen multicollineariteit. In de dataset was sprake van een multilevel-structuur. Doordat de conditie per groep

werd toegewezen en niet aan individuele proefpersonen, kon niet worden aangenomen dat de observaties volledig onafhankelijk van elkaar waren. In onderzoek van Stoevenbelt et al. (2025) was besloten de multilevel-structuur niet te modelleren. Verder werden de lineariteit en homoscedasticiteit beoordeeld aan de hand van een spreidingsdiagram van de gestandaardiseerde residuen tegenover de gestandaardiseerde voorspelde waarden. Ook werd de normaliteit van de residuen getoetst met een normal P-P plot van de gestandaardiseerde residuen en de multicollineariteit gecontroleerd aan de hand van de VIF (Variance Inflation Factor). Uiteindelijk werd de effectgrootte (Cohen's f^2) gerapporteerd om te bepalen hoeveel extra variantie de interactieterm verklaarde in wiskundeprestaties.

Resultaten

Beschrijvende statistiek

De steekproef bestond uit 1275 vrouwelijke studenten. Van de 1275 deelnemers hadden vier vrouwen de lijst met vragen over genderidentificatie niet volledig ingevuld. Deze gevallen werden uitgesloten van analyses waarin genderidentificatie werd meegenomen (listwise deletion). De gemiddelde scores op wiskundeprestatie en genderidentificatie per experimentele conditie zijn weergegeven in Tabel 1. De beschrijvende statistieken lieten zien dat de groepen weinig van elkaar verschilden in genderidentificatie en dit suggereert dat het willekeurig verdelen van de deelnemers op deze variabele goed is verlopen. Opvallend was dat de deelnemers in de stereotiepe bedreiging conditie gemiddeld iets hogere scores behaalden op de wiskundetoets dan deelnemers in de twee controlecondities. Dit resultaat is tegenstrijdig met eerdere bevindingen over stereotiepe bedreiging, waarin vaak lagere prestaties bij vrouwen werden gevonden onder stereotiepe bedreigende instructies. Daarbij viel ook op dat de spreiding (SD) in wiskundeprestaties bij de conditie stereotiepe bedreiging het grootst was. Dit geeft aan dat deelnemers binnen deze conditie meer verschillend presteerden dan in de andere condities, mogelijk als gevolg van individuele gevoeligheid voor stereotiepe bedreiging.

Correlatie tussen Genderidentificatie en Wiskundeprestaties

In de eerste analyse werd een zwakke, maar significante negatieve correlatie gevonden tussen genderidentificatie en wiskundeprestaties ($r = -.074$, $p = .008$). Dit wijst erop dat vrouwen die zich sterker met hun gender identificeren, gemiddeld iets lagere prestaties behalen op de wiskundetoets. Hoewel het effect statistisch significant was, was de effectgrootte klein (Cohen, 1992). Deze bevinding vormde aanleiding om de relatie tussen genderidentificatie en wiskundeprestatie verder te toetsen binnen een regressiemodel, waarin ook de experimentele conditie werd meegenomen als voorspeller.

Tabel 1*Beschrijvende statistiek van de variabelen per conditie*

Behandeling		Wiskundeprestatie	Genderidentificatie
PS	Mean	,51	3,22
	N	332	332
	Std. Deviation	,28	,51
ST	Mean	,54	3,20
	N	591	588
	Std. Deviation	,32	,46
TI	Mean	,51	3,23
	N	352	351
	Std. Deviation	,32	,45
Total	Mean	,52	3,21
	N	1275	1271
	Std. Deviation	,31	,47

Noot. PS = probleemoplossende conditie; ST = stereotiepe bedreiging conditie; TI = onderwijsinterventie conditie.

Regressieanalyse

Er werd een lineaire regressieanalyse uitgevoerd om te onderzoeken in hoeverre genderidentificatie en conditie stereotiepe bedreiging gezamenlijk bijdragen aan het voorspellen van wiskundeprestaties bij vrouwelijke studenten. Het regressiemodel (model 1) bleek slechts een zeer klein deel van de variantie in wiskundeprestaties te verklaren ($R^2 = .007$, adjusted $R^2 = .006$), wat overeenkomt met 0,7% verklaarde variantie. Dit betekent dat 0,7% van de verschillen in wiskundeprestaties verklaard worden door de conditie (stereotiepe bedreiging versus controle) en de mate van genderidentificatie. Het verschil tussen R^2 en adjusted R^2 was heel klein, wat aangeeft dat de verklarende kracht zeer beperkt is.

Hoewel er een zwakke negatieve correlatie werd gevonden tussen genderidentificatie en wiskundeprestaties ($r = -.074$), bleek uit de regressieanalyse dat genderidentificatie ook een unieke voorspeller was, onafhankelijk van de experimentele conditie. De regressiecoëfficiënten van de analyse lieten zien dat genderidentificatie een significante negatieve voorspeller was van wiskundeprestatie ($B = -.048$, $SE = .019$, $\beta = -.073$, $t = 2,609$, $p = .009$). Dit betekent dat deelnemers met een sterkere identificatie met hun gender gemiddeld iets lager scoren op de wiskundetoets. De conditie (stereotiepe bedreiging versus controle) was geen significante voorspeller ($B = .027$, $SE = .018$, $\beta = .043$, $t = 1.549$, $p = .122$), hoewel de richting van het effect positief was. Dit betekent dat de groep met stereotiepe bedreiging (1) gemiddeld iets hoger scoort op de wiskundetoets dan de controlegroepen PS en TI (0).

Regressieanalyse met interactieterm

Ondanks dat de hoofdeffecten in het model relatief zwak of niet significant waren, werd ervoor gekozen om toch een regressiemodel met interactieterm (model 2) te schatten. Op deze manier werd de hypothese getoetst en sloot het aan bij de onderzoeksvraag. Dus werd in een tweede regressiemodel een interactieterm toegevoegd om te onderzoeken of genderidentificatie het effect van stereotiepe bedreiging op wiskundeprestaties modereerde. De predictoren in dit model waren: conditie (stereotiepe bedreiging vs controle), genderidentificatie en hun interactie. Met dit model werd gemeten of het effect van stereotiepe bedreiging op wiskundeprestatie veranderde, afhankelijk van hoe sterk iemand zich identificeerde met haar gender.

Het regressiemodel met conditie, genderidentificatie en hun interactieterm was significant ($p = .024$) en verklaarde 0,7% van de variantie in wiskundeprestaties ($R^2 = .007$, adjusted $R^2 = .005$). De toevoeging van de interactieterm leidde tot een significante verbetering van het model, maar geen toename in verklaarde variantie ($\Delta R^2 = 0$, $p = .024$). Dit suggereert dat de combinatie van conditie en genderidentificatie mogelijk van invloed is op wiskundeprestaties. Bij nadere beschouwing, te zien in Tabel 2, blijkt dat de interactieterm tussen conditie en genderidentificatie niet significant is ($B = .011$, $SE = .062$, $t = .183$, $p = .855$), wat suggereert dat genderidentificatie het effect van stereotiepe bedreiging op prestaties niet modereert. Met andere woorden betekent dit dat het effect van stereotiepe bedreiging niet verschilt afhankelijk van de mate van genderidentificatie.

De R^2 van het regressiemodel bleef gelijk na toevoeging van de interactieterm ($R^2 = .007$ in beide modellen). Dit resulteerde in een Cohen's f^2 van 0, wat aangeeft dat de interactie tussen stereotiepe bedreiging en genderidentificatie geen extra voorspellende waarde heeft en dus niet bijdraagt aan een betere verklaring van de wiskundeprestaties. Verschillen in wiskundeprestaties zijn daarom waarschijnlijk het resultaat van andere factoren die buiten dit model vallen. Daarbij is het zo dat door de grote steekproefomvang zelfs kleine effecten statistisch significant kunnen worden.

Tabel 2

Regressiemodel met stereotiepe bedreiging, genderidentificatie en de interactieterm

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,51	,01		42,57	<.001		
	Stereotiepe bedreiging	,03	,02	,04	1,55	,122	1,00	1,00
	Genderidentificatie	-,05	,04	-,08	-1,49	,138	,26	3,86
	Interactie	,01	,06	,01	,18	,855	,26	3,86

*Afhankelijke variabele: Wiskundeprestaties

Assumpties

Om de betrouwbaarheid en validiteit van de regressieresultaten te waarborgen, is gecontroleerd of voor beide modellen (met en zonder interactieterm) aan de onderliggende assumpties van multiple lineaire regressie werd voldaan. Wanneer één of meerdere assumpties geschonden werden, moeten de resultaten van de analyses met meer voorzichtigheid worden geïnterpreteerd.

Onafhankelijkheid residuen

Doordat de conditie per groep is toegewezen en niet aan individuele proefpersonen, kon voor beide modellen niet worden aangenomen dat de observaties volledig onafhankelijk van elkaar waren. In onderzoek van Stoevenbelt et al. (2025) is besloten deze multilevel-structuur niet te modelleren.

Lineariteit

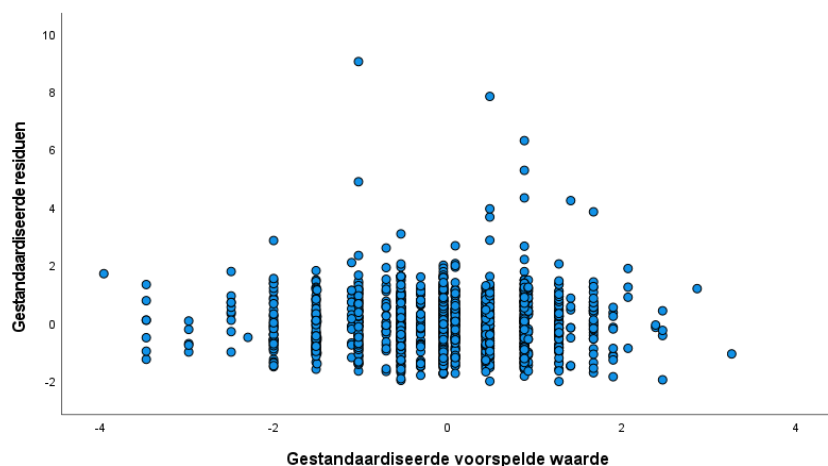
De resultaten zijn betrouwbaar wanneer er een lineaire relatie is tussen de afhankelijke en de onafhankelijke variabelen. Door middel van een spreidingsdiagram wordt gecheckt of deze aanname wordt geschonden of niet. Figuur 2 en 3 laten zien dat de residuen willekeurig verspreid lijken rond de nullijn en dat geen duidelijke kromme, S-vorm of patroon zichtbaar is. De assumptie van lineaire relaties wordt dus voor beide modellen niet geschonden.

Homoscedasticiteit

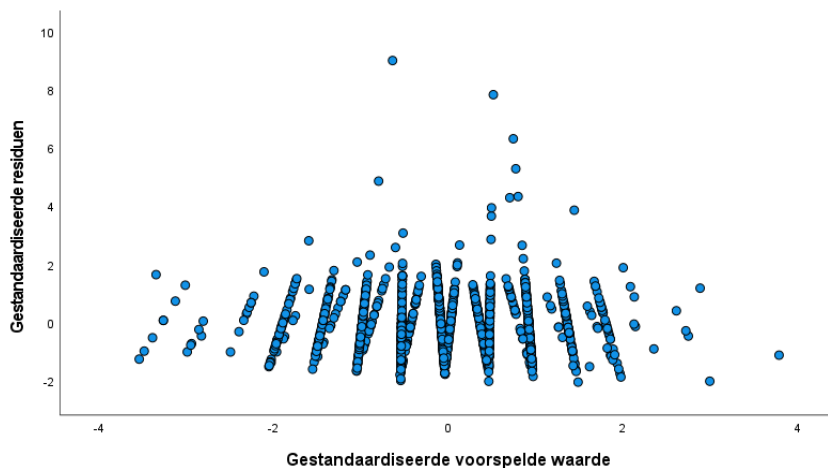
De spreiding van de residuen van het regressiemodel is redelijk constant over het bereik van voorspelde waarden (figuur 2). De verticale lijnen in de figuur zijn mogelijk te verklaren doordat er een regressiemodel is gebruikt met categorische predictoren en één continue variabele. Hierdoor maakt SPSS een beperkt aantal unieke voorspelde waarden, waardoor veel deelnemers dezelfde voorspelde waarde krijgen. Hoewel enkele uitbijters aanwezig zijn, is er geen structureel patroon dat duidt op heteroscedasticiteit. De assumptie homoscedasticiteit wordt dus niet geschonden. Figuur 3 toont dat de spreiding van de residuen bij het regressiemodel met interactieterm licht toeneemt, wat kan wijzen op enige mate van heteroscedasticiteit. Al is het patroon beperkt, de assumptie van homoscedasticiteit wordt daarmee niet volledig bevestigd.

Figuur 2

Spreadingsdiagram regressiemodel van gestandaardiseerde residuen tegen gestandaardiseerde voorspelde waarden

**Figuur 3**

Spreadingsdiagram regressiemodel met interactieterm van gestandaardiseerde residuen tegen gestandaardiseerde voorspelde waarden

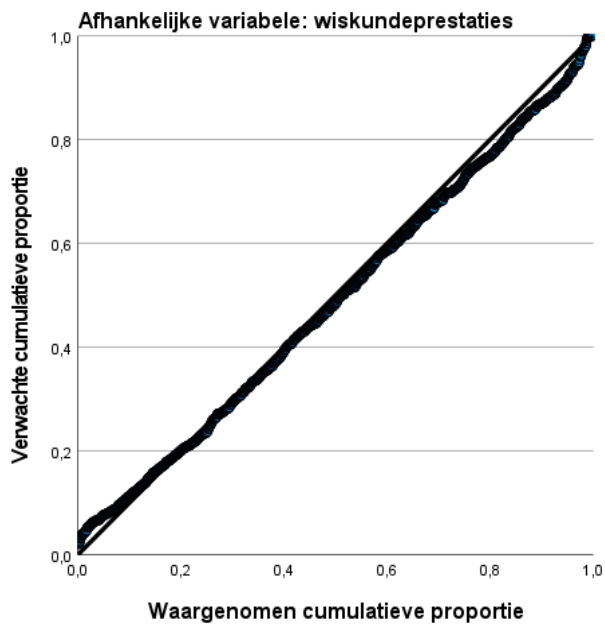


Normaliteit

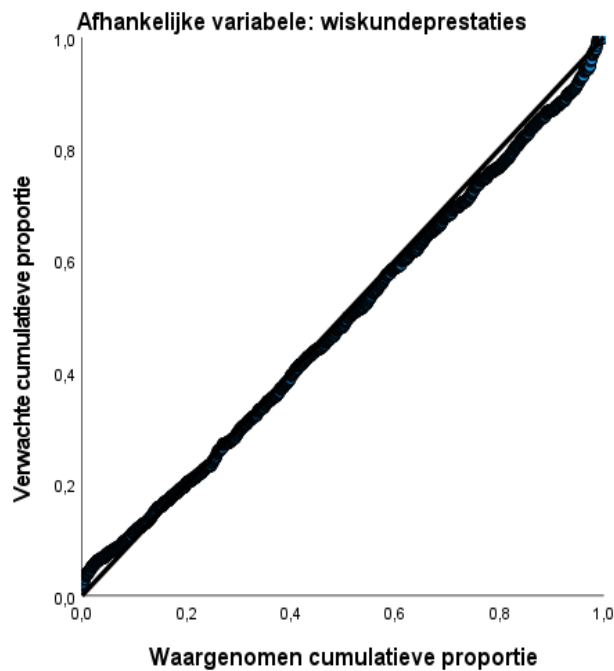
De normaliteit van de residuen werd gecontroleerd via een P-P plot. Zoals Figuur 3 en 4 laten zien, liggen de residuen dicht bij de diagonaal, wat wijst op een acceptabele normaliteit voor beide modellen.

Figuur 4

Normaliteitsplot van gestandaardiseerde residuen regressiemodel (P-P plot)

**Figuur 5**

Normaliteitsplot van gestandaardiseerde residuen regressiemodel met interactieterm (P-P plot)



Multicollineariteit

Tot slot wordt voldaan aan de assumptie van geen multicollineariteit. Multicollineariteit werd beoordeeld aan de hand van de Variance Inflation Factor (VIF), waarbij waarden < 5 als acceptabel werden beschouwd (Field, 2018). Tabel 3 laat zien dat de VIF-waarden van beide modellen onder de kritische waarde van 5 liggen, wat wijst op een aanvaardbare mate van multicollineariteit. De licht verhoogde VIF van het regressiemodel met de interactieterm is een gevolg van een hoge correlatie tussen de interactieterm en de hoofdeffecten. Dit is gebruikelijk aangezien de interactieterm een product is van deze hoofdeffecten.

Tabel 3

VIF-waarden per predictor in Model 1 en 2

Predictor	VIF (model 1)	VIF (model 2)
Stereotype bedreiging	1,00	1,00
Genderidentificatie	1,00	3,86
Interactieterm	-	3,86

Noot. Model 1 bevat alleen de hoofdeffecten; Model 2 bevat ook de interactieterm tussen stereotiepe bedreiging en genderidentificatie. VIF = Variance Inflation Factor.

Discussie

Het doel van dit onderzoek was om te onderzoeken in hoeverre genderidentificatie de relatie tussen stereotiepe bedreiging en wiskundeprestaties bij vrouwelijke studenten beïnvloedt. Met kennis vanuit eerdere studies werd verwacht dat vrouwen die zich sterker met hun gender identificeren, gevoeliger zouden zijn voor stereotiepe bedreiging, wat zou leiden tot lagere prestaties op een wiskundetoets. Deze verwachting is echter niet bevestigd door de resultaten.

De regressieanalyses lieten zien dat genderidentificatie een kleine maar significante negatieve voorspeller was van wiskundeprestaties. Dit betekent dat vrouwen die zich sterker identificeren met hun gender, gemiddeld iets lager op de wiskundetoets scoren. In tegenstelling tot de bevindingen van Spencer et al. (1999) werd in dit onderzoek geen statistisch significant effect gevonden van de experimentele conditie (stereotiepe bedreiging versus controle) op wiskundeprestaties. De effectgrootte ($\beta = .043$) was klein, wat suggereert dat het verschil tussen de condities in de praktijk beperkt is. Zoals Schmader (2002) suggereerde, zou stereotiepe bedreiging sterker zijn bij vrouwen met een hoge mate van genderidentificatie. In dit onderzoek werd die hypothese echter niet bevestigd. Hoewel de toevoeging van de interactieterm leidde tot een kleine, maar statistisch significante verbetering van de verklaarde variantie in het regressiemodel ($\Delta R^2 = .000$, $p = .024$), bleek de interactieterm zelf niet significant. Dit geeft aan dat het effect van stereotiepe bedreiging op wiskundeprestaties niet verschilt afhankelijk van de

mate van genderidentificatie. Met andere woorden: het maakte in deze studie niet uit hoe sterk iemand zich identificeert met haar gender, stereotiepe bedreiging heeft gemiddeld genomen geen effect op iedereen. Samengevat tonen de resultaten aan dat genderidentificatie op zichzelf een voorspellende factor is, maar dat deze factor de relatie tussen stereotiepe bedreiging en wiskundeprestaties niet modereert. Aangezien de steekproefomvang relatief groot was ($N = 1275$), is het mogelijk dat de statistische verbetering van het model is toe te schrijven aan het feit dat minimale variantieverschillen bij een grote steekproef sneller significant worden, zonder dat er sprake is van een betekenisvol effect in de praktijk.

Een mogelijke verklaring voor het uitblijven van een effect van stereotiepe bedreiging is dat de manipulatie te subtiel was. In dit onderzoek werd stereotiepe bedreiging opgewekt door middel van een audio-instructie voorafgaand aan de wiskundetaak, waarin impliciet werd verwezen naar het meten van cognitieve vaardigheden. Studies van onder andere Steele en Aronson (1995) lieten zien dat stereotiepe bedreiging sterker werd geactiveerd wanneer expliciet verwezen werd naar negatieve groep stereotypen of wanneer deelnemers zich bewust waren van mogelijke beoordeling op basis van hun identiteit. Het is daarom mogelijk dat niet alle deelnemers de situatie als stereotiep bedreigend ervaren hebben, waardoor het verwachte effect is uitgebleven. Het zou dus voor vervolgonderzoek aan te raden zijn om expliciet te verwijzen naar stereotiepe bedreiging. Ook zou toevoeging van een Impliciete Associatie test (IAT) mogelijk meer inzicht kunnen geven. Deze test meet hoe sterk individuen automatisch bepaalde stereotypen (zoals 'wiskunde = mannelijk') aan zichzelf koppelen of juist afwijzen. Eerder onderzoek van Nosek et al. (2002) liet zien dat vrouwen die onbewust sterker de associatie 'wiskunde = mannelijk' hadden, zich minder aangetrokken voelden tot wiskunde en minder geneigd waren om zich in dat domein verder te ontwikkelen. Het zou een aanwijzing kunnen zijn voor de mate waarin iemand gevoelig is voor stereotiepe bedreiging. Toekomstig onderzoek zou hiermee mogelijk beter kunnen voorspellen bij wie en wanneer stereotiepe bedreiging invloed heeft op prestaties.

Interessant is de bevinding dat genderidentificatie ook zonder stereotiepe bedreiging samenhangt met wiskundeprestaties. Dit suggereert dat vrouwen die zich sterker met hun gender identificeren, mogelijk gevoeliger zijn voor verwachtingen rondom "mannelijke domeinvakken". Eerder onderzoek liet zien dat een sterke genderidentificatie kan leiden tot verhoogde prestatiedruk of zelftwijfel in domeinen waarin het stereotiepe actief is (Schmader, 2002). Ook zonder de expliciete activatie van stereotiepe bedreiging kunnen onderliggende processen dus van invloed zijn op wiskundeprestaties. Dit benadrukt het belang van breder kijken naar sociale en psychologische factoren als het gaat over wiskundig functioneren.

Voor vervolgonderzoek is het belangrijk om te kijken naar een meer diverse gendergroep. In de huidige studie werd genderidentificatie gemeten als één dimensionale schaal bij cisgender vrouwelijke deelnemers, hierdoor blijft het effect van stereotiepe bedreiging op personen met een andere

genderidentiteit onduidelijk. Daarbij zou vervolgonderzoek kunnen overwegen om mogelijke andere factoren op te nemen, aangezien de onderzochte variabelen slechts in beperkte mate bijdragen aan verschillen in wiskundeprestaties. In eerder onderzoek wordt gesuggereerd dat de effecten van stereotiepe bedreiging sterker zijn bij personen die zich in hoge mate identificeren met het betreffende domein, zoals wiskunde (Steele, 1997). De domeinidentificatie werd niet direct gemeten, maar andere factoren zoals zelfvertrouwen in wiskunde (Pajares & Miller, 1994), faalangst (Ashcraft & Kirk, 2001) en wiskundige voorkennis (Geary, 2011) kunnen ook een belangrijke rol spelen. Deze factoren beïnvloeden in hoeverre iemand zich competent voelt binnen het domein en kunnen daarmee indirect bijdragen aan de gevoeligheid voor stereotiepe bedreiging. Door deze variabelen mee te nemen komen onderzoekers mogelijk tot een completer verklaringsmodel.

Een beperking van dit onderzoek is dat de assumptie van onafhankelijke residuen is geschonden. De data kende een multilevel-structuur, wat inhoudt dat de conditie niet individueel maar per groep is toegewezen. Dit vergoot de kans dat de residuen binnen dezelfde groep met elkaar samenhangen. Volgens Field (2018) en Hox et al. (2017) kan deze afhankelijkheid leiden tot een onderschatting van, in dit geval de standaardfouten voor het effect van genderidentificatie op wiskundeprestaties, waardoor de bijbehorende p -waarden te laag kunnen uitvallen. Dit vergroot de kans dat de effecten die statistisch significant lijken, in werkelijkheid niet significant zijn (Type I-fout). Toekomstig onderzoek zou daarom gebaat zijn bij het toepassen van een multilevel-analyse die deze groepsstructuur meeneemt.

Het blijft maatschappelijk relevant om de rol van genderstereotypen in het onderwijs te volgen, ondanks dat deze studie niet het verwachte effect opleverde. Vrouwen blijven ondervertegenwoordigd in STEM-richtingen en stereotiepe bedreiging kan bijdragen aan studie- of carrièrekeuzes. Onderzoek naar verschillende factoren helpt beter te begrijpen welke psychologische processen hier invloed op hebben en hoe onderwijsinstellingen hierop kunnen inspelen. Daarbij is deze studie relevant geweest voor de wetenschap. Het ontbreken van een interactie-effect, stuurt nieuwe onderzoekslijnen aan om wel antwoord te krijgen over wanneer en bij wie stereotiepe bedreiging effect heeft.

De verwachting dat stereotiepe bedreiging een negatiever effect zou hebben op wiskundeprestaties bij vrouwen die zich sterker met hun gender identificeren, werd in dit onderzoek dus niet bevestigd. Dit suggereert dat de relatie tussen genderidentiteit en wiskundeprestaties complexer is dan eerder werd aangenomen en dat dergelijke modererende effecten mogelijk afhangen van andere psychologische of contextuele factoren.

Referenties

- Ashcraft, M. H., & Kirk, E. P. (2001). *The relationships among working memory, math anxiety, and performance*. *Journal of Experimental Psychology: General*, 130(2), 224–237. <https://doi.org/10.1037/0096-3445.130.2.224>
- Beasley M. A., Fischer M. J. (2012). Why they leave: The impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Social Psychology of Education*, 15, 427–448. doi:10.1007/s11218-012-9185-3
- Brown, C.S., & Leaper, C. (2010). Latina and European American girls' experiences with academic sexism and their self-concepts in mathematics and science during adolescence. *Sex Roles*, 63, 860-870. doi:10.1007/s11199-010-9856-5
- Casad, B.J., Hale, P., & Wachs, FL (2017). Stereotype Threat Among Girls: Differences by Gender Identity and Math Education Context. *Psychology of Women Quarterly*, 41 (4), 513-529. <https://doi.org/10.1177/0361684317711412> (Origineel werk gepubliceerd in 2017)
- Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., & Kim, S. (2011). *Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? Social Psychological and Personality Science*, 2(6), 656–664. <https://doi.org/10.1177/1948550611405218>
- Cohen, J. (1992). Statistical Power Analysis. *Current Directions in Psychological Science*, 1(3), 98-101. <https://doi.org/10.1111/1467-8721.ep10768783> (Original work published 1992)
- Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences*, 47, 103–116. <https://doi.org/10.1016/j.lindif.2015.12.018>
- Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1999). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, 70(3), 670–685. <https://doi.org/10.1111/1467-8624.00048>
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5e editie). SAGE Publications
- Geary, D. C. (2011). *Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study*. *Developmental Psychology*, 47(6), 1539–1552. <https://doi.org/10.1037/a0025510>
- Hox, J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications*, Third Edition (3rd ed.). Routledge. <https://doi.org/10.4324/9781315650982>
- Jiang, S., Simpkins, S. D., & Eccles, J. S. (2020). Individuals' math and science motivation and their subsequent STEM choices and achievement in high school and college: A longitudinal study of gender and college generation status differences. *Developmental psychology*, 56(11), 2137–2151. <https://doi.org/10.1037/dev0001110>
- Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science*, 16, 175-179. doi:<https://doi.org/10.1111/j.0956-7976.2005.00799.x>

- Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British Journal of Educational Psychology*, 77, 323-338. <https://doi.org/10.1348/000709906X113662>
- Kolen, M. J., & Brennan, R. L. (2014). *Test Equating, Scaling, and Linking* (3 ed.). Springer-Verlag. <https://doi.org/10.1007/978-1-4939-0317-7>
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, 18, 302-318. <https://doi.org/10.1177/014616729218300>
- Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, 93(6), 1314–1334. <https://doi.org/10.1037/a0012702>
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, 83(1), 44–59. <https://doi.org/10.1037/0022-3514.83.1.44>
- Pajares, F., & Miller, M. D. (1994). *Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis*. *Journal of Educational Psychology*, 86(2), 193–203. <https://doi.org/10.1037/0022-0663.86.2.193>
- Schmader T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, 38, 194–201. doi:10.1006/jesp.2001.1500
- Selimbegovic, L., Karabegovic, M., Blažev, M., & Burušić, J. (2019). The Independent Contributions of Gender Stereotypes and Gender Identification in Predicting Primary School Pupils' Expectancies of Success in STEM Fields. *Psychology in the Schools*, 56(10), 1614–1632. <https://doi.org/10.1002/pits.22296>
- Shapiro J. R., Neuberg S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, 11, 107–130. doi:10.1177/1088868306294790
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35(1), 4–28. <https://doi.org/10.1006/jesp.1998.1373>
- Steele, C. M., & Aronson, J. (1995). *Stereotype threat and the intellectual test performance of African Americans*. *Journal of Personality and Social Psychology*, 69(5), 797–811. <https://doi.org/10.1037/0022-3514.69.5.797>
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613-629. doi:10.1037/0003-066X.52.6.613
- Steele, C. M., & Aronson, J. (1998). Stereotype threat and the test performance of academically successful African Americans. In C. Jencks & M. Phillips (Eds.), *The Black–White test score gap* (pp. 401– 427). Washington, DC: Brookings.
- Stoevenbelt, A. H., Wicherts, J. M., Flore, P. C., Phillips, L. A. T., Pietschnig, J., Verschuere, B., Voracek, M., & Schwabe, I. (2022). Are Speeded Tests Unfair? Modeling the Impact of Time Limits on the Gender Gap in

Mathematics. Educational and Psychological Measurement, 83(4), 684-709.

<https://doi.org/10.1177/00131644221111076> (Original work published 2023)

Tajfel, H. (1981). Human groups and social categories: Studies in social psychology. Cambridge, UK: Cambridge Univ. Press.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), The psychology of intergroup relations (pp. 7–24). Chicago: Nelson–Hall.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). Rediscovering the social group: A self-categorization. Oxford, UK: Blackwell.

Woodcock, A., Hernandez, P. R., & Schultz, P. W. (2016). Diversifying Science: Intervention Programs Moderate the Effect of Stereotype Threat on Motivation and Career Choice. *Social psychological and personality science*, 7(2), 184–192. <https://doi.org/10.1177/1948550615608401>