



# **Konferenz- beiträge 2021**

16. Jenaer  
Empiriepraktikumskongress



# Danksagung

Es ist wieder soweit: Heute, am 16. Juli 2021, findet bereits zum 16. Mal der Jenaer Empiriepraktikums-(EmPra)-Kongress statt. Ein zweites Mal in Folge präsentieren die EmPra-Gruppen dabei ihre Ergebnisse des zurückliegenden Studienjahres aufgrund der Corona-Pandemie nicht als Poster im Foyer des Campus Carl-Zeiss-Straße, sondern als Kurzvortrag online.

Vielmehr als diese spezielle Umsetzung des Kongresses, müssen jedoch die Umstände, unter denen die präsentierten Arbeiten zustande gekommen sind, hervorgehoben werden. Jedes Jahr betonen wir als Organisatoren des EmPra-Kongresses an dieser Stelle, dass die Studierenden mit Fug und Recht stolz sein können auf ihre Leistungen. Jedes Jahr hat es Probleme und Hindernisse gegeben, die überwunden werden mussten, da Forschung nun mal selten reibungslos verläuft und es noch dazu für die meisten Teilnehmenden das erste Mal in ihrer akademischen Laufbahn ist, dass überhaupt empirisch gearbeitet wird. Jedes Jahr entstehen trotz verschiedenster Rückschläge und Pannen spannende und innovative Forschungsarbeiten von überragender Qualität.

Während all dies auch in diesem Jahr zutrifft, kamen durch die verschiedenen Maßnahmen zur Einschränkung der Ausbreitung des Virus SARS-Cov-2 während des Sommersemesters große zusätzliche Herausforderungen dazu. Und trotz allem sind auch in diesem Jahr wieder so viele spannende, kreative und relevante Arbeiten entstanden, dass diese Konferenz sich vor der wissenschaftlichen Öffentlichkeit sehen lassen kann. Deswegen möchten wir allen EmPra-Gruppen ganz besonders mit auf den Weg geben: Blicken Sie auf das zurückliegende Studienjahr und Ihre Leistungen zurück und seien Sie stolz auf sich angesichts dessen, was Sie erreicht haben!

Eines hat sich auch unter den besonderen Umständen dieses Kongresses nicht geändert: ein solcher Kongress kann nur aufgrund der Unterstützung und des Engagements zahlreicher Personen zustande kommen. Ein besonderer Dank gilt an dieser Stelle stets den Mitgliedern der Professorenschaft, des akademischen Mittelbaus und der Studierendenschaft, welche sich als Jury die Zeit nehmen, die einzelnen Arbeiten mit viel Sorgfalt und Eifer zu begutachten. Darüber hinaus gilt unser Dank allen teilnehmenden Gruppen und ihren Betreuern und Betreuerinnen, die durch die bunte Themenvielfalt, die kreativen Forschungsideen und nicht zuletzt durch die Freude am Austausch und dem Generieren neuer Ideen den EmPra-Kongress prägen und unvergesslich machen. Für die Bereitstellung der Preisgelder und der Gutscheine zur Prämierung der besten Arbeiten bedanken wir uns beim Institut für Psychologie und Herrn Christopher Schnell von der Thalia Universitätsbuchhandlung GmbH. Des Weiteren danken wir den Studierenden aus der Fachschaft Psychologie für die Hilfe bei der Suche studentischer Jury-Mitgliedern.

Für die Abteilung Allgemeine Psychologie II,  
Adrian Jusepeitis und Merlin Urban.

# Inhaltsverzeichnis

Kurzberichte	6 - 45
Gruppe 1	6
<hr/>	
<i>Nur gut gemeint!? – Sexismus in unseren Komplimenten</i> Sophie Asmus, Johanna Biele, Luise Frommelt, Lydia Kahle, Jana Leonhardt	
Gruppe 2	8
<hr/>	
<i>Bitte steh mir bei! – Bystander-Effekt bei Online-Mobbing</i> Marie Hoffmann, Sebastian Goerke, Katharina Bronn, Celine Henning, Victoria Gerhart	
Gruppe 3	10
<hr/>	
<i>Propositionale Maße im Duell - Effekte von Leistungsfeedback auf indirekt und direkt gemessenen Selbstwert</i> Till Biedermann, Arabella Kolmitz, Mathilda Kraft, Viktoria Partschefeld, Marina Petrovic	
Gruppe 4	12
<hr/>	
<i>Komponenten von Bestrafung und deren Wirkung - Wie nehmen wir Täter*innen wahr?</i> Aljoscha Labonte, Flavia Laqua, Selina Marten, Jacob Schuster	
Gruppe 5	14
<hr/>	
<i>Wer ist der Täter? Wie Attraktivität die Auswahl in Lineups beeinflusst</i> Antonia Grüner, Anja Büttner, Nicole Keller, Christin Strutz	
Gruppe 6	16
<hr/>	
<i>Wenn Stimmen hervorstechen - Effekte von Karikierung und selektiver Aufmerksamkeit auf die Wahrnehmung von Gesangsgruppen</i> Lisa Gärtner, Sinha Kirchhoff, Marie Limmer, Vivien Schaffer	
Gruppe 7	18
<hr/>	
<i>Gesellschaftliche Normen – Die Mutter aller Vorurteile?</i> Nadin Merten, A.V., Laura Müller, Leonie Art, Philipp Bornschein	
Gruppe 8	20
<hr/>	
<i>Traumurlabsziel vs. Mordgelüste: Wie verändern sich automatische Gedächtnisprozesse des Lügens bei intimen Themen im Vergleich zu oberflächlichen Themen?</i> Melanie Bomm, Sissy Höbelt, Selina Hußner, Rebecca Matthias, Lea Ulverich	
Gruppe 9	22
<hr/>	
<i>Effects of Personality on Perceived Stress during the COVID-19 Pandemic</i> Benjamin Eggs, Friederike Witt, Philipp Wenzel	
Gruppe 10	24
<hr/>	
<i>Effects of punishment on offenders' self-perception</i> Jessica Heinz, Ann-Kathrin Kohl, Freyja Krull, Paula Molina Sander, Julia Schön	
Gruppe 11	26
<hr/>	
<i>Observational SR-Binding – „Mein, Dein oder Unser?“ – wie beeinflusst unser Mindset das Beobachtungslernen</i> Johanna Dorra, Joanna Hein, Aurelia Holfeld, Julian Huber, Greta Sippel, Teresa Wiesner	

Gruppe 12	28
<i>Let's face it! Wie die unidirektionale Manipulation konfiguraler Information das Lernen von Gesichtern beeinträchtigt</i>	
Svenja Andrade Breslauer, Maria Arabadzhieva, Lucy Braxmeier, Michelle Gutteck, Verena Kolb, Eva Reiter, Saida Saibou, & Lisa-Marie Schröder	
Gruppe 13	30
<i>Die Schwierigkeit und die Vorhersagekraft im IAT</i>	
Luisa Drechsel, Natalie Hülßner, Annelie Rodestock, Lena Wetzel, Elena Zuschlag	
Gruppe 14	32
<i>Die magische evaluative Konditionierung - Produzieren asymmetrische Relationspaare asymmetrische EC-Effekte?</i>	
Laura Dienemann, Lara Fangradt, Hannah Hoffmann, Viktoria Kämpfe	
Gruppe 15	34
<i>Social Networks and Stress During the Covid-19 Pandemic</i>	
Laura Maas, Moritz Ratzke, Hannes Thomas	
Gruppe 16	36
<i>Sind Polizisten die besseren Augenzeugen? Der Einfluss des Waffen Fokus Effekts auf die Zuverlässigkeit von Augenzeugenaussagen</i>	
Pauline Endler, Anne Fischer, Maximilian Francke, Lisa Hensel, Sümeyye Kocak	
Gruppe 17	38
<i>Implizites Beobachtungslernen unter Konformitätseinfluss</i>	
Niclas Dittmann, Nicole Klubkova, Marco Klug, Karl Schramm	
Gruppe 18	40
<i>„Singen auf der Cocktail-Party“: Effekte von Stimmenklangfarbe und -karikierung auf die auditive Szenenanalyse</i>	
Anna Knospe, Natalie Freiboth, Jonas Dragendorf	
Gruppe 19	42
<i>Anderthalb Meter Gesellschaft: Wie verändert COVID-19 unsere Gesichtswahrnehmung?</i>	
Erik Eyring, Dorothea Eichentopf, Lydia Krause, J.S., Jana Bodenbinder	
Gruppe 20	44
<i>Hat die unterschiedliche Symmetrie relationaler Wortpaare einen Einfluss auf EC-Effekte?</i>	
Friederike Brabetz, Jan Goebel, Lydia Hoffmeier, Clara Ladig	
Programm	46

# Nur gut gemeint!? – Sexismus in unseren Komplimenten

Sophie Asmus, Johanna Biele, Luise Frommelt, Lydia Kahle, Jana Leonhardt

Leitung: Prof. Dr. Thomas Kessler

## 1. Einleitung

„Das ist doch kein Sexismus, das ist ein Kompliment...“ Aussagen über wohlwollende Stereotype, wie etwa die „besondere Einfühlsamkeit“ von Frauen, begegnen uns fortwährend im Alltag. Positive Stereotypisierung auf der Wärme-Dimension werden als sexistisch wahrgenommen und kommen bei Frauen nicht gut an, denn es wird weniger Kompetenz unterstellt (Siy & Cheryan, 2016). Allerdings beschränken sich diese Studien auf positive Stereotype auf der Wärmedimension. Was passiert aber, wenn Frauen auf der Kompetenzdimension positiv stereotypisiert werden? Dem wollten wir in unserer Studie nachgehen, indem wir Frauen nicht nur mit positiven Stereotypen auf der Wärmedimension, sondern auch mit positiven Stereotypen auf der Kompetenzdimension (z.B. "Frauen sind so intelligent") konfrontieren. Kommt ein solches Stereotyp besser an? Folgende zwei konkurrierenden Hypothesen sind naheliegend: Nach Fiskes Stereotype Content Model (2002) sollte die Stereotypisierung auf der Wärmedimension dazu führen, dass Frauen sich in ihrer Kompetenz herabgesetzt fühlen. Unsere 1. Hypothese besagt demnach, dass Frauen sich nur bei positiver Stereotypisierung auf der Wärmedimension negativ stereotypisiert, depersonalisiert und sexistisch behandelt fühlen. Stereotypisierung auf der Kompetenzdimension führt nicht zu dieser Einschätzung und wird als besser als die Kontrollbedingung ohne Stereotypisierung bewertet. Die konkurrierende 2. Hypothese wäre ebenso plausibel: Jede Art von positiver Stereotypisierung (egal ob auf der Wärme- oder Kompetenzdimension) führt dazu, dass Frauen sich negativ stereotypisiert, depersonalisiert und sexistisch behandelt fühlen, da sie sich auf ihre Gruppenzugehörigkeit reduziert sehen. In der Wärme- und Kompetenzbedingung fallen die Werte signifikant höher aus als in der Kontrollbedingung ohne Stereotypisierung (Hypothese 2).

## 2. Methode

Unsere Studie haben wir online über die Plattform sosci-survey über einen Erhebungszeitraum von drei Wochen durchgeführt. Insgesamt nahmen 137 weibliche Versuchspersonen an der Studie teil. Das Durchschnittsalter der Teilnehmerinnen liegt bei 38.9 (SD = 20.21 [18 - 90]). Es wurde ein between-subjects Design verwendet, welches an die Studie von Siy und Cheryan (2016) angelehnt ist, wobei wir jedoch eine zusätzliche Bedingung („Kompetenz“) aufgenommen haben. Zufällig wurden die Versuchspersonen einer von drei Bedingungen (Wärme-Bedingung, Kompetenz-Bedingung oder Kontrollbedingung) zugeordnet, in welcher sie dann eine von vier Situationen geschildert bekamen, innerhalb derer sie (in den Experimentalbedingungen aufgrund ihres Geschlechtes) für eine Aufgabe ausgewählt wurden. In der Wärme-Bedingung wurde ihr Geschlecht mit einem aus fünf positiven, warmen Stereotypen (herzlich, freundlich, einfühlsam, rücksichtsvoll, gutmütig) in Verbindung gebracht („Frauen sind sehr herzlich.“), in der Kompetenz-Bedingung mit einem von fünf Adjektiven, welche Kompetenz ausdrücken (professionell, intelligent, kompetent, fähig, selbstbewusst). In der Kontrollbedingung wird die Versuchsperson ohne Begründung für die Aufgabe ausgewählt. Im Anschluss wurden den Teilnehmerinnen vier Fragen dazu gestellt, inwieweit sie sich durch das dargebotene Szenario auf ihr Geschlecht reduziert fühlen. Diese vier Items lassen sich zur Skala „Grad der empfundenen Depersonalisierung“ zusammenfassen ( $\alpha = .77$ ). Weiterhin wurde erfasst, ob sie glauben, dass man ihnen auch 27 andere Eigenschaften zuschreibt. Hierbei wurden 7 negative weibliche Stereotype mit 20 Füllitems gemischt dargeboten. Diese Items wurden zur Skala „Negative Stereotype“ zusammengefasst ( $\alpha = .86$ ). Die Antworten wurden jeweils auf einer 7-stelligen Skala erfasst. Zusätzlich wurden die Probandinnen gefragt, für wie sexistisch sie die Person halten und welches Geschlecht diese hat (1=wenig sexistisch, 7=sehr sexistisch).

### 3. Ergebnisse

*Tabelle 1: Mittelwerte in den verschiedenen Bedingungen*

	Wärme <i>M(SD)</i>	Kompetenz <i>M(SD)</i>	Kontrolle <i>M(SD)</i>
Sexistisch	4.46(1.85)	3.76(2.1)	2.96(1.71)
Depersonalisierung	4.97(1.25)	4.58(1.25)	3.59(1.27)
Negative Stereotypisierung	3.71(1.49)	3.22(1.22)	2.92(1.02)

Tab. 1: Die Tabelle zeigt die Mittelwerte der drei Skalen in den drei Bedingungen.

Um die Unterschiede zwischen den Bedingungen zu untersuchen, führten wir eine Kontrastanalyse durch, die die beiden Stereotypen Bedingungen mit der Kontrollgruppe verglichen. Der Unterschied zwischen den beiden Experimentalbedingungen und der Kontrollgruppe ist über alle drei Skalen hinweg signifikant. Auf der Skala „sexistisch“ mit  $t(134) = -3.37$ ,  $p < .001$ . Auf der Skala „Depersonalisierung“ mit  $t(133) = -5.20$ ,  $p < .001$  und auf der Skala „Negative Stereotypisierung“ mit  $t(134) = -2.45$ ,  $p = .016$ . Der Unterschied zwischen den beiden Experimentalbedingungen „Wärme“ und „Kompetenz“ wird auf keiner Skala signifikant. Auf der Skala „sexistisch“ mit  $t(134) = -1.77$ ,  $p = .0796$ . Auf der Skala „Depersonalisierung“ mit  $t(133) = -1.48$ ,  $p = .142$  und auf der Skala „Negative Stereotypisierung“ mit  $t(134) = -1.88$ ,  $p = .063$ . In der Bedingung „Wärme“ gaben ca. 76% der Probandinnen an, dass sie vermuten, dass die Person in dem Szenario männlich sei, in der Bedingung „Kompetenz“ waren es ca. 69%. In der Kontrollbedingung waren es ca. 50%.

### 4. Diskussion

Frauen, die im fiktiven Szenario von einer Person positiv stereotypisiert wurden, fühlten sich depersonalisiert, beschrieben die Person eher als sexistisch und glaubten eher, auch negative weibliche Stereotype zugeschrieben zu bekommen, als in der Kontrollbedingung. Das war sowohl bei positiven Stereotypen auf der Wärme-, als auch auf der Kompetenzdimension der Fall. Daher kann die 2. Hypothese angenommen und die 1. verworfen werden: Jede Art von Stereotypisierung, egal ob auf der Wärme- oder Kompetenzdimension, wird als negativ empfunden. Insgesamt fielen die Werte für die Wärmebedingung etwas höher aus, als in der Kompetenzbedingung. Auf Wärme stereotypisiert zu werden, kam in der Stichprobe also etwas schlechter an, als auf Kompetenz. Die Unterschiede wurden jedoch nicht signifikant. Die Ergebnisse der Frage nach dem Geschlecht des Targets in den Szenarien weisen darauf hin, dass Frauen sich eher von Männern stereotypisiert fühlen. Stereotype haben also Intergruppencharakter. Eine Frage, die für zukünftige Forschung offenbleibt, ist inwieweit diese Ergebnisse auch im Alltag eine Rolle spielen. In unserer Studie wurden die Personen zum Beispiel direkt gefragt, ob sie das Target "sexistisch" finden. Das löst möglicherweise Gedankengänge aus, die im Alltag nicht aufkommen würden, wenn man mit einem positiven Stereotyp auf der Kompetenzdimension konfrontiert wird. Außerdem sollte untersucht werden, ob sich diese Stereotypisierungseffekte auch auf Männer übertragen lassen.

### 5. Literatur

- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878–902. <https://doi.org/10.1037/0022-3514.82.6.878>
- Siy, J. O., & Cheryan, S. (2016). Prejudice Masquerading as Praise: The Negative Echo of Positive Stereotypes. *Personality & social psychology bulletin*, 42(7), 941–954. <https://doi.org/10.1177/0146167216649605>

## Bitte steh mir bei! – Bystander-Effekt bei Online-Mobbing

Marie Hoffmann, Sebastian Goerke, Katharina Bronn,  
Celine Henning, Victoria Gerhart

Leitung: M.Sc. Anne Grigutsch

### 1. Einleitung

Gibt es irgendwo einen Notfall mit mehreren Zeugen, so kann man häufig beobachten, dass es keinen Konsens darüber gibt, wer zuerst helfen oder eingreifen soll. In der Wissenschaft ist dieses Phänomen als Bystander-Effekt bekannt: Die Wahrscheinlichkeit eines Opfers in Not- oder Unrechtssituation Hilfe zu erhalten nimmt paradoxer Weise ab, je mehr Personen (und damit potenzielle Helfer) anwesend sind<sup>1)</sup>. Zahlreiche Studien zu diesem Thema haben Erklärungsansätze wie pluralistische Ignoranz, Verantwortungsdiffusion und Bewertungsangst untersucht und sich traditionell mit Merkmalen einer Situation befasst, die diese begünstigen<sup>1)</sup>. Vor dem Hintergrund der wachsenden Bedeutung sozialer Medien und der nahezu allgegenwärtigen Präsenz (mitlesender) Bystander im digitalen Raum hat das Phänomen im online Kontext, besonders für die Problematik des Cybermobbing, erneut an Relevanz gewonnen. Dabei rücken zunehmend Fragen nach dem Einfluss von *Personenmerkmalen* auf Bystander-Verhalten in den Fokus der Forschung. So fand sich, dass Empathie<sup>2),3),4)</sup>, Extraversion<sup>2)</sup> und Selbstwirksamkeit<sup>3)</sup> (SWE) die Wahrscheinlichkeit erhöhen unterstützend in einer beobachteten Cybermobbing-Situation zu intervenieren.

Die folgende Untersuchung soll diese Befunde sowie den Bystander-Effekt als solchen im online Kontext replizieren. Darüber hinaus sollen explorativ als weitere Personenvariablen das Anschluss- und Machtmotiv sowie die Vorhersagekraft der Prädiktoren für konkrete Interventionsarten betrachtet werden.

*Hypothesen.* Es werden ein positiver Einfluss von Empathie, Extraversion und SWE, sowie ein negativer Einfluss der Bystander-Bedingung auf Interventionsbereitschaft erwartet.

### 2. Methode

*Stichprobe.* Die Stichprobe umfasste 63 Versuchspersonen (53 Frauen,  $M_{Alter} = 23.1$ ,  $SD_{Alter} = 3.6$ ). Diese wurden zufällig einer von zwei experimentellen Bedingungen zugeordnet: Bystander-Bedingung (N=27), noBystander-Bedingung (N=36).

*Material, Vorgehen.* Die Datenerhebung erfolgte über das Umfragetool "SoSci Survey". In einer 25-minütigen Befragung wurden zunächst die Personenvariablen Anschluss- und Machtmotiv (MMG), Extraversion (BFI-10), Empathie (SPF-IRI) und Allgemeine Selbstwirksamkeitserwartung (SWE) gemessen. Danach wurde den Versuchspersonen ein Auszug einer fiktiven Facebook Gruppen-Konversation zum Thema Sexismus gezeigt, worin eine Userin zunehmend von zwei anderen Usern gemobbt wurde. Je nach Bedingung erhielten die Versuchspersonen unterschiedliche Informationen über die Anzahl der Bystander – wie viele Personen also neben ihnen und den Postenden Teil der Gruppe waren (Bystander-Bedingung: 17, noBystander-Bedingung: keine). Anschließend wurden die Versuchspersonen gebeten anzugeben, ob sie selbst an der Konversation teilnehmen würden, ggf. einen entsprechenden Kommentar zu erfassen und ihre Entscheidung zu begründen. Ebenso wurde erfasst, mit welcher der anderen Personen sie sich am meisten identifizierten und ob die konkrete Anzahl der Gruppen-Mitglieder erinnert wurde.

### 3. Ergebnisse

Versuchspersonenkommentare wurden anhand inhaltlicher Kriterien dahingehend kodiert, ob sie eine (helfende) Intervention darstellten (binär: ja/nein) und welcher von 4 Kategorien (supportive-offline/supportive-online/confrontational/mediate) diese zuzuordnen war.



Zur Untersuchung des Einflusses der erhobenen Prädiktoren und der Experimentalbedingung auf die Wahrscheinlichkeit einer Intervention wurde eine multiple logistische Regression berechnet. Hierfür wurden in vier Schritten zunächst die aus der Literatur ermittelten Prädiktoren (Empathie, SWE, Extraversion), danach die Experimentalbedingung, anschließend Anschluss- und Machtmotiv-Facetten des MMG und zuletzt alle zweifach-Interaktionen mit dem Bedingungsfaktor aufgenommen und jeweils die Modellverbesserung mittels Likelihood-Ratio-Test evaluiert. Dabei wurde in den letzten beiden Schritten keine Modellverbesserung erzielt ( $ps > .559$ ) und Motivfacetten und Interaktionsterme daher als Prädiktoren ausgeschlossen. Von den übrigen Variablen erwiesen sich alle als signifikante Prädiktoren der Interventionswahrscheinlichkeit (s. Tabelle 1). Hypothesenkonform zeigte sich ein positiver Einfluss von Empathie. Die Effekte von Selbstwirksamkeitserwartung, Extraversion und der Experimentalbedingung fielen jedoch entgegen der Hypothesen aus. Darüber hinaus hatten die einzelnen Prädiktoren keine Vorhersagekraft für die konkrete Interventionsart.

	$\beta$	SE	p	Odds Ratio ( $e^{\beta}$ )
(Intercept)	0.365	0.466	.433	1.441
Empathie <sup>a</sup>	0.202	0.066	.002**	1.224
Extraversion <sup>a</sup>	-1.534	0.765	.045*	0.216
SWE <sup>a</sup>	-0.369	0.131	.005**	0.692
Bedingung <sup>b</sup>	1.409	0.706	.046*	4.093

**Tabelle 1:** AV: Intervention (0=nein, 1=ja);  $R^2_{\text{Nagelkerke}} = .411$ ,  $\log\text{Lik} = -28.42659$

<sup>a</sup>mittelwerts-zentriert; <sup>b</sup>dummy-codiert (0 = noBystander-Bedingung, 1 = Bystander-Bedingung)

#### 4. Diskussion

Wie zahlreiche Vorstudien<sup>2),3),4)</sup> fanden wir einen positiven Einfluss von Empathie auf die Wahrscheinlichkeit helfender Interventionen bei Beobachtung von Cybermobbing, was die Robustheit dieses Befundes unterstreicht. Entgegen unserer Erwartungen zeigten Extraversion und Selbstwirksamkeit einen negativen Einfluss auf die Interventionswahrscheinlichkeit. Ebenfalls überraschend war die Beobachtung eines „reversed Bystander-Effects“. Die Wahrscheinlichkeit einer helfenden Intervention war also in der Bedingung ohne weitere Bystander *geringer* als in der Bedingung mit Bystandern. Mögliche Gründe hierfür liegen im speziellen online-Kontext der Studie, welcher zugrundeliegenden Prozessen<sup>1)</sup> des Bystander-Effekts entgegengewirkt haben könnte. Zum einen wurde die Konversation im Rahmen einer geschlossenen Facebook-Gruppe präsentiert, was das Ausmaß der Öffentlichkeit und damit verbundener Verantwortungsdiffusion, pluralistischer Ignoranz und Bewertungsangst<sup>2)</sup> reduziert haben könnte. Letztere wird möglicherweise durch den Abstand zu anderen Personen in online Diskursen zusätzlich gemindert. Daneben könnte auch das spezielle Mobbing-Thema (Sexismus) eine Rolle beim Zustandekommen der unerwarteten Befunde gespielt haben. So dürfte dieses Thema für die vorrangig weibliche Stichprobe hohe persönliche Relevanz haben und dadurch zu Selbstkategorisierung und Identifikation mit dem (ebenfalls weiblichen) Opfer geführt haben – Prozesse, die dem klassischen Bystander-Effekt entgegenwirken<sup>1)</sup>. Nicht signifikante Befunde zur Rolle der Motivdisposition und Vorhersage konkreter Interventionsarten müssen vor dem Hintergrund geringer Power mit Vorsicht interpretiert werden.

#### 5. Literatur

- 1) Urschler, D. F., Fischer, J., Kastenmüller, A., & Fischer, P. (2015). Bystander Effect. *Psychology*, 29, 203-206.
- 2) Freis, S. D., & Gurung, R. A. R. (2013). A Facebook analysis of helping behavior in online bullying. *Psychology of Popular Media Culture*, 2(1), 11–19. <https://doi.org/10.1037/a0030239>
- 3) Pöyhönen, V., Juvonen, J., & Salmivalli, C. (2010). What does it take to stand up for the victim of bullying? the interplay between personal and social factors. *Merrill-Palmer Quarterly*, 56(2), 143–163.
- 4) Shultz, E., Heilman, R., & Hart, K. J. (2014). Cyber-bullying: An exploration of bystander behavior and motivation. *Cyberpsychology*, 8(4). <https://doi.org/10.5817/CP2014-4-3>

# **Propositionale Maße im Duell - Effekte von Leistungsfeedback auf indirekt und direkt gemessenen Selbstwert**

**Till Biedermann, Arabella Kolmitz, Mathilda Kraft, Viktoria Partschefeld, Marina Petrovic**

Leitung: M.Sc. Adrian Jusepeitis

## 1. Einleitung

Indirekte Selbstwertmaße wollen jene Facetten des Selbstwertes messen, die die Personen selbst nicht berichten können oder wollen (Buhrmester et al., 2009). Hierfür wurde bisher der Implizite Assoziations-Test (IAT) als beliebtes Maß eingesetzt. Ein wesentliches Problem ist hierbei, dass assoziative Maße wie der IAT blind gegenüber den Beziehungen zwischen Konzepten (z.B. ICH - GUT) sind. So beeinflussen auch irrelevante Propositionen wie „Ich will gut sein“ den Selbstwert-IAT. Beliefs wie „Ich bin gut“ können gezielter direkt in Fragebögen und kürzlich publizierten, propositionalen, indirekten Maßen wie dem „Propositional Evaluation Paradigm“ (PEP; Müller & Rothermund, 2019) abgebildet werden. Wir haben uns die Frage gestellt, inwieweit direkte und indirekte propositionalen Maße unterschiedliche Erkenntnisse bezüglich des Selbstwertes erlauben. Bisherige Studien zeigten, dass indirekte, propositionale Verfahren eine inkrementelle Varianzaufklärung für relevante Kriterien bieten (Dentale et al., 2020) und spontanes Verhalten besser vorhersagen als direkte Maße (Müller & Rothermund, 2019). Diese Befunde konnten wir in unserer ersten Studie bestätigen, in der das PEP die Varianz in depressiven Symptomen über die Ergebnisse eines Selbstwertfragebogens hinaus aufklären konnte. In unserer zweiten Studie wollen wir die Unterschiede zwischen diesen beiden Maßen durch eine Manipulation weiter herausarbeiten. Wir stellen die Hypothese an, dass durch einmaliges selbstwertrelevantes Leistungsfeedback nur die indirekte, nicht aber die direkte Selbstwertmessung manipuliert wird. Wir vermuten, dass die PEP-Effekte nach positivem Feedback im Durchschnitt einen höheren Selbstwert anzeigen als nach neutralem. Unsere Hypothese basiert auf dem Befund, dass direkte Messungen des Selbstwertes zeitlich stabiler als indirekte sind (Dentale et al., 2016). Wir vermuten, dies liegt daran, dass bei direkten Messungen auf langjährige Einstellungen zurückgegriffen wird, während bei indirekter Messung aktuell saliente Informationen, wie etwa das soziale Feedback, entscheidend sind.

## 2. Methode

Die Studie fand als computerbasierte Online-Studie statt. Von anfangs 106 Versuchspersonen wurden zwei ausgeschlossen, deren Muttersprache nicht Deutsch war und zwei, deren PEP Daten über 20% Fehler oder Ausreißer enthielten. Das mittlere Alter der finalen Stichprobe (N = 102, davon 64 weiblich, 37 männlich, 1 divers) betrug 25.64 Jahre (SD = 8.37). 67 Personen waren Psychologiestudierende. Anfangs absolvierten die Probanden den Reading Mind In The Eyes-Test, welcher zur Messung der Empathiefähigkeit eingesetzt wird. Im Anschluss erhielten sie ein zufälliges schriftliches und auditives Feedback zu ihrer Leistung im Test (Experimentalgruppe: überdurchschnittliche Leistung, Kontrollgruppe: durchschnittliche Leistung). Dann absolvierten die Probanden das PEP zur indirekten Messung ihres Selbstwerts. Dieses Messinstrument bestand aus insgesamt 120 Durchgängen in vier Blöcken. In jedem Durchgang wurde ein Satz (verkürzte Items aus der Rosenberg Self-Esteem Scale) Wort für Wort präsentiert, gefolgt von einer Eingabeaufforderung: Die Versuchspersonen sollten ihre Maus von einer blauen Start-Box in der unteren Mitte des Bildschirms entweder zu „JA“ oder „NEIN“ in der linken oder rechten oberen Ecke des Bildschirms bewegen. Die Probanden sollten den präsentierten Satz ignorieren und auf das danach gezeigte „JA“ oder „NEIN“ ihre Maus zum entsprechenden Feld bewegen. Durchgänge, in denen ein Satz gezeigt wurde, der auf einen positiven (negativen) Selbstwert hinweist und der Aufforderung, mit „JA“ („NEIN“) zu reagieren, bezeichnen wir als kongruent und Durchgänge, in

denen ein Satz gezeigt wurde, der auf einen positiven (negativen) Selbstwert hinweist und der Aufforderung, mit „NEIN“ („JA“) zu reagieren, als inkongruent. Als indirektes Maß für die Ausprägung des Selbstwerts verwendeten wir die Differenz der mittleren Reaktionszeiten zwischen kongruenten und inkongruenten Durchgängen (größere Reaktionszeiteffekte = höherer Selbstwert). Nach dem PEP absolvierten die Probanden die RSES. Zum Schluss wurden sie über die Manipulation und den Studieninhalt aufgeklärt. Diese Prozedur und die folgenden Analysen waren präregistriert ([osf.io/6sx79](https://osf.io/6sx79)).

### 3. Ergebnisse

Die Reliabilität des PEPs erreichte ein akzeptables Maß von  $\alpha = .75$ , 95% Konfidenzintervall = [0.68, 0.83]. PEP und RSES korrelierten signifikant hoch,  $r = .51$ ,  $p < .01$ , miteinander. Unsere zentrale Hypothese testeten wir in einer mixed Design ANOVA mit einem between subjects Faktor (Gruppe: Experimentalgruppe vs. Kontrollgruppe) und einem within subjects Faktor (Messmethode: PEP vs. RSES). Abhängige Variable waren die standardisierten Messwerte von PEP und RSES. In der ANOVA zeigte sich ein marginal signifikanter Interaktionseffekt von Gruppe und Messmethode,  $F(1, 100) = 3.61$ ,  $p = .06$ ,  $\eta^2 = .01$ . Die PEP-Effekte waren im Durchschnitt in der Experimentalgruppe signifikant niedriger als in der Kontrollgruppe,  $t(100) = 2.05$ ,  $p = .04$ ,  $d = -.41$ . Für die RSES bestand kein Unterschied  $t(100) = 0.16$ ,  $p = .88$ ,  $d = -.03$ .

### 4. Diskussion

In der vorliegenden Studie wurde der Selbstwert durch eine indirekte (PEP) und eine direkte Methode (RSES) in einem Kontrollgruppendesign gemessen. Wie erwartet, zeigte sich kein signifikanter Unterschied im direkten Maß (RSES) zwischen Experimental- und Kontrollgruppe, jedoch ein signifikanter Unterschied im indirekten Maß (PEP). Der Unterschied der Differenzwerte aus direkter und indirekter Messung zwischen den Gruppen zeigte sich jedoch nur marginal signifikant und in entgegengesetzter Richtung als erwartet, die PEP-Effekte waren in der Experimentalgruppe negativer. Dies lässt vermuten, dass die Methoden nicht das Gleiche messen, jedoch nicht wie erwartet auf selbstwertrelevantes Feedback reagieren. Eine mögliche Erklärung dafür besteht in bekannten Kompensationsmechanismen (Rudman et al., 2007): In der Kontrollgruppe wird ein ernüchterndes Feedback wahrgenommen und die PEP-Effekte fallen deswegen höher aus, weil das positive Selbstkonzept erhalten werden will. Dies ist jedoch eine Spekulation, die in weiteren Studien untersucht werden muss. Außerdem könnte die unterschiedliche Manipulierbarkeit der Maße auch durch zeitlichen Abstand von der Manipulation erklärt werden, da das PEP der RSES für alle Versuchspersonen vorausging. Dennoch konnten wir erste vorläufige experimentelle Evidenz sammeln, dass indirekte und direkte propositionale Selbstwertmaße nicht die gleichen Prozesse abbilden.

### 5. Literatur

- Buhrmester, M. D., Blanton, H., & Swann, W. B., Jr. (2011). Implicit self-esteem: Nature, measurement, and a new way forward. *Journal of Personality and Social Psychology*, *100*(2), 365–385. <https://doi.org/10.1037/a0021341>.
- Dentale, F., Vecchione, M., Ghezzi, V., Spagnolo, G., Szemenyi, E., & Barbaranelli, C. (2020). Beyond an Associative Conception of Automatic Self-Evaluations: Applying the Relational Responding Task to Measure Self-Esteem. *The Psychological Record*, *70*(2), 227+. <https://doi.org/10.1007/s40732-020-00392-4>.
- Dentale, F., Vecchione, M., Ghezzi, V. & Barbaranelli, C. (2016) Applying the Latent State-Trait Analysis to Decompose State, Trait, and Error Components of the Self-Esteem Implicit Association Test. *European Journal of Psychological Assessment*, *35*, 78-85. <https://doi.org/10.1027/1015-5759/a000378>.
- Müller, F., & Rothermund, K. (2019). The Propositional Evaluation Paradigm: Indirect Assessment of Personal Beliefs and Attitudes. *Frontiers in Psychology*, *10*. <https://doi.org/10.3389/fpsyg.2019.02385>.
- Rudman, L. A., Dohn, M. C., & Fairchild, K. (2007). Implicit self-esteem compensation: Automatic threat defense. *Journal of Personality and Social Psychology*, *93*, 798–813.

# Komponenten von Bestrafung und deren Wirkung - Wie nehmen wir Täter\*innen wahr?

Labonte, A., Laqua, F., Marten, S., Schuster, J.

Leitung: Dr. Stefanie Hechler

## 1. Einleitung

In Anbetracht eines Fehlverhaltens sehen Außenstehende Täter\*innen als agentisch (d.h., ihnen wird Kontrolle über ihr eigenes Handeln zugeschrieben; im Folgenden: *Agency*) und gleichzeitig unmoralisch, da sie geteilte Werte verletzt haben (*Moral*; SimanTov-Nachlieli et al., 2013). Beobachter\*innen strafen diejenigen Täter\*innen hart, die sie als intentionale Akteur\*innen und verantwortlich für ihr Fehlverhalten und dessen Ergebnis sehen (Cushman, 2008). Aktuelle Forschungsarbeiten gehen davon aus, dass Strafende durch die Bestrafung mit den Täter\*innen kommunizieren und dadurch soziale Lernprozesse anregen (Sarin et al., 2020). Während die symbolische Komponente die kommunikative Funktion übernimmt (z.B. „Du hast falsch gehandelt.“), verleiht ihr die behaviorale Komponente Nachdruck (z.B. „Du musst den Flur reinigen.“; Feinberg, 1965). Wir erwarten, dass die symbolische Komponente die wahrgenommene *Agency* der Täter\*innen erhöht, während die behaviorale Komponente diese Attribution lediglich unterstützt (H1: Haupt- und Interaktionseffekt). In Bezug auf die wahrgenommene *Moral* der Täter\*innen erwarten wir, dass Täter\*innen nach einer symbolischen Bestrafung als weniger moralisch wahrgenommen werden. Auch hier sollte die behaviorale Komponente unterstützend wirken (H2: Haupt- und Interaktionseffekt).

## 2. Methode

Die Gesamtstichprobe besteht aus 114 Teilnehmenden. Davon gaben elf Personen die falsche Antwort auf einem Aufmerksamkeitstest und wurden deshalb von der Analyse ausgeschlossen. Die Teilnehmenden (64 Frauen, 30 Männern; Alter: 18-77,  $M = 27.89$ ,  $SD = 12.61$ ; 62% Studierende) wurden über persönliche Kontakte der Autor\*innen und Mailverteiler der Universität rekrutiert. Die Studie folgt einem 2 (symbolische Komponente: ja/nein) x 2 (behaviorale Komponente: ja/nein) Zwischensubjekt-Design. Proband\*innen lasen ein Szenario, in dem eine Person Fehlverhalten zeigte und das Opfer (nicht) reagierte. Die Reaktion unterschied sich zwischen den vier Bedingungen: traditionelle Bestrafung (symbolische und behaviorale Komponente), symbolische Komponente, behaviorale Komponente oder keine Konsequenzen. Um die Generalisierbarkeit der Ergebnisse zu erhöhen, lasen die Versuchspersonen zufällig eines von fünf unterschiedlichen Szenarien. Anschließend gaben die Proband\*innen anhand von mehreren Eigenschaften an, wie sie die *Agency* (9 Items, z.B.: „als hätte er\*sie viel Kontrolle“,  $\alpha = .68$ , 95% CI [.59, .77]) und die *Moral* (9 Items, „moralisch“,  $\alpha = .87$ , 95% CI [.84, .91]) der Täter\*innen einschätzten. Alle Items wurden auf einer Skala von 1 = *gar nicht* bis 6 = *sehr* erhoben. Ergebnisse von zusätzlich erhobenen Variablen werden an dieser Stelle nicht berichtet. Die Studie wurde präregistriert: <https://aspredicted.org/blind2.php>. Um unsere Hypothesen zu testen, führten wir zwei zwei-faktorielle Varianzanalysen (ANOVAs) mit den Faktoren „symbolische Komponente“ und „behaviorale Komponente“ von Bestrafung und den abhängigen Variablen wahrgenommene *Agency* oder *Moral* der Täter\*innen durch.

## 3. Ergebnisse

Die Auswertung der Manipulationschecks zu Bestrafung und deren beiden Komponenten zeigen, dass die Manipulationen erfolgreich waren und die Proband\*innen die Operationalisierung wie erwartet verstanden. Die ANOVA zur Hypothesenprüfung (H1) zeigte einen signifikanten Haupteffekt der symbolischen Komponente auf die wahrgenommene *Agency* des\*der Täter\*in ( $F(1,99) = 5.00$ ,  $p =$

0.028,  $\eta^2 = .05$ ). Es gab keinen signifikanten Haupteffekt der behaviorale Komponente ( $F(1, 99) = 0.87, p = .352, \eta^2 = .009$ ) und keinen Interaktionseffekt auf *Agency* ( $F(1, 99) > 0.01 (0.0001), p = .989, \eta^2 = .000$ ). Entgegen unserer Hypothese schätzten die Versuchspersonen die *Agency* des\*der Täter\*in als niedriger ein, wenn die Reaktion des Opfers eine symbolische Bestrafungskomponente enthielt ( $M_{traditionell} = 2.90, SD = 0.72; M_{symbolisch} = 3.03, SD = 0.78$ ) als wenn diese nicht vorhanden war ( $M_{behavioral} = 3.22, SD = 0.76; M_{keine} = 3.36, SD = 0.76$ ). Die behaviorale Komponente hatte keinen Einfluss auf die zugeschriebene *Agency*. Die ANOVA zur Hypothesenprüfung (H2) der abhängigen Variablen *Moral* zeigte weder einen signifikanten Haupteffekte der symbolischen Komponente ( $F(1,99) = 0.29, p = 0.590, \eta^2 = .003$ ), noch der behavioralen Komponente ( $F(1, 99) = 0.04, p = .836, \eta^2 = .000$ ). Es gab keinen Interaktionseffekt auf *Moral* ( $F(1, 99) = 2.21, p = .140, \eta^2 = .022$ ). Entsprechend unterschied sich die wahrgenommene *Moral* des\*der Täter\*in nicht signifikant hinsichtlich der beiden Komponenten von Bestrafung ( $M_{traditionell} = 2.87, SD = 1.12; M_{symbolisch} = 2.64, SD = 0.9; M_{behavioral} = 2.70, SD = 0.89; M_{keine} = 3.01, SD = 0.84$ ).

#### 4. Diskussion

Das Ziel unserer Studie war es, Erkenntnisse zu erlangen, wie Bestrafung sich auf die Wahrnehmung von Täter\*innen auswirkt. Eine rein behaviorale Aktion der Opfer hatte keinen Einfluss auf die Wahrnehmung der Täter\*innen. Unsere Erwartung, dass eine symbolische Bestrafung die wahrgenommene *Agency* der Täter\*innen erhöht, konnte nicht bestätigt werden. Im Gegenteil nahmen Außenstehende Täter\*innen sogar als weniger agentisch wahr, wenn das Opfer ihre Tat als falsch und sie als verantwortlich bezeichnete. Das könnte bedeuten, dass eine behaviorale Konsequenz, vor allem bei kleinen Fehlritten, über eine verbale Zurechtweisung hinaus nicht unbedingt sinnvoll ist. Entgegen unseren Erwartungen zeigte sich, dass Beobachter\*innen die Täter\*innen unabhängig von der Bestrafung als ähnlich unmoralisch wahrnehmen. Möglicherweise waren die beschriebenen Vergehen in den Szenarien nicht schwerwiegend genug, um zu bewirken, dass Täter\*innen in der einen Bedingung als unmoralischer als in den anderen Bedingungen wahrgenommen werden. Wir haben die genutzten Szenarien speziell für diese Studie entwickelt und auf die vorliegende Studie angepasst. Dennoch muss beachtet werden, dass diese empirisch validiert werden müssten, um mögliche Fehlerquellen zu identifizieren und zu minimieren. Für zukünftige Studien könnte es interessant sein, ob wahrgenommene Unterschiede in den Szenarien einen Einfluss auf die Beurteilung haben. Über die verschiedenen Szenarien hinweg unterschied sich die Wahrnehmung der *Agency* des\*der Täter\*in nach einer symbolischen Bestrafung von keiner symbolischen Bestrafung. Zusammenfassend kann gesagt werden, dass sich vor allem die Kommunikation von Missbilligung darauf auswirkt, wie Täter\*innen hinsichtlich ihrer *Agency* wahrgenommen werden. Wenn diese vorhanden ist, wird den Täter\*innen weniger Handlungsfähigkeit zugeschrieben. Was für Konsequenzen könnte es haben, dass Bestrafung die Wahrnehmung von Täter\*innen beeinflusst? Welche Auswirkungen hat die Wahrnehmung von Täter\*innen als nicht agentisch auf den Umgang mit ihnen? Wann werden sie wieder in die Gemeinschaft aufgenommen und welchen Beitrag hat hier die Bestrafung?

#### 5. Literatur

- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Science Direct, Cognition (108)*(2), 353-380. <https://doi.org/10.1016/j.cognition.2008.03.006>
- Feinberg, J. (1965). The Expressive Function of Punishment. *Monist 49*(3), 397-423. <https://doi.org/10.5840/monist196549326>
- Sarin, A. et al. (2020). Punishment is Organized around Principles of Communicative Inference. *Cognition 208*, 104544. <https://doi.org/10.1016/j.cognition.2020.104544>
- Shnabel, N., & Nadler, A. (2008). A needs-based model of reconciliation: Satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. *Journal of Personality and Social Psychology, 94*(1), 116–132. <https://doi.org/10.1037/0022-3514.94.1.116>

# Wer ist der Täter? Wie Attraktivität die Auswahl in Lineups beeinflusst

Antonia Grüner, Anja Büttner, Nicole Keller, Christin Strutz

Leitung: Dr. Jürgen Kaufmann und B.Sc. Pauline Gürtler

## 1. Einleitung

„What is beautiful is good“ ist der Titel einer Studie von Dion et al. (1972), die als eine der ersten zeigte, dass attraktiven Menschen eher positive und unattraktiven Menschen eher negative Charakteristika zugeschrieben werden. Dieser Effekt wurde später in verschiedenen Studien in unterschiedlichen Zusammenhängen repliziert. So bekommen zum Beispiel unattraktive Kriminelle im Schnitt härtere Strafen als attraktive (Darby & Jeffers, 1988). Es ist also anzunehmen, dass dieser Bias auch in anderen Situationen existiert. Wenig untersucht ist dabei die Auswirkung des Attraktivitäts-Bias in Augenzeugensituationen. Diese Situation ist gerade deswegen von Brisanz, da in den letzten Jahren einer Reihe von Verurteilten ihre Unschuld bewiesen werden konnte, welche nur aufgrund von Augenzeugenberichten verurteilt wurden (Eysenck, 2015).

In dieser Studie ging es also darum zu überprüfen, ob die Attraktivität der Personen in einem Lineup bei einer unklaren Tatsituation eine Rolle für die Auswahl des mutmaßlichen Täters durch Augenzeugen spielt. Dabei wurde basierend auf vorheriger Forschung vermutet, dass eher unattraktive Personen ausgewählt werden. Des Weiteren sollte untersucht werden, ob attraktive Täter eine kürzere Strafe bekommen.

## 2. Methode

*Stichprobe.* Bisher bearbeiteten 35 Versuchspersonen das Experiment vollständig. Von ihnen musste eine Person wegen Nichtbefolgen der Anweisungen ausgeschlossen werden, sodass die Daten von insgesamt 34 Versuchspersonen (25 weiblich, 9 männlich und 1 divers) für die vorläufige Analyse herangezogen wurden. Der Großteil dieser Probanden war zwischen 18 und 24 Jahre alt (27), vier waren zwischen 25 und 30, drei zwischen 31 und 40 und eine Person war über 40. Die Versuchspersonen wurden aus dem eigenen Bekanntenkreis und über die Mailingliste des FSR Psychologie der FSU Jena rekrutiert.

*Material.* Das Stimulusmaterial bestand aus 30 verschiedenen mittelalten männlichen Gesichtern, deren Attraktivität in einer Vorstudie bewertet wurde. Sie wurden in fünf Gruppen à sechs Gesichter aufgeteilt, von denen je zwei als attraktiv, zwei als neutral und zwei als unattraktiv bewertet wurden. Aus diesen sechs Gesichtern wurde anschließend online mit WebMorph (v0.0.0.9001) ein Gesicht erstellt, das zu gleichen Teilen aus den verwendeten Bildern bestand (DeBruine, 2018). Dieser Prozess wurde für alle fünf Gruppen durchgeführt, sodass insgesamt fünf gemorphte Gesichter entstanden. Aus den jeweils sechs zusammengehörigen Bildern wurden mit PicCollage Lite - Collage Maker (1.1) Collagen erstellt, welche als Lineups dienten (Softase Tech Co., Limited, 2013). Außerdem wurden 5 Bilder von männlichen Personen, die beispielsweise von Sicherheitskameras aufgenommen wurden, verwendet. Über deren Gesichter wurde mit GIMP (2.10.24) jeweils eins der gemorphten Gesichter eingefügt (The GIMP Team, 2021). Die resultierenden Bilder wurden als Täterbilder verwendet. Durch den vorherigen Prozess des Morphens sollte vermieden werden, dass einige der sechs Gesichter eines Lineups dem angeblichen Tätergesicht ähnlicher sahen als andere.

*Ablauf.* Die Datenerhebung fand online über die Software PsyToolkit (3.3.2) statt (Stoet, 2010, 2017). Das Experiment bestand aus fünf Blöcken, die in randomisierter Reihenfolge präsentiert wurden. Nach einer kurzen Einweisung, in der den Versuchspersonen unter anderem erklärt wurde, dass sie im Folgenden Zeugen einer Straftat sein würden, begann das Experiment mit dem ersten Block. Die Probanden lasen die Beschreibung einer fiktiven Straftat und sahen 4000ms lang ein Bild des angeblichen Täters. Dieses wurde in Wirklichkeit von uns erstellt. Danach folgten zwei

Distraktorfragen und anschließend ein Lineup, aus dem die Versuchspersonen den Täter auswählen sollten, was in Wirklichkeit allerdings unmöglich war. Im Lineup waren zwar die sechs Gesichter, aus denen das Tätergesicht erstellt wurde, enthalten, dieses selbst jedoch nicht. Nachdem eins der sechs Gesichter gewählt wurde, erfragten wir wie sicher sich die Versuchsperson war. Anschließend wurde erklärt zu wie vielen Jahren Haftstrafe der Schuldige verurteilt wurde und die Versuchspersonen sollten angeben, ob sie mit dem Urteil einverstanden waren oder selbst mehr oder weniger Strafe gegeben hätten. Dieser Ablauf wiederholte sich insgesamt viermal, sodass alle fünf Blöcke durchlaufen wurden.

### 3. Ergebnisse

Die Daten wurden in RStudio (1.3.1093) ausgewertet (RStudio Team, 2020). Mit einem einseitigen Einstichproben-*t*-Test wurden die über alle fünf Lineups gemittelten Attraktivitätsratings des von einer Versuchsperson als Tätergesicht ausgewählten Gesichts mit dem Mittelwert der Attraktivitätsratings aller 30 verwendeten Gesichter verglichen. Der Test zeigte, dass die Attraktivitätsratings der ausgewählten Gesichter ( $M = 2.47$ ) signifikant niedriger waren als der Mittelwert aller Attraktivitätsratings ( $M = 2.77$ ),  $t(33) = -3.07$ ,  $p = .002$ .

### 4. Diskussion

In mehreren Studien, wie beispielsweise Dion et al. (1972), konnte bereits gezeigt werden, dass attraktiven Menschen eher positive Eigenschaften zugeschrieben werden, während dies bei unattraktiven Menschen eher mit negativen Eigenschaften der Fall ist. Daraus ergab sich die Frage, ob dieser Attraktivitäts-Bias auch bei der Auswahl eines Täters aus einem Lineup durch Augenzeugen eine Rolle spielt, also angenommen wird, dass unattraktive Personen eher kriminell sind. Es zeigte sich, dass nach den bisherigen Daten in unserer Studie, wie erwartet, eher die unattraktiven Gesichter aus den Lineups ausgewählt wurden. Dies entspricht der bisherigen Fachliteratur. Die Frage inwiefern unattraktive Täter eine längere Strafe bekommen, ist aufgrund fehlender Daten noch nicht zu beantworten. Ebenfalls sollten die aktuellen Ergebnisse aufgrund der noch nicht erreichten Zielanzahl der Probanden noch nicht als final, sondern eher als Hinweis auf einen Zusammenhang angesehen werden. Abschließend muss erwähnt werden, dass trotz bestmöglichem Bemühen durch das Morphen der Gesichter in den Lineups nicht ganz ausgeschlossen werden kann, dass die Probanden die mutmaßlichen Täter auch nach Ähnlichkeit, beispielsweise der Kinnform, ausgewählt haben.

### 5. Literatur

- Darby, B. W. & Jeffers, D. (1988). The effects of defendant and juror attractiveness on simulated courtroom trial decisions. *Social Behavior and Personality: an international journal*, 16(1), 39–50. <https://doi.org/10.2224/sbp.1988.16.1.39>
- DeBruine, L. (2018). *WebMorph: Beta Release 2* (v0.0.0.9001) [Computer Software]. Zenodo. <https://doi.org/10.5281/zenodo.1162670>
- Dion, K., Berscheid, E. & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24(3), 285–290. <https://doi.org/10.1037/h0033731>
- Eysenck, M. W. (2015). Eyewitness testimony. In A. Baddeley, M. W. Eysenck & M. C. Anderson (Hrsg.), *Memory* (2. Aufl., S. 329–358). Psychology Press.
- RStudio Team. (2020). *RStudio: Integrated Development Environment for R* (1.3.1093) [Computer Software]. RStudio, PBC. <https://www.rstudio.com/>
- Softase Tech Co., Limited. (2013). *PicCollage Lite - Collage Maker* (1.1) [App]. Mac App Store. <https://apps.apple.com/us/app/piccollage-lite-collage-maker/id724109472>
- Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096–1104. <https://doi.org/10.3758/brm.42.4.1096>
- Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, 44(1), 24–31. <https://doi.org/10.1177/0098628316677643>
- The GIMP Team. (2021). *GIMP* (2.10.24) [Computer Software]. <https://www.gimp.org/>

# Wenn Stimmen hervorstechen - Effekte von Karikierung und selektiver Aufmerksamkeit auf die Wahrnehmung von Gesangsgruppen

Lisa Gärtner, Sinha Kirchhoff, Marie Limmer, Vivien Schaffer

Leitung: PD Dr. Romi Zäske

## 1. Einleitung

In unserem Alltag treffen wir auf viele verschiedene Stimmen, von denen einige deutlich mehr herausstechen als andere. Während uns diese Eigenschaft im Alltag intuitiv hilft, Einzelstimmen aus einer Menge herauszuhören und zu unterscheiden, sind distinkte Stimmen im Chorgesang eher unerwünscht. Der Einfluss von wahrgenommener Distinktheit auf die Segregation mehrerer gleichzeitiger Schallquellen [1] wurde für (Sing-)Stimmen jedoch noch nicht experimentell untersucht. Ausgehend von der Modellannahme, dass sich distinkte Stimmen im Vergleich zu nicht-distinkten Stimmen stärker voneinander, aber auch stärker von durchschnittlichen Stimmen in einem akustischen ‘Voice Space’ unterscheiden [2], manipulieren wir Stimmendistinktheit mithilfe akustischer Karikierung. Wir vermuten, dass karikierte (distinkte) Stimmen stärker aus einer Gruppe von Singstimmen hervorstechen als ihre jeweilige Originalversion bzw. ihre anti-karikierte (nicht-distinkte) Version. Zur Untersuchung nutzen wir das Disappearance-Paradigma [3], bei welchem zwei aufeinander folgende, auditive Szenen verglichen werden müssen. Das Verschwinden einer Teststimme in der zweiten Szene sollte am ehesten für Karikaturen bemerkt werden, am wenigsten für Antikarikaturen. Basierend auf Befunden, dass Aufmerksamkeitslenkung die ‘change-deafness’ bei dieser Aufgabe reduzieren kann [3], sagen wir zudem vorher, dass sich die Performanz erhöht, wenn die Probanden zuvor einen Hinweis über die Art der Veränderung bekommen (Exp. 1), im Vergleich zu der Aufgabe ohne diesen Hinweis (Exp. 2).

## 2. Methode

Pro Experiment dieser Online-Studie [5,6] wurden 14 Datensätze ausgewertet (Exp. 1: 6 Frauen,  $M = 22,6$  Jahre, 19-30 Jahre; Exp. 2: 9 Frauen,  $M = 23,3$  Jahre, 19-28 Jahre). Alle Probanden waren deutsche Muttersprachler ohne Hörschädigung. Eine a priori Poweranalyse ergab, dass ein mittelgroßer Effekt einen Stichprobenumfang von  $N = 28$  pro Experiment benötigt, um mit einer Teststärke von 0.8 aufgedeckt zu werden. Grundlage waren 50 Vokal-Stimuli [a, e, i, o, u] von Männern und Frauen (18 – 25 Jahre) aus der JESS-Datenbank [4]. Jede Originalaufnahme dauerte 1,5 s und lag zusätzlich als karikierte und anti-karikierte Version vor.

*Distinktheitsratings* (Manipulationscheck): zunächst schätzen die Probanden die Distinktheit für alle 150 Einzelstimmen auf einer 6-stufigen Likert-Skala ein: entweder hinsichtlich der Abweichung der jeweiligen Stimmen von ihnen bekannten Stimmen (DEV-Distinktheit) oder hinsichtlich des Grades an Herausstechen aus der Menge (VITC-Distinktheit).

*Disappearance-Task*: In jedem Trial hörten die Probanden zwei aufeinanderfolgende Multisänger-Stimuli (S1 und S2), welche entweder identisch waren und 3 Stimmen enthielten oder bei denen in S2 eine Stimme aus S1 fehlte. Die Probanden entschieden per Tastendruck entweder, ob S2 “weniger” Sänger enthielt als S1 (Exp. 1, spezifische Aufgabe, d.h. mit Hinweis auf Art der Veränderung) oder die Sängergruppen sich “unterscheiden” (Exp. 2, unspezifische Aufgabe, d.h. ohne Hinweis auf Art der Veränderung). Als S1 Stimuli dienten alle Kombinationen von 3 Originalstimmen innerhalb desselben Geschlechts und Vokals sowie Versionen, in denen jeweils eine der Stimmen durch ihre jeweilige Karikatur bzw. Anti-Karikatur ersetzt wurde, sodass 3 Karikaturgrade entstanden (Anti-karikiert, Original, Karikatur). Für S2 Stimuli, die aus nur 2 Stimmen bestanden, verwendeten wir ausschließlich Originale. Die Disappearance-Task bestand aus 120 (Exp. 1) bzw. 150 (Exp. 2) voll-randomisierten Trials. Inklusive Pausen dauerte jedes Experiment ca. 35-40 min und wurde auf Wunsch mit einer Vpn-Stunde vergütet.



### 3. Ergebnisse

*Distinktheitsratings*: Der Manipulationscheck mittels ANOVA mit MW auf Karikaturgrad (3) und mit Distinktheitsmaß (2) als Gruppenfaktor bestätigte eine sign. Zunahme der wahrgenommenen Distinktheit mit Erhöhung des Karikaturgrades: HE Karikaturgrad  $F(2, 52) = 45.43, p < .001, \eta^2 = .64$ . Alle paarweisen Vergleiche der 3 Stufen des Faktors Karikaturgrad wurden sign. ( $-7,2 \leq ts[27] \leq -2,94, ps \leq .007$ ). Es ergab sich kein Effekt des Distinktheitsmaßes und keine Interaktion beider Faktoren. DEV und VITC Distinktheit korrelierten sign. positiv ( $\rho = .36, p < .001$ ). Getrieben wurde die Korrelation von karikierten Stimmen ( $\rho = .66, p < .001$ ).

*Disappearance-Task*: Eine ANOVA für Akkuratheit mit MW auf Karikaturgrad (3) und mit Aufgabe (2) als Gruppenfaktor ergab einen sign. HE des Karikaturgrades,  $F(2, 52) = 28.95, p < .001, \eta^2 = .53$ , wobei auf Karikaturen genauer reagiert wurde als auf Originalstimmen ( $t(27) = -7.15, p < .001$ ) und Antikarikaturen ( $t(27) = -6.07, p < .001$ ). Zudem ergab sich ein unerwarteter Trend für genauere Antworten bei Antikarikaturen im Vgl. zu Originalen ( $t(27) = 1.85, p = 0.076$ ). Für den HE Aufgabe zeigte sich ein Trend,  $F(1, 26) = 4.22, p = .050, \eta^2 = .14$  in die erwartete Richtung, mit genaueren Antworten für die spezifische vs. der unspezifischen Aufgabe. Es gab keine Interaktion zwischen beiden Faktoren und keine sign. Effekte in den Reaktionszeiten.

### 4. Diskussion

Je stärker sich akustisch manipulierte Stimmen von einer Durchschnittsstimme unterschieden, desto distinkter wurden sie jeweils wahrgenommen im Einklang mit dem Voice Space Modell [2]. Wie auch [4] fanden wir eine pos. Korrelation beider Distinktheitsmaße (VITC und DEV), sodass sie teilweise das gleiche Konstrukt abbilden und dazu dienen könnten, passende Sänger für ein Ensemble auszuwählen. Hörer können Unterschiede zwischen Stimmenszenen bestehend aus 3 bzw. 2 Stimmen gut, jedoch nicht perfekt wahrnehmen, da wir ein gewisses Maß von „change deafness“ beobachten konnten [3]. Veränderungen zwischen zwei Stimmenszenen wurde leichter bemerkt, wenn eine (distinkte) karikierte Stimme im Vergleich zu einer Originalstimme wegfiel. Im Widerspruch zu unserer Hypothese wurde das Wegfallen einer (nicht-distinkten) Anti-Karikatur im Vergleich zu einer Originalstimme im Trend ebenfalls leichter bemerkt. Ob dies ein genuiner Effekt ist, sollte mit einer größeren Stichprobe und Power untersucht werden, was uns aufgrund des Zeitlimits nicht möglich war. So könnte auch abschließend geklärt werden, ob ein Hinweis auf die Art der Veränderung tatsächlich hilft, Unterschiede zwischen Sängergruppen zu bemerken, wie wir es hier nur trendweise feststellen konnten. Unsere Ergebnisse lassen vermuten, dass distinkter wahrgenommene Stimmen das Segregieren verschiedener, gleichzeitig präsentierter Stimmen erleichtern und die „change deafness“ reduzieren. Stimmendistinktheit sowie Top-down Prozesse könnten demnach die Segregation von Stimmen unterstützen [3] und infolgedessen helfen, einem Gespräch mit mehreren Sprechern zu folgen. In Kontexten, in denen ein homogenes Klangbild angestrebt wird, beispielsweise in Chören, sollten eher weniger distinkte Stimmen gewählt werden.

### 5. Literatur

1. Bregman, A.S., *Auditory scene analysis: The perceptual organization of sound*. 1990: The MIT Press.
2. Latinus, M., et al., Norm-Based Coding of Voice Identity in Human Auditory Cortex. *Current Biology*, 2013. 23(12): p. 1075-1080.
3. Eramudugolla, R., et al., Directed attention eliminates 'Change deafness' in complex auditory scenes. *Current Biology*, 2005. 15(12): p. 1108-1113.
4. Zäske, R., et al., The Jena Speaker Set (JESS)-A database of voice stimuli from unfamiliar young and old adult speakers. *Behavior Research Methods*, 2020. 52(3): p. 990-1007.
5. Stoet, G., PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 2010. 42(4): p. 1096-1104.
6. Stoet, G., PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, 2017. 44(1): p. 24-31.

## **Gesellschaftliche Normen – Die Mutter aller Vorurteile?**

**Nadin Merten, A.V., Laura Müller, Leonie Arlt, Philipp Bornschein**

Leitung: Prof. Dr. Thomas Kessler

### 1. Einleitung

Ausländerfeindlichkeit kann zu verschiedenen individuellen, sozialen und gesellschaftlichen – typischerweise negativen – Konsequenzen führen. Es ist deshalb von gesellschaftlicher Bedeutung, Vorurteile genauer zu verstehen und vorherzusagen. In unserer Studie untersuchen wir, ob Normen Vorurteile erklären können. Darüber hinaus betrachten wir, welche Rolle individuelle Persönlichkeitsfaktoren wie etwa Autoritarismus bei der Vorhersage von Vorurteilen spielen. Beide Ansätze vergleichen wir miteinander, vor allem um zu erfahren, welcher sich eher dazu eignet, Vorurteile zu erklären.

Unter Vorurteilen versteht man die Bewertung von Individuen aufgrund ihrer Zugehörigkeit zu einer sozialen Gruppe, ungeachtet ihrer individuellen Merkmale (Kessler & Fritsche, 2018). Laut Sherif (1948) sind Vorurteile Ausdruck unserer gesellschaftlichen Normen, die als Regeln und Standards verstanden werden, die innerhalb einer Gruppe geteilt werden und die unser soziales Verhalten leiten (Cialdini & Trost, 1998). Wer ein „gutes“ Gruppenmitglied sein möchte, muss die Werte und Einstellungen der anderen Mitglieder in dieser Gruppe teilen. Ausgehend von Sherifs Überlegungen untersuchen wir in unserer Studie die Annahme, dass Normen und Vorurteile eng zusammenhängen. Wir fokussieren hierbei auf Ausländerfeindlichkeit und erheben Vorurteile gegenüber vielen unterschiedlichen Vorurteiltargets. Unsere Hypothese lautet: Normen und Vorurteile gegenüber Einwohner\*innen anderer Länder und Angehöriger ethnischer Minoritäten korrelieren hoch positiv miteinander.

Für die Untersuchung interindividueller Unterschiede erfassen wir Autoritarismus, der in der Vergangenheit reliabel aber nur gegenüber ausgewählten Targets mit Vorurteilen korreliert (Sibley & Duckitt, 2008). In unserer Studie untersuchen wir Vorurteile gegenüber vielen verschiedenen Targets, um mögliche Selektionseffekte zu minimieren. Unsere Annahme ist, dass einige Targets positiv und andere negativ mit der Autoritarismus-Skala korrelieren. Je autoritärer eine Person ist, desto ausländerfeindlicher wird sie sein. Unsere abschließende Hypothese besagt, dass Normen und Vorurteile gleichermaßen mit dem Autoritarismus (RWA) korrelieren, da Normen und Vorurteile hoch miteinander korrelieren.

### 2. Methode

Bei dem Studiendesign handelt es sich um ein Between-Subject-Design mit den Bedingungen Vorurteile und Normen für die 55 Targets. Die Teilnehmer bestanden aus 63 Frauen, 80 Männern und 3 diversen Personen im Alter von 18 bis 72 Jahren.

Erfragt wurde vorab die politische Gesinnung auf einer Skala von 1 (links orientiert) bis 10 (rechts orientiert) sowie die Bewertung von Aussagen zur Identifikation der Teilnehmer\*innen mit Europa und Deutschland. Autoritarismus wurde mittels einer deutschen Version der RWA-Skala (Altmeyer, 1981) gemessen, die aus 9 Items besteht,  $\alpha=.84$ .

Für die Erhebung der Daten nutzten wir das Online-Portal Social Science Survey, auf welchem wir beide Fragebögen getrennt nach den Bedingungen veröffentlichten.

Für die Erfassung der Vorurteile erstellten wir eine Liste mit Targets, die sowohl in der Norm- als auch der Vorurteilsbedingung verwendet wurden. Wir nahmen alle Staaten des Europarates als Targets in unsere Liste auf. Zusätzlich wurden noch Staaten aufgenommen, die zwar keine Mitglieder des Europarats sind, aber dennoch eine starke Präsenz in Europa haben (Vietnam, Syrien und Afghanistan). Weiterhin fügten wir der Target-Liste einige Europäische Minderheiten hinzu (Lappen, Katalanen, Basken, Friesen und Sorben).

### 3. Ergebnisse

Für jedes Target wurde der Mittelwert der Normen über alle Versuchspersonen hinweg berechnet. Ebenso für die Vorurteile. Die Mittelwerte der explizit erhobenen Normen und Vorurteile gegenüber spezifischen Targetgruppen korrelieren hoch negativ miteinander ( $r = -.76$ ,  $p < 0.01$ ). Je mehr gesellschaftliche Normen negative Gefühle gegenüber einer Gruppe dulden, desto höher sind die negativen Vorurteile gegenüber dieser Gruppe.

Betrachtet man die Korrelationen zwischen den Vorurteilen und Autoritarismus, so finden sich vor allem für Syrien ( $r = .35$ ), für Afghanistan ( $r = .40$ ) und für Deutschland ( $r = -.36$ ) auffällig hohe Werte, die den Trend in der Literatur bestätigen. Schaut man sich jedoch die Korrelationen für die weiteren Targets an, so wird deutlich, dass Autoritarismus Vorurteile für viele Gruppen kaum zuverlässig vorhersagt,  $r = .26$  bis  $r = -.24$ . Zudem korreliert Autoritarismus relativ hoch mit dem politischen Spektrum ( $r = .48$ ) und mäßig hoch mit der Identifikation mit dem Erhebungsland Deutschland ( $r = .30$ ). Die Angaben zu Normen und Vorurteilen zeigen, dass negative Gefühle gegenüber Ethnien allgemein eher als negativ betrachtet werden, am negativsten gegenüber Deutschland.

### 4. Diskussion

Auf Grundlage von Sherifs Annahmen (1948) vermuteten wir, dass Normen und Vorurteile gegenüber Einwohner\*innen anderer Länder und Mitglieder ethnischer Minderheiten hoch positiv miteinander korrelieren. Die Analyse des Zusammenhangs zwischen Normen und Vorurteilen gegenüber mehr als 50 unterschiedlichen Ländern ergibt eine signifikante positive Korrelation ( $r = -.76$ ) zwischen beiden Variablen, was unsere Hypothese bestätigt. Da es sich hierbei nur um Korrelationen handelt, können keine Aussagen über die Wirkrichtung getätigt werden. Besonders zu beachten bleibt jedoch das Ergebnis, dass Normen und Vorurteile über das gesamte politische Spektrum ähnlich vertreten sind. Das bedeutet, dass Normen scheinbar von den meisten Individuen innerhalb einer Gesellschaft angenommen werden und Personen weniger individuell als konform der Gruppennorm handeln.

Weiterhin untersuchten wir die Annahme, dass autoritäre Persönlichkeiten mehr Vorurteile gegenüber Einwohner\*innen aus anderen Ländern oder ethnischen Minoritäten haben (Sibley und Duckitt, 2008). Die Datenanalyse zeigte dabei für die meisten Länder eine Korrelation zwischen der RWA-Skala und den Mittelwerten von Normen bzw. Vorurteilen nahe null, was für ein Verwerfen der Hypothese spricht. Es zeigt sich also entgegen den Annahmen der vorherrschenden Literatur, dass der Autoritarismus die Ausprägtheit von Vorurteilen nicht gut vorhersagt. Allerdings stachen einige Länder mit nicht signifikanten, aber trotzdem höheren Korrelationen heraus (z.B. Deutschland, Afghanistan). Dies ist ein Hinweis auf die Bestätigung unserer Hypothese, dass keine universelle, sondern lediglich eine gruppenspezifische Ausländerfeindlichkeit existiert. Da die Erhebung der Daten ausschließlich in unserem Familien- und Bekanntenkreis erfolgte und keine weiteren inferenzstatistischen Analysen durchgeführt wurden, bleiben die Ergebnisse aber nur eingeschränkt repräsentativ.

### 5. Literatur

- Cialdini, R. B. & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Hrsg.), *The handbook of social psychology* (4. Aufl., Bd. 2, S. 151-192). New York: McGraw-Hill.
- Kessler, T. & Fritsche, I. (2018). *Sozialpsychologie*. Springer.
- Sherif, M. (1948). The necessity of considering current issues as part and parcel of persistent major problems: Illustrated by the problem of prejudice. *International Journal of Opinion and Attitude Research*, S. 63-68.
- Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: a meta-analysis and theoretical review. *Personality and Social Psychology Review*, 12, 248-279.

# Traumurlaubsziel vs. Mordgelüste: Wie verändern sich automatische Gedächtnisprozesse des Lügens bei intimen Themen im Vergleich zu oberflächlichen Themen?

Melanie Bomm, Sissy Höbelt, Selina Hußner, Rebecca Matthias, Lea Ulverich

Leitung: M.Sc. Franziska Schreckenbach

## 1. Einleitung

Das Wissen darüber, eine Frage unehrlich beantwortet zu haben, wird automatisch aus unserem Langzeitgedächtnis abgerufen, wenn wir dieser Frage erneut begegnen (Koranyi et al., 2015). Doch ist das bei allen Lügen gleich oder wird dieser Abruf je nach Frageninhalt beeinflusst? Ist das Wissen über eine Lüge auf intime Fragen im Vergleich zu oberflächlichen präziser, sodass der Abruf schneller geschieht, um uns vor potenziellen Blamagen zu schützen? Vor dem Hintergrund dieser Fragestellung untersuchten wir, ob es Unterschiede in diesen automatischen Gedächtnisprozessen gibt, je nach thematischem Inhalt der Frage. Eine Vorstudie, in der Fragen nach ihrer Intimität und Wahrscheinlichkeit, darauf zu lügen, eingeschätzt wurden, sowie ein digitales Probandeninterview mit anschließendem Priming-Experiment dienten der Hypothesenüberprüfung. Wir erwarteten aufgrund bisheriger Forschungsergebnisse kürzere Reaktionszeiten bei kongruenten Prime-Target-Kombinationen. Des Weiteren nahmen wir an, dass dieser Effekt stärker bei Fragen mit intimmem Inhalt zu Tage tritt.

## 2. Methode

*Stichprobe.* Anhand einer Poweranalyse mit G\*Power ( $\eta_p^2 = .11$ ,  $\alpha\text{-level} = .05$ , Power  $1-\beta = .8$ ) wurde eine erforderliche Stichprobe von 66 ermittelt. Aufgrund der begrenzten Ressourcen und somit kurzem Erhebungszeitraum erreichten wir nur eine Probandenzahl von 40. Davon wurden aufgrund zu hoher Fehlerraten weitere 7 Proband\*innen ausgeschlossen. Die übrigen 33 Teilnehmenden (25 Frauen) hatten ein durchschnittliches Alter von  $M = 21.61$  ( $SD = 3.66$ ) und waren Studierende verschiedener Studiengänge (überwiegend Psychologie).

*Material.* Bei unserer Vorstudie wurden von 32 Teilnehmenden Fragen auf zwei 5-stufigen Skalen bezüglich ihrer Intimität und Wahrscheinlichkeit, diese in einer anonymisierten Studie ehrlich zu beantworten, bewertet. Dabei wurden die Fragen, die mindestens einen Mittelwert von 3 auf der Ehrlichkeitskala aufwiesen, in 2 Kategorien eingeteilt: intime Fragen (Intimität  $> 3.7$ ) und oberflächliche Fragen (Intimität  $< 1.4$ ). Diese wurden dann gleichmäßig auf zwei Fragenblöcke aufgeteilt, sodass jeder Block zur Hälfte aus intimen ( $M = 4.1$ ) und zur anderen Hälfte aus oberflächlichen Fragen ( $M = 1.3$ ) bestand. In einem standardisierten Interview über Zoom sollten die Proband\*innen, ausbalanciert anhand ihrer VPN-Nummer, zu einem der beiden Blöcke lügen, ohne dass die Versuchsleiterin wusste, wann sie lügen. Im darauffolgenden E-Prime Go Experiment dienten die Fragen aus dem Interview als Prime und als Targets die verschlüsselt präsentierten Worte „gelogen“ und „ehrlich“.

*Design.* Wir untersuchten eine within-subject 2 (Prime: Frage, auf die ehrlich oder gelogen geantwortet wurde) x 2 (Target: ehrlich vs. gelogen) x 2 (Topic: intim vs. oberflächlich) ANOVA. Als abhängige Variablen dienten die Reaktionszeiten in Millisekunden und die Accuracy.

*Ablauf.* Nach dem Debriefing und Instruktionen über E-Prime und dem standardisierten Interview über Zoom fand das Priming-Experiment in E-Prime statt. Die Trials begannen mit einem Fixationskreuz, das für 500ms präsentiert wurde. Anschließend folgte die Wort-für-Wort-Präsentation einer Frage aus dem Interview und zuletzt eines der verschlüsselten Targets „ehrlich“ oder „gelogen“. Auf diese sollte mit den Tasten „f“ oder „d“ reagiert werden, denen, über die Probanden ausbalanciert, die Konzepte „ehrlich“ oder „gelogen“ zugeordnet waren. Zwischen den Trials wurden Kontrolltrials eingebaut. Das Experiment wurde auf [aspredicted.org](https://aspredicted.org) präregistriert.

### 3. Ergebnisse

Wir führten für unsere beiden abhängigen Variablen, Reaktionszeit in Millisekunden und die Accuracy, je eine 2x2x2 ANOVA durch.

Die erste ANOVA mit den within-Faktoren Prime, Target und Intimität und Reaktionszeit in Millisekunden als abhängige Variable ergab einen signifikanten Effekt bei Intimität,  $F(1,32) = 6.95$ ,  $p = .013$ ,  $\eta_p^2 = .179$ . Dieser Effekt belegt, dass die Proband\*innen auf intime Fragen signifikant schneller reagierten als auf oberflächliche Fragen.

Die zweite ANOVA mit den gleichen Faktoren und Accuracy als abhängige Variable ergab einen signifikanten Effekt bei Prime,  $F(1,32) = 5.34$ ,  $p = .027$ ,  $\eta_p^2 = .143$ . Dieses Ergebnis spricht dafür, dass es signifikante Unterschiede in der Accuracy gibt, abhängig davon, ob die Teilnehmenden auf Lüge oder Wahrheit geprint wurden. Deskriptiv lässt sich in Bezug auf die Accuracy eine leichte Tendenz erkennen mit Prime x Target  $F(1,32) = 2.71$ ,  $p = .11$ ,  $\eta_p^2 = .078$ , insofern dass kongruente Bedingungen akkurater beantwortet wurden als inkongruente. Die Dreifachinteraktion Prime x Target x Intimität lieferte keine signifikanten Ergebnisse mit  $F(1,32) = .021$ ,  $p = .88$ ,  $\eta_p^2 < .001$ , was bedeutet, dass sich Kongruenzeffekte nicht in Abhängigkeit von der Intimität verstärken würden.

Auf Basis dieser Ergebnisse müssen wir unsere Hypothese, dass es geringere Reaktionszeiten bei kongruenten Prime-Target Passungen gäbe als bei inkongruenten Prime-Target Passungen, verwerfen. Dass sich dieser Kongruenzeffekt bei intimen Fragen verstärken würde, können wir ebenfalls nicht bestätigen. Es findet sich lediglich der Trend, dass die Reaktionszeiten bei intimen Fragen geringer sind, welche aber nicht im Zusammenhang mit Prime und Target stehen.

### 4. Diskussion

In unserer Studie untersuchten wir, ob das Wissen, auf eine Frage gelogen zu haben und dessen automatischer Abruf aus dem Gedächtnis vom Inhalt der Frage beeinflusst wird. Ist die kognitive Repräsentation stärker bei Lügen auf intime Fragen im Vergleich zu Lügen auf oberflächliche?

Die Auswertung ergab keine signifikanten Effekte bezüglich unserer Hypothese, gleichwohl wir kürzere Reaktionszeiten bei Präsentation intimer Fragen beobachteten ( $F(1,32) = 6.95$ ,  $p = .013$ ,  $\eta_p^2 = .179$ ). Jedoch konnten im Vergleich zu oberflächlichen Themen keine stärkeren Kongruenzeffekte bei Präsentation intimer Fragen gefunden werden ( $F(1,32) = .021$ ,  $p = .88$ ,  $\eta_p^2 < .001$ ).

Ursachen hierfür finden sich zum Beispiel in der unzureichenden Stichprobengröße, die die Vorgaben der A-Priori Poweranalyse nicht erfüllt.

Da die Proband\*innen das Experiment zu Hause auf ihren eigenen Endgeräten durchführten, können Störvariablen wie Konzentrationsprobleme durch externe Ablenkung nicht ausgeschlossen werden. Der Versuchsablauf verlangte außerdem einige Wechsel zwischen den Plattformen Zoom, JotForm und E-Prime Go, was neben auftretenden technischen Problemen die reibungslose Durchführung des Experiments erschwerte. Es waren keine Laborbedingungen gegeben.

Da die Versuchsleiterin nicht wusste, auf welche Fragen die Versuchsperson log und auf welche sie wahrheitsgemäß antwortete, war es nicht möglich, zu überprüfen, ob die Versuchsperson gemäß der Instruktion auf die Frage reagierte. Hier fehlte ein zusätzlicher Manipulationscheck.

Wieso unsere Hypothese nicht bestätigt werden konnte, lässt sich hier nicht abschließend klären. Fest steht, dass weitere Studien nötig sind, um die sich ergebenden Fragen zu beantworten: Wären signifikante Ergebnisse bei größerer Power möglich? Besteht überhaupt ein Zusammenhang zwischen der Intimität des Frageninhalts und dessen automatischem Abruf im Langzeitgedächtnis? Trotz unzureichender Ergebnisse sind wir zuversichtlich, dass nachfolgende Untersuchungen diese und weitere Fragen abdecken werden.

### 5. Literatur

Koranyi, N., Schreckenbach, F. & Rothermund, K. (2015). The Implicit Cognition of Lying: Knowledge about Having Lied to a Question is Retrieved Automatically. *Social Cognition*, 33, 67-84. doi: 10.1521/soco.2015.33.1.67.

# Effects of Personality on Perceived Stress during the COVID-19 Pandemic

Benjamin Eggs, Friederike Witt, Philipp Wenzel

Supervision: Dr. Jennifer A. Bellingtier

## 1. Introduction

Greater stress is associated with a lower quality of life, poorer mental health, and mortality (Folkman, 2011). Nearly everyone has experienced objective stressors as a result of the Covid-19 pandemic. However, the subjective experience of these stressors varies between individuals. Based on the transactional theory of stress that specifies that perceptions of stress are the result of the interaction of environmental and personal variables (Lazarus & Folkman, 1987), we proposed that individual differences in personality traits would affect stress perceptions during the pandemic. Past research has established that neuroticism is associated with greater stress perceptions, whereas extraversion is associated with more resilient responses (e.g., Penley & Tomaka, 2002). However, due to the social restrictions of the pandemic, we investigated if extraversion, like neuroticism, would also be related to greater perceptions of stress. In line with the Five-Factor-Model (McCrae & Costa, 1987), we divided personality into five major domains composed of smaller facets. In particular, we were interested in the facets of extraversion: sociability, assertiveness, and activity. Following from the transactional theory of stress, we predicted that the severity of corona-related stressors (i.e., environmental variables) would interact with personality (i.e., personal variables) to predict stress appraisals. We hypothesized that individuals higher in neuroticism, extraversion, and especially the extraversion facet of socialability, would report greater perceived stress when corona stressors are high (vs. low) than those lower in these personality traits.

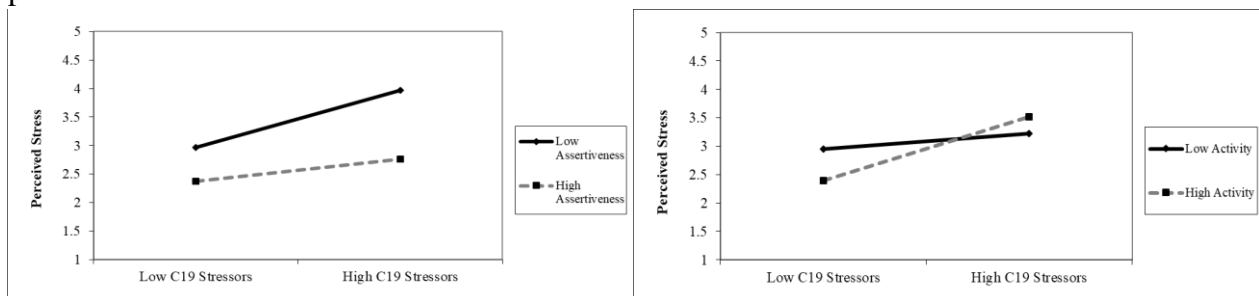
## 2. Methods

In May, 2021, we conducted an online survey including 130 participants between 19 and 40 ( $M = 21.73$ ,  $SD = 3.57$ ) years of age (101 women, 28 men, 1 diverse). An additional 4 participants were excluded for failing attention checks. As incentives for participation we offered a giftcard drawing or experimental participation credits for students at FSU Jena. All participants provided informed consent. Corona stressors were measured using a German-translation of the Coronavirus Impact Scale (Stoddard & Kaufman, 2020) which captured the extent participants' lives have changed due to the pandemic with 12 items. Personality was measured using the German version of the 60-item Big Five Inventory 2 (BFI-2; Danner et al., 2016). Mean scores were created at the domain and facet levels. Perceived stress was measured using a German version of the Perceived Stress Scale (PSS; Cohen et al., 1983), capturing the degree to which various circumstances were subjectively appraised as stressful during the past month of the pandemic. PSS scores were obtained by reversing responses to the four positively-worded items and then averaging across all scale items. To test our hypotheses, we conducted bias-corrected bootstrapped linear regression analyses with perceived stress as the dependent variable being predicted by corona stressors, personality, and the interaction of corona stressors and personality.

## 3. Results

There was a main effect of corona stressors ( $\beta = .49$ ,  $p < .001$ ,  $R^2 = .24$ ) indicating that individuals who reported more stressors related to Covid-19 perceived greater stress in the preceding month. Neuroticism ( $\beta = .51$ ,  $p < .001$ ) was also associated with higher perceived stress and explained an additional 33% of its variance, but it did not interact with corona stressors. When examining the domain of extraversion, there were no main effects or interactions with corona stressors (all  $p > .05$ ).

However, when looking at the extraversion facets, there was a main effect of sociability ( $\beta = .36, p < .001$ ) indicating that individuals who were higher in sociability perceived more stress in the last month than those lower in this facet. In addition, there was a main effect of assertiveness ( $\beta = -.41, p < .001$ ) indicating that individuals who were higher in assertiveness perceived less stress in the last month than those lower in this facet. Activeness was not directly associated with perceived stress ( $\beta = -.12, p = .191$ ). In addition, there were two interaction effects. Covid-19 stressors interacted with assertiveness ( $\beta = -.19, p = .041$ ) and activeness ( $\beta = .27, p = .004$ ). Higher levels of assertiveness and lower levels of activity buffered the relationship between Covid-19 stressors and perceived stress (see figures below). The extraversion facets accounted for an additional 18% of the variance in perceived stress.



#### 4. Discussion

Our findings suggest that personality has both a moderating and direct effect on perceived stress. Individuals high in neuroticism perceived greater stress in the last month regardless of differences in the experience of particular corona stressors. This aligns with past research (e.g., Penley & Tomaka, 2002) and the conceptual definition of neuroticism (McCrae & Costa, 1987) as indicating greater emotional volatility. Although extraversion as a domain was not associated with perceived stress, the individual facets revealed a more complex relationship. Similar to neuroticism, individuals high in sociability perceived the last month as more stressful regardless of experienced corona stressors. We found that higher levels of assertiveness helped reduce stress perceptions, possibly due to their ability to facilitate coping. On the other hand, higher score on the activity facet exacerbated perceptions of stress when corona stressors were high, but were associated with lower perceived stress when corona stressors were low. These individuals, who are high in enthusiasm and energy, may have been able to take advantage of the greater flexibility provided by the pandemic, assuming they did not face too many stressor. Our findings capture a particular sample, German young adults, at a particular moment in time, over a year into a global pandemic. They support the transactional theory of stress (Lazarus & Folkman, 1987) as perceptions of stress were best predicted by a combination of environment and personal factors.

#### 5. References

- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*(4), 385-396. <https://doi.org/10.2307/2136404>
- Danner, D., Rammstedt, B., Bluemke, M., Lechner, C., Berres, S., Knopf, T., Soto, C. J., & John, O. P. (2019). Das Big Five Inventar 2. *Diagnostica, 65*(3), 121–132. <https://doi.org/10.1026/0012-1924/a000218>
- Folkman, S. (2011). *The Oxford handbook of stress, health, and coping* (S. Folkman (Ed.)). Oxford University Press.
- Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality, 1*(3), 141–169. <https://doi.org/10.1002/per.2410010304>
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81–90. <https://doi.org/10.1037/0022-3514.52.1.81>
- Penley, J. A., & Tomaka, J. (2002). Associations among the Big Five, emotional responses and coping with acute stress. *Personality and Individual Differences, 32*(7), 1215–1128. [https://doi.org/10.1016/S0191-8869\(01\)00087-3](https://doi.org/10.1016/S0191-8869(01)00087-3)
- Stoddard, J., & Kaufman, J. (2020). Coronavirus Impact Scale. Retrieved from [https://www.phenxtoolkit.org/toolkit\\_content/PDF/CIS\\_Stoddard.pdf](https://www.phenxtoolkit.org/toolkit_content/PDF/CIS_Stoddard.pdf)

## Effects of punishment on offenders' self-perception

Jessica Heinz, Ann-Kathrin Kohl, Freyja Krull, Paula Molina Sander, Julia Schön

Leitung: Dr. Stefanie Hechler

### 1. Einleitung

Mit einem Fehlverhalten verstoßen Täter:innen gegen moralische Vorstellungen der Gruppe. Bestrafung zielt häufig darauf ab, dass Täter:innen ihre Einstellungen und ihr Verhalten ändern sollen. Die Übernahme moralischer Verantwortung könnte es den Täter:innen erleichtern, sich wieder in die Gruppe zu integrieren. In der aktuellen Studie untersuchen wir, wie Bestrafung die Selbstwahrnehmung von Täter:innen als agierende und moralische Personen beeinflusst. Dies kann durch die wahrgenommene *Agency* (Handlungsfähigkeit) und *Morality* (moralisches Selbstbewusstsein) konzeptualisiert werden. Bestrafung untersuchen wir anhand seiner beiden Komponenten: die symbolische und die behaviorale Komponente. Die symbolische Komponente kommuniziert den Täter:innen, dass ihr Verhalten falsch ist und sie dafür verantwortlich sind. Bei der behavioralen Komponente handelt es sich um eine negative Behandlung der Täter:innen, die alleine keine kommunikative Funktion besitzt (Feinberg, 1965; Sarin et al., 2021). Deshalb vermuten wir, dass die Komponenten der Bestrafung die Selbstwahrnehmung von Täter:innen in Bezug auf ihre *Agency* und *Morality* unterschiedlich beeinflussen. Während die symbolische Komponente die Täter:innen auf ihre *Agency* hinweist, spricht sie auch den Bruch der gemeinsamen Moralvorstellungen, ihre fehlende *Morality*, an. Die behaviorale Komponente dagegen entzieht Täter:innen ihrer *Agency*, zeigt aber nicht, dass sie unmoralisch gehandelt haben. Die behaviorale Komponente könnte außerdem die Wirkung der symbolischen Komponente unterstützen, im Sinne von „Wer nicht hören will, muss fühlen“. Wir nehmen an, dass die wahrgenommene *Morality* geringer ist, wenn die Bestrafung eine symbolische Komponente beinhaltet, als wenn sie keine symbolische Komponente beinhaltet (H1). Die *Agency*-Wahrnehmung der Täter:innen hingegen sollte geringer sein, wenn die Bestrafung eine behaviorale Komponente beinhaltet, als wenn sie keine behaviorale Komponente beinhaltet (H2a). Zudem könnte eine Bestrafung mit behavioraler Komponente zu höherer Wahrnehmung von *Agency* führen, wenn sie zusätzlich die symbolische Komponente beinhaltet, als wenn sie diese nicht beinhaltet (H2b).

### 2. Methode

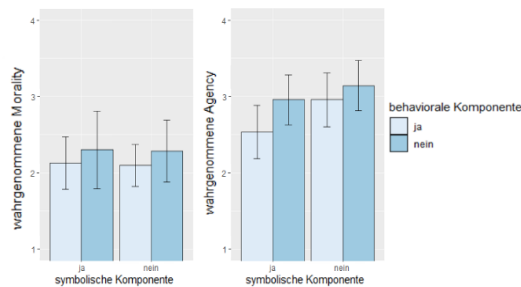
104 Proband:innen (71 Frauen, 21 Männer, 3 Divers; Alter: 18 - 61,  $M = 26,6$ ,  $SD = 10,48$ ; 71% Studierende) haben an der Studie teilgenommen. Von den ursprünglichen 104 Teilnehmenden wurden 9 aufgrund von Aufmerksamkeitschecks oder unvollständigen Daten ausgeschlossen. Eine Analyse mit G\*Power ( $f = .20$ ,  $\alpha = .05$ ,  $\beta = .80$ ) ergab eine empfohlene Stichprobengröße von 199 Teilnehmenden. Diese Anzahl an Teilnehmenden haben wir nicht erreicht. Die Studie hat ein 2 (symbolische Komponente: ja/nein) x 2 (behaviorale Komponente: ja/nein) between-participants Design mit zwei Faktoren. Die Teilnehmenden sollten sich in ein Szenario hineinversetzen, in dem sie sich anderen gegenüber falsch verhalten. Die Reaktion ihres Gegenübers fällt dabei in jedem Szenario auf vier unterschiedliche Weisen aus. Um die Selbstwahrnehmung der Täter:innen zu erheben, schätzten diese sich selbst auf einer Skala mit 22 Adjektiven in zufälliger Reihenfolge ein (Bilewicz et al., 2019). Diese Skala ging von 1 = "überhaupt nicht" bis 6 = "sehr viel". Zehn Items bezogen sich dabei auf die Selbstwahrnehmung der *Morality* und neun Items auf die Selbstwahrnehmung der *Agency*. Drei weitere Adjektive dienten als Füllitems. Um die Daten auszuwerten, führten wir zwei 2x2 Varianzanalysen mit den Faktoren symbolische Komponente und behaviorale Komponente auf wahrgenommene *Agency* und *Morality* durch. Wir haben weitere explorative Variablen erhoben, deren Ergebnisse wir an dieser Stelle nicht berichten. Die Studie wurde präregistriert: <https://aspredicted.org/blind.php?x=mi2635>.



### 3. Ergebnisse

Die Ergebnisse zum Manipulationscheck zeigten, dass unsere Manipulation erfolgreich war - die Täter:innen waren sich ihres Fehlverhaltens bewusst. Eine Varianzanalyse zur Hypothesentestung (H1) zeigte keine signifikanten Effekte auf die Wahrnehmung von *Morality*. Weder die symbolische Komponente,  $F(1,91) = .015, p = .903, \eta^2 < .001$ , noch die behaviorale Komponente,  $F(1,91) = .891, p = .348, \eta^2 = .009$ , hatten einen signifikanten Haupteffekt auf *Morality*. Der Interaktionseffekt auf *Morality* ist ebenfalls nicht signifikant,  $F(1,91) = .002, p = .965, \eta^2 < .001$ . Zudem gab es keine

Abbildung 1:  
Mittelwerte der wahrgenommenen *Agency* und *Morality* in allen Bedingungen. KI: 95%



Effekte der Komponenten von Bestrafung auf die wahrgenommene *Agency* der Proband:innen (H2). Es gab weder einen signifikanten Effekt der symbolischen Bestrafung,  $F(1,91) = 3.390, p = .069, \eta^2 = .036$ , noch der behavioralen Komponente,  $F(1,91) = 3.390, p = .069, \eta^2 = .036$ , oder einen Interaktionseffekt auf *Agency*,  $F(1,91) = .524, p = .471, \eta^2 = .006$  sind ebenfalls nicht signifikant. Es zeigte sich eine Tendenz, dass die *Agency* sowohl durch die behaviorale als auch symbolische Komponente geringer ist, als wenn diese nicht vorhanden sind, als wenn sie vorhanden sind (siehe Abbildung 1).

### 4. Diskussion

Nach einem Fehlverhalten könnte eine Selbsteinschätzung der Täter:innen als agentisch und unmoralisch ihr zukünftiges Verhalten beeinflussen. Die vorliegenden Ergebnisse zeigen entgegen unserer Hypothesen keine signifikanten Effekte einer Bestrafung auf die Selbstwahrnehmung von Täter:innen. Es deuten sich jedoch Effekte auf die wahrgenommene *Agency* der Täter:innen an. Wird den Täter:innen gesagt, dass ihr Verhalten falsch war (symbolische Komponente) oder hatte die Situation negative Konsequenzen (behaviorale Komponente) scheint dies zu einer niedrigeren wahrgenommenen *Agency* zu führen. Beide Komponenten wirken unabhängig voneinander darauf, wie agentisch sich die Täter:innen fühlen. Der Einfluss der behavioralen Komponente auf die *Agency* könnte dadurch erklärt werden, dass die Handlungsfähigkeit der Täter:innen durch die negative Konsequenz der Bestrafung oder die soziale Kontrolle durch andere eingeschränkt wird. Durch das Szenario stellten die Täter:innen sich ein Verhalten vor, für das sie sich nicht aktiv entscheiden konnten. Wurden sie von der bestrafenden Person darauf hingewiesen, dass sie sich falsch verhalten haben, könnte ihnen die fehlende Kontrolle über die Tat auffallen. Dies könnte zu einem niedrigeren Gefühl von *Agency* geführt haben. Die Werte von *Morality* liegen unterhalb der Skalenmitte. Eine mögliche Interpretation ist, dass sich das moralische Bewusstsein der Versuchspersonen durch eine vorgestellte Tat, die gegen die eigenen Moralvorstellungen spricht, negativ auf die wahrgenommene *Morality* auswirkt - unabhängig von der Reaktion des Gegenübers. Letztendlich wurden einige Effekte möglicherweise aufgrund der zu geringen Stichprobengröße nicht aufgedeckt. In künftiger Forschung könnte die Wirkung der verschiedenen Komponenten von Bestrafung auf die wahrgenommene *Agency* der Täter:innen mit ausreichender Stichprobengröße untersucht werden.

### 5. Literatur

- Bilewicz, M., Stefaniak, A., Barth, M., Witkowska, M., & Fritsche, I. (2019). The Role of Control Motivation in Germans' and Poles' Interest in History. *Social Psychological Bulletin*, 14(2), 1-31. <https://doi.org/10.32872/spb.v14i2.33399>
- Feinberg, J. (1965). The Expressive Function of Punishment. *Monist*, 49(3), 397-423. <https://doi.org/10.5840/monist196549326>
- Sarin, A., Ho, M. K., Martin, J. W. & Cushman, F. A. (2021). Punishment is Organized around Principles of Communicative Inference. *Cognition*, 208 (Issue), 104544. <https://doi.org/10.1016/j.cognition.2020.104544>

# Observational SR-Binding – „Mein, Dein oder Unser?“ – wie beeinflusst unser Mindset das Beobachtungslernen

Johanna Dorra, Joanna Hein, Aurelia Holfeld, Julian Huber, Greta Sippel, Teresa Wiesner

Leitung: Dr. Carina G. Giesen

## 1. Einleitung

Menschen nehmen ihre Umwelt unterschiedlich wahr, verarbeiten Informationen und beobachten andere. Kann allein durch das Beobachten von Reaktionen eines Anderen, diese kognitiv an einen Stimulus gebunden werden? Der *observational stimulus-response (SR) Binding* Effekt tritt auf, wenn ein Proband die beobachtete Reaktion einer anderen Person auf einen Stimulus in seinem Gedächtnis speichert. Wird anschließend der gleiche Stimulus erneut gezeigt, werden die Erinnerungen abgerufen und beeinflussen die Reaktion. Faktoren, die bei der Entstehung eine Rolle spielen, sind zum Beispiel die Interdependenz der Personen (Giesen et al., 2014). Eine bisher in diesem Zusammenhang noch ungeklärte und spannende Frage bleibt, welchen Einfluss die Induktion unterschiedlicher Mindsets hat. Wird ein kollektivistisches Mindset geprimed (Fokus auf „wir“, „uns“, etc.) führt das zu einer stärkeren Integration von anderen in das eigene Selbstkonzept (Colzato et al., 2012) als ein individualistisches Mindset (Fokus auf „ich“, „mein“, etc.) was sich in den Reaktionszeiten zeigt. Inwiefern verschiedene Mindsets durch Fokusinduktionen observational SR-Binding beeinflussen, ist bis jetzt unklar und soll in dieser Studie untersucht werden. Davon ausgehend, lautet unsere Hypothese, dass der observational SR-Binding Effekt nur auftreten sollte, wenn ein kollektivistisches Mindset induziert wird, nicht jedoch, wenn ein individualistisches Mindset induziert wird.

## 2. Methode

*Probanden.* An der Studie nahmen 11 Versuchspersonen (VP; 7 Frauen, 4 Männer,  $M = 22.3$  Jahre,  $SD = 2.5$ ) teil. Die VP wurden für die Teilnahme mit Versuchspersonenstunden entlohnt.

*Material und Ablauf.* Zuerst fand eine Präregistrierung des Online-Experiments statt (<https://aspredicted.org/6hf6r.pdf>). Unser Experiment bestand aus Manipulationsaufgaben und Reaktionszeitaufgaben, welche von den VP durchgeführt wurden. Zusätzlich wurde den VP vermittelt, dass sie die Aufgabe gemeinsam mit einer Hilfswissenschaftlerin als Interaktionspartnerin durchführen würden, welche online mit ihnen verbunden sei. Bei den Manipulationsaufgaben wurde entweder ein individualistisches oder ein kollektivistisches Mindset induziert, indem die VP in einem Text individualistisch oder kollektivistisch assoziierte Pronomen zählen mussten. Alle VP durchliefen beide Mindset-Bedingungen (Reihenfolge ausbalanciert). Bei den Reaktionszeitaufgaben handelte es sich um ein observational SR-Binding Paradigma; es gab *Prime*- und *Probe*-Durchgänge. In den *Prime*-Durchgängen sah die VP ein Wort in weißer Schriftfarbe. Den VP wurde suggeriert, die angebliche Partnerin sehe dasselbe Wort in grüner oder roter Schriftfarbe und kategorisiere die Wortfarbe durch Tastendruck. In diesen *Prime*-Durchgängen sah die VP lediglich die Reaktion ihrer Interaktionspartnerin (roter oder grüner Knopf auf dem Bildschirm leuchtet auf). In den folgenden *Probe*-Durchgängen sah nun die VP ein Wort entweder in grüner oder in roter Schriftfarbe. Es konnte dasselbe Wort wie im *Prime*-Durchgang oder ein anderes sein (Faktor Stimulusrelation). Das Wort erschien entweder in der gleichen (rot-rot, grün-grün) oder in entgegengesetzter (rot-grün, grün-rot) Farbe (Faktor Reaktionskompatibilität). Die VP musste das Wort per Tastendruck so schnell wie möglich farblich kategorisieren (Taste A oder L), die Reaktion konnte kompatibel oder inkompatibel mit der beobachteten Reaktion sein. Nach 25% der *Probe*-Durchgänge erfolgte ein Gedächtnistest, welcher die Aufmerksamkeit der Probanden während der *Prime*-Durchgänge testete.

Jede VP absolvierte 160 *Prime-Probe*-Sequenzen, je 80 pro Fokusindikation (nach jeweils 40 Sequenzen wurde erneut manipuliert).

Die Reaktionszeit der VP in den *Probe*-Durchgängen wurde gemessen. Nach je 40 Sequenzen folgte auf die Reaktionszeitaufgabe eine Selbsteinschätzung der Versuchspersonen auf der „Inclusion of Others in the Self (IOS)“, welche als Manipulationscheck diente.

*Design.* Die Studie war nach einem 2 (Stimulusrelation: Wortwiederholung [Wdh.] vs. Wortwechsel [Wec.])  $\times$  2 (Reaktionskompatibilität: kompatibel [K] vs. inkompatibel [IK])  $\times$  2 (Fokus: individualistisch vs. kollektivistisch) within-Design aufgebaut.

### 3. Ergebnisse

*Analyse der Reaktionszeiten.* Vor der Analyse wurden *Probe*-Durchgänge mit fehlerhaften Reaktionen (2.4 %), nachfolgend falsch beantworteten Gedächtnistests (1.7 %) und Reaktionszeit-Ausreißerwerten (3.6 %) ausgeschlossen. Letztere wurden mithilfe individuenbasierter Tukeygrenzen ermittelt. Eine 2 $\times$ 2 $\times$ 2 Varianzanalyse (ANOVA) mit Messwiederholung ergab einen signifikanten Haupteffekt der Stimulusrelation,  $F(1,10) = 905.43, p < .001, \eta_p^2 = .99$ , welcher zeigt, dass Reaktionszeiten bei Wortwiederholungen im Mittel kürzer waren (489 ms) als solche bei Wortwechseln (492 ms). Ebenso konnte eine signifikante Zweifachinteraktion von Stimulusrelation und Reaktionskompatibilität ( $F[1,10] = 8.86, p = .014, \eta_p^2 = .47$ ; S $\times$ R Interaktionseffekt = 24.5 ms) gefunden werden, die den observational SR-Binding Effekt anzeigt. Die erwünschte Dreifachinteraktion dieser beiden mit dem Fokus wurde jedoch nicht signifikant ( $F[1,10] = .33, p = .579, \eta_p^2 = .03$ ; S $\times$ R<sub>individualistisch</sub> = 19.7 ms; S $\times$ R<sub>kollektivistisch</sub> = 29.4 ms).

*Manipulationscheck.* Ein *t*-Test für abhängige Stichproben ergab keine signifikanten Unterschiede auf der IOS-Skala zwischen individualistischem und kollektivistischem Fokus ( $M_{\text{individualistisch}} = 2.7$ ;  $M_{\text{kollektivistisch}} = 2.5$ ;  $t[10] = .585, p = .572$ ). Dementsprechend kann davon ausgegangen werden, dass die Manipulation nicht funktioniert hat.

### 4. Diskussion

Unsere Hypothese, dass der observational SR-Binding Effekt nur bei kollektivistischem, aber nicht bei individualistischem Mindset auftritt, konnte nicht bestätigt werden. Es spiegelt sich in den Daten sogar ein gegenteiliges Muster wider, wobei sich der observational SR-Binding Effekt bei der individualistischen Mindset-Induktion stärker abbildet als bei der kollektivistischen Mindset-Induktion.

*Interpretation und Probleme.* Eines unserer größten Probleme, die uns die Interpretation dieser Ergebnisse erschwert, war die Rekrutierung der Probanden. Während wir laut *a-priori* Poweranalyse:  $n=41$  Probanden erheben wollten, nahmen nur  $n=11$  Probanden teil. Aufgrund der daraus folgenden geringen (*post-hoc*) Power von  $1-\beta = 0.341$  ist keine plausible Interpretation der Ergebnisse möglich. Außerdem erschwerte die Suggestion der Interaktion mit einer Hilfwissenschaftlerin die Rekrutierung, da deshalb die Probanden nur in einem festgelegten Zeitraum teilnehmen konnten. Leider gab es auch einige Probanden, die das Experiment nicht im gegebenen Zeitraum durchführten, weshalb hier auch die Manipulation der Interaktion nicht funktioniert haben kann, was eine Erklärung für die widersprüchlichen Ergebnisse zu unserer Hypothese sein könnte. Außerdem fällt auf, dass der Manipulationscheck für die Manipulation der Fokusinduktion nicht wie gewünscht ausfällt, was darauf hindeutet, dass auch diese Manipulation nicht gelungen ist.

### 5. Literatur

- Colzato, L.S., De Bruhn, E.R.A. & Hommel, B. (2012). Up to “me” or up to “us”? The impact of self-construal priming on cognitive self-other integration. *Front. Psychology* 3:341. doi: 10.3389/fpsyg.2012.00341
- Giesen, C., Hermann, J., & Rothermund, K. (2014). Copying Competitors? Interdependency Modulates Stimulus-Based Retrieval of Observed Responses. *Journal of Experimental Psychology: Human Perception & Performance*, 40(5), 1978-1991. doi: 10.1037/a0037614

## **Let's face it! Wie die unidirektionale Manipulation konfigurationaler Information das Lernen von Gesichtern beeinträchtigt.**

**Svenja Andrade Breslauer, Maria Arabadzheva, Lucy Braxmeier, Michelle Gutteck, Verena Kolb, Eva Reiter, Saida Saibou, & Lisa-Marie Schröder**

Leitung: Dr. Jürgen Kaufmann und B.Sc. Pauline Gürtler

### 1. Einleitung

Im Alltag begegnen uns viele Menschen, manche sehen durchschnittlich aus, andere heben sich von der Masse ab. Solche „distinkten“ Gesichter werden normalerweise besser und schneller gelernt (Schulz, Kaufmann, Schweinberger & Kurt, 2012). Ein Nachteil findet sich aber für other-„race“ Gesichter, die meist schlechter gelernt werden. Dieser other-„race“ Effekt konnte bisher trotz intensiver Forschung nicht hinreichend erklärt werden, oft wird aber von einem qualitativen Verarbeitungsunterschied ausgegangen (Rossion & Michel, 2011). Im Gegensatz dazu vermuteten wir ein generelles Prinzip beim Lernen von Gesichtern als Ursache. Hierbei spielt, unabhängig von der „race“ eine irreführende Distinktheit die entscheidende Rolle: Beim Lernen von Gesichtern benutzt das visuelle System die Abweichung eines Gesichts von der Norm, um es im Gedächtnis abzuspeichern. Da bei other-„race“ Gesichtern die distinkten Merkmale aber vorwiegend die „ethnischen“ Gruppenmerkmale betreffen und dadurch eher von individuellen Merkmalen ablenken, könnte es hier aufgrund der erhöhten Distinktheit sogar zu Nachteilen kommen. Dies führte zu unserer Frage, ob man für same-„race“ Gesichter einen vergleichbaren Effekt herstellen kann, wenn man deren Distinktheit immer auf die gleiche Art und Weise steigert. In einem Lernexperiment präsentierten wir dazu europäische Gesichter, deren Distinktheit wir durch unidirektionale Manipulation der Konfiguration der internen Features erhöhten. Die Lernleistungen für diese Gesichter verglichen wir mit denen für asiatische und originale europäische Gesichter. Wir erwarteten Nachteile sowohl für asiatische, als auch für unidirektional manipulierte europäische Gesichter.

### 2. Methode

Das Online-Experiment wurde mit PsyToolkit Version 3.3.2 (Stoet, 2010; 2017) erstellt und durchgeführt. Bisher nahmen 28 gültige VPn am Lernexperiment teil (basierend auf einer Poweranalyse planten wir die Erhebung von 50 VPn). Das Experiment bestand aus einer Lernphase mit zwei Durchgängen, einer Übungsphase und einer Testphase. In beiden Lerndurchgängen wurden den VPn in randomisierter Reihenfolge jeweils 26 europäische, europäisch manipulierte und asiatische Gesichter gezeigt, die sie sich einprägen und hinsichtlich ihrer Distinktheit beurteilen sollten. Die Manipulation der europäischen Gesichter erfolgte durch eine Veränderung des Augenabstandes und des Abstandes von Mund und Nase. In der darauffolgenden Testphase wurden 78 gelernte und ebenso viele neue Gesichter gezeigt und es sollte durch Tastendruck möglichst schnell entschieden werden, welche Gesichter in der Lernphase gezeigt worden waren und welche nicht. Im Anschluss wurde der Cambridge Face Memory Test (CFMT, Duchaine & Nakayama, 2006) zur Erfassung der individuellen Gesichterlernfähigkeiten durchgeführt.

### 3. Ergebnisse

Die Analyse der Daten erfolgte mit R-Studio (Version 1.4.1103). Für die Distinktheitsratings der Lernphase wurde eine 2x3-ANOVA mit Messwiederholung auf den Faktoren Lerndurchgang (1 vs. 2) und Gesichtertyp (europäisch vs. europ. manipuliert vs. asiatisch) gerechnet. Sie zeigte einen Haupteffekt für den Faktor Gesichtertyp  $F(2,54) = 26.287, p < .001, \eta_p^2 = .493$ . Eine Überprüfung durch t-Tests ergab, dass europäisch-manipulierte Gesichter distinkter wahrgenommen wurden als europäisch-originale,  $t(27) = -7.597, p < .001$  und asiatische Gesichter,  $t(27) = -3.609, p = 0.001$ .

Zudem wurden asiatische Gesichter wie erwartet distinkter eingeschätzt als europäisch-originale Gesichter,  $t(27) = 3.580, p = 0.001$ . Für die Testphase wurde für die Antwortgenauigkeiten eine  $2 \times 3$ -ANOVA mit Messwiederholung auf den Faktoren Bekanntheit (gelernt vs. neu) und Gesichtertyp durchgeführt. Es zeigten sich Haupteffekte für Bekanntheit und Gesichtertyp,  $F(1,27) = 5.199, p = .031, \eta_p^2 = .161$ ;  $F(2,54) = 4.299, p = .019, \eta_p^2 = .137$ , und eine zweifach-Interaktion,  $F(2,54) = 5.413, p = .007, \eta_p^2 = .167$ . Zur Weiterverfolgung wurden separate ANOVAs mit dem Faktor Gesichtertyp für gelernte und neue Gesichter durchgeführt. Für neue Gesichter ergab sich ein Effekt für den Faktor Gesichtertyp  $F(2,54) = 11.483, p < .001, \eta_p^2 = .298$ . Anschließende t-Tests zeigten, dass europäisch-originale Gesichter häufiger korrekt zurückgewiesen wurden als europäisch-manipulierte,  $t(27) = 2.990, p = 0.006$ , und asiatische Gesichter,  $t(27) = -5.441, p < .001$ . Die ANOVA für gelernte Gesichter ergab keinen Haupteffekt für den Faktor Gesichtertyp. Eine ANOVA für die Reaktionszeiten der Testphase ergab Haupteffekte für Bekanntheit,  $F(1,27) = 6.264, p = .019, \eta_p^2 = .188$ , und Gesichtertyp,  $F(2,54) = 5.652, p = .012, \eta_p^2 = .173$ , sowie eine zweifach-Interaktion,  $F(2,54) = 3.706, p = .031, \eta_p^2 = .121$ . Auch für die RTs waren die Effekte des Gesichtertyps für die neuen Gesichter ausgeprägter, als für gelernte. Es zeigte sich im Vergleich zu europäischen Originalgesichtern ein Nachteil für europäische manipulierte und asiatische Gesichter, welche sich wiederum nicht voneinander unterschieden,  $t(27) = 0.091, p = 0.928$ . Eine ANOVA für den Sensitivitätsparameter  $d'$  ergab einen signifikanten Effekt für den Faktor Gesichtertyp,  $F(2,54) = 7.088, p = .002, \eta_p^2 = .208$ . Daran anschließende t-Tests zeigten höhere  $d'$  Werte für europäische originale Gesichter als für manipulierte,  $t(27) = 3.127, p = 0.004$ , und für asiatische Gesichter,  $t(27) = -3.902, p < 0.001$ . Letztere unterschieden sich nicht voneinander. Eine ANOVA für das Maß des Antwortbias (Kriterium C) ergab einen Effekt für den Faktor Gesichtertyp  $F(2,54) = 7.470, p = .001, \eta_p^2 = .217$ , der von anschließenden t-Tests auf konservativere Antworten für europäische originale und europäisch-manipulierte im Vergleich zu asiatischen Gesichtern zurückgeführt wurde,  $t(27) = -4.197, p < .001$  und  $t(27) = -2.37, p = 0.025$ .

#### 4. Diskussion

Wir untersuchten, ob der other-„race“-Effekt durch eine unidirektionale und dadurch irreführende Distinktheit erklärt werden kann. Neben einer Replikation des other-„race“-Effekts ergab die Studie auch den Befund geringerer Antwortgenauigkeiten für die manipulierten europäischen same-„race“-Gesichter, als für die weniger distinkten Originale. Vergleichbare Ergebnisse zeigten sich für das Sensitivitätsmaß  $d'$  und für die Reaktionszeiten. Offensichtlich konnte also durch die Manipulationen eine Art other-„race“-Effekt induziert werden. Unsere vorliegenden Ergebnisse deuten somit stark darauf hin, dass der other-„race“-Effekt gar nicht spezifisch für other-„race“-Gesichter ist, sondern zumindest teilweise durch eine unidirektionale und dadurch irreführende Distinktheit erklärt werden kann. Man sollte deswegen vielleicht besser von einem generellen other-appearance Effekt sprechen, bei dem die unidirektionale Distinktheit von Merkmalen zum schlechteren Gesichterlernen führt.

#### 5. Literatur

- Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44*(4), 576–585. doi.org/10.1016/j.neuropsychologia.2005.07.001
- Rossion, B., & Michel, C. (2011). An Experience-Based Holistic Account of the Other-Race Face Effect. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford Handbook of Face Perception* (pp. 215-243). Oxford New York: Oxford University Press. doi:10.1093/oxfordhb/9780199559053.013.0012
- Schulz, C., Kaufmann, J.M., Schweinberger, S.R., & Kurt, A. (2012). Faces forming traces: Neurophysiological correlates of learning naturally distinctive and caricatured faces. *NeuroImage, 63*(1), 491-500. doi.org/10.1016/j.neuroimage.2012.06.080
- Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096-1104. doi:10.3758/BRM.42.4.1096
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*(1), 24-31. doi.org/10.1177/0098628316677643

# Die Schwierigkeit und die Vorhersagekraft im IAT

Luisa Drechsel, Natalie Hülßner, Annelie Rodestock, Lena Wetzel, Elena Zuschlag

Leitung: M.Sc. Merlin Urban

## 1. Einleitung

Die prädiktive Validität des impliziten Assoziationstests (IAT) ist gering und Untersuchungen von Moderatoren zur Erklärung der geringen prädiktiven Validität liefern nur wenig vielversprechende Ergebnisse (Greenwald, Poehlman, Uhlmann & Banaji, 2009; Kurdi et al., 2019). Wir haben in dieser Studie einen bisher nicht betrachteten Moderator, die Schwierigkeit aus der klassischen Testtheorie (KTT), untersucht, um das Problem der geringen prädiktiven Validität aus einer neuen Perspektive zu betrachten. Aus der KTT ist bekannt, dass mittelschwere Items/Tests bessere Vorhersagen treffen als Items/Tests mit extremen Schwierigkeiten, da sie mehr Varianz haben (Bühner, 2011). Ein IAT ist leicht, wenn der IAT-Effekt positiv ist, schwierig wenn der IAT-Effekt negativ ist und mittelschwer, wenn der IAT-Effekt bei null liegt. Allerdings nur dann, wenn der Block mit der Tastenzuordnung relevante Targetkategorie und positiv von dem Block mit der Tastenzuordnung relevante Targetkategorie und negativ abgezogen wird. Dann wird im Falle eines positiven IAT-Effekts im Sinne der relevanten Targetkategorie geantwortet, sie wird positiver bewertet, und im Falle eines negativen IAT-Effekts nicht im Sinne der relevanten Targetkategorie geantwortet, sie wird negativer bewertet. Es ist anzunehmen, dass mehr Personen im Sinne der relevanten Targetkategorie antworten, je negativer die Referenzkategorie ist und mehr Personen nicht im Sinne der relevanten Targetkategorie antworten, je positiver die Referenzkategorie ist. Folglich nehmen wir an, dass die Schwierigkeit über die Valenz der Referenzkategorie manipulierbar ist und dass die prädiktive Validität eines IATs steigt, je näher der IAT einer mittleren Schwierigkeit ist.

## 2. Methode

*Stichprobe.* Insgesamt gingen 73 Proband:innen (weiblich = 57, männlich = 15, divers = 1, Altersdurchschnitt = 23.1 Jahre, SD = 7.91) in die Analysen ein.

*Ablauf.* Die Studie wurde Online als Computerexperiment durchgeführt. Die Proband:innen bearbeiteten drei IATs und mehrere Fragebogenmaße, u.a. zur Messung der Verhaltenskriterien.

*Design.* Die unabhängige Variable unsere Studie war die Schwierigkeit der IATs mit drei Stufen (sehr leicht = sehr negative Referenzkategorie „Umweltschädigung“; leicht = negative Referenzkategorie „Konsumgesellschaft“; mittelschwer = positive Referenzkategorie „Wirtschaftswachstum“). Die abhängige Variable waren die Korrelationen zwischen den drei IATs und zwei Verhaltenskriterien. Die Reihenfolge der IATs sowie der IAT-Blöcke wurde nicht randomisiert, da wir im Wesentlichen an interindividuellen Unterschieden interessiert waren.

*Messinstrumente.* Drei IATs mit der relevanten Targetkategorie „Umweltschutz“ wurden entwickelt, die wir im Folgenden entsprechend ihrer Referenzkategorien benennen werden: Umweltschädigung-IAT ( $\alpha = .53$ ), Konsumgesellschaft-IAT ( $\alpha = .41$ ) und Wirtschaftswachstum-IAT ( $\alpha = .52$ ). Alle drei IATs hatten gemeinsam, dass die Stimuli der Targetkategorien Bilder und die Stimuli der Attributkategorien Wörter waren. Nach einer Übung der Kategorisierung der Attribute durchlief jede Person nacheinander die zwei kombinierten Testblöcke eines jeden IATs. Insgesamt gab es drei Testdurchgänge pro IAT. Das Umweltverhalten wurde zum einen mit der Kurzversion B der „General Ecological Behavior Scale“ (GEB; Kaiser, Merten & Wetzel, 2018;  $\alpha = .76$ ), und zum anderen mit der „Pro-Environmental Behavior Scale“ (PEBS; Markle, 2013;  $\alpha = .74$ ) erfasst. Letztere haben wir modifiziert, sodass das Verhalten Bezug auf den letzten Monat hat.

## 3. Ergebnisse

Wie angenommen ist der Umweltschädigung-IAT der leichteste IAT und der Konsumgesellschaft-IAT etwas schwieriger aber immer noch leicht. Der Wirtschaftswachstum-IAT ist jedoch nicht mittelschwer, sondern liegt mit seiner Schwierigkeit zwischen den beiden anderen IATs (vgl. mittlere D-Scores in Tabelle 1). Eine einfaktorielle ANOVA mit Messwiederholung ergab, dass die Unterschiede zwischen den mittleren D-Scores signifikant waren,  $F(2,146)=54.09$ ,  $p < .001$ . Post-hoc Tests nach Benjamini-Hochberg zeigten, dass sich alle IAT-Effekte signifikant voneinander unterschieden, alle  $ps < .05$ . Wir konnten also die Schwierigkeit der IATs manipulieren, wenn auch nicht genau in die a priori angenommene Richtung. Weiterhin ergaben Z-Tests, dass die Unterschiede zwischen dem Konsumgesellschaft- und dem Umweltschädigung-IAT in ihren Korrelationen mit der PEBS-Skala signifikant,  $p = .05$ , und mit der GEB-Skala marginal signifikant,  $p = .08$ , waren. Der IAT, der mehr in Richtung mittlere Schwierigkeit geht, scheint also tendenziell wie angenommen höhere Korrelationen zu haben als der IAT mit extremer Schwierigkeit. Die Korrelationen des Wirtschaftswachstum-IATs lagen in ihrer Höhe wie zu erwarten zwischen denen der anderen beiden IATs, unabhängig von dem betrachteten Verhaltenskriterium. Allerdings wurde nur der Unterschied zur Korrelation des Umweltschädigung-IAT mit der PEBS-Skala signifikant,  $p = .03$ .

Tabelle 1: Deskriptive Statistiken der IATs und ihre Zusammenhänge mit den Verhaltenskriterien

	Mittelwert	Varianz	GEB	PEBS
Umweltschädigung-IAT	0.64	0.07	.21	.05
Konsumgesellschaft-IAT	0.35	0.07	.37*	.25*
Wirtschaftswachstum-IAT	0.41	0.08	.28*	.24*

Anmerkung. \*  $p < .05$

#### 4. Diskussion

Die Manipulation der Schwierigkeit hat funktioniert, wenn auch nur teilweise in die a priori angenommene Richtung. Außerdem sind Reihenfolgeeffekte nicht auszuschließen. Folglich lässt sich die Manipulation der Schwierigkeit nicht eindeutig auf die Valenz der Referenzkategorien zurückführen. Weiterhin zeigt sich, dass der Konsumgesellschaft-IAT, welcher am ehesten in Richtung einer mittleren Schwierigkeit geht, rein deskriptiv die höchsten Korrelationen hat, wenn auch nur teilweise signifikant von den anderen Korrelationen verschieden. Allerdings hat der Konsumgesellschaft-IAT nicht die höchste Varianz. Die von uns angenommene Erklärung für die bessere prädiktive Validität mittelschwerer IATs scheint also nicht korrekt zu sein. Eine Alternativerklärung könnte sein, dass bei mittelschweren IATs Rekodierung weniger stattfindet und mittelschwere IATs dadurch mehr konstruktrelevante Varianz messen (vgl. Rothermund, Teige-Mocigemba, Gast und Wentura, 2009).

#### 5. Literatur

- Bühner, M. (2011). Einführung in die Test- und Fragebogenkonstruktion (3. Auflage). München: Pearson
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97(1), 17–41.
- Kaiser, F. G., Merten, M., & Wetzel, E. (2018). How do we know we are measuring environmental attitude? Specific objectivity as the formal validation criterion for measures of latent attributes. *Journal of Environmental Psychology*, 55, 139–146.
- Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., Tomezsko, D., Greenwald, A. G., & Banaji, M. R. (2019). Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *American Psychologist*, 74(5), 569–586.
- Markle, G. (2013). Pro-Environmental Behavior: Does It Matter How It's Measured? Development and Validation of the Pro-Environmental Behavior Scale (PEBS). *Human Ecology*, 41(6), 905–914.
- Rothermund, K., Teige-Mocigemba, S., Gast, A.; Wentura, D. (2009): Minimizing the Influence of Recoding in the Implicit Association Test: The Recoding-free Implicit Association Test (IAT-RF). *The Quarterly Journal of Experimental Psychology*, 62(1), 84–98.

# Die magische evaluative Konditionierung - Produzieren asymmetrische Relationspaare asymmetrische EC-Effekte?

Laura Dienemann, Lara Fangradt, Hannah Hoffmann, Viktoria Kämpfe

Leitung: M.Sc. Karoline Bading

## 1. Einleitung

Ein aktuell viel beforschter Ansatz innerhalb der evaluativen Konditionierung ist das relationale EC-Paradigma. Dabei werden ein neutraler Stimulus und ein US zusammen mit Relationswörtern gepaart präsentiert. Die Bewertung des CS bewegt sich in Richtung des US in Abhängigkeit von ihrer Relation. Folglich handelt es sich hierbei um ein 2x2 Forschungsdesign (US-Valenz: positiv vs. negativ x CS-US-Relation: assimilativ vs. kontrastiv). Assimilative Relationen sind CS-US-Paarungen, durch die sich die Bewertung des CS in Richtung der US-Valenz verschiebt. Bei kontrastiven Relationen entfernt sich die Bewertung des CS von der Valenz des US.

Das typische Befundmuster der evaluativen Konditionierung ist eine disordinale Interaktion von US-Valenz und Relation (Zanon et al., 2014). In vorherigen Studien wurde vermehrt eine Asymmetrie in den EC Effekten gefunden: kontrastive Relationen erzeugten schwächere EC Effekte als assimilative Relationen (Hu et al., 2017). Statistisch drückt sich dies in einem signifikanten Haupteffekt von US Valenz aus. Nach dem Zwei-Prozess-Modell ließe sich dieses Muster durch ein entgegengesetztes Wirken von Assoziation *und* Proposition in der kontrastiven Bedingung erklären (Hu et al., 2017). Unserer Vermutung entsprechend, könnte die Asymmetrie der EC Effekte jedoch ebenso durch die Wortpaare selbst und somit dem Ein-Prozess-Modell folgend, ausschließlich durch propositionale Prozesse entstanden sein. Bei asymmetrischen Wortpaaren erzeugt die jeweilige kontrastive Relation deutlich schwächere Bewertungen als ihr assimilatives Gegenstück. Bei symmetrischen Wortpaaren erzeugen beide Relationen folglich vergleichbar starke Bewertungen. Dementsprechend lautete die zentrale Frage für unsere Studie: Induzieren Wortpaare mit stärkerer Asymmetrie auch dementsprechend starke Asymmetrien in den EC-Effekten? Dies könnte empirisch gestützt werden, sofern der Haupteffekt von US-Valenz durch die Symmetrie/Asymmetrie des jeweils verwendeten Relationspaares moderiert wird.

## 2. Methode

Die Studie ist im 2 (Valenz: positiv vs. negativ) x 2 (Relation: assimilativ vs. kontrastiv) x 2 (Symmetrie: symmetrisch vs. asymmetrisch) x 2 (Blockreihenfolge: zuerst symmetrische Paarung vs. zuerst asymmetrische Paarung) – Design aufgebaut. Die ersten drei Faktoren „Valenz“, „Relation“ und „Symmetrie“ sind *within-subject-Faktoren*, während die Blockreihenfolge ein *between-subject-Faktor* ist. Insgesamt nahmen 18 Personen an der Studie teil ( $M_{\text{Alter}} = 30.6$  Jahre,  $SD_{\text{Alter}} = 13.8$  Jahre), welche als Kompensation eine halbe Versuchspersonenstunde bekamen. Die Proband\*innen wurden über einen Email-Verteiler der Universität Jena und über private Kontakte rekrutiert. Die Zaubersprüche wurden im Prinzip *one-to-one* mit je einem spezifischen valenten Stimulus gepaart. Somit wurden einer Versuchsperson hierbei nur zwei valente Stimuli – ein positiv und ein negativ valenter - präsentiert, da diese über beide Blöcke hinweg konstant blieben. Zwischen Personen wurden die präsentierten CS und US zufällig variiert. Die Stimuli wurden durch die Wortpaare „stärken“ – „schwächen“ (im asymmetrischen Block) und „verursachen“ – „verhindern“ (im symmetrischen Block) miteinander verknüpft. Die Proband\*innen wurden so instruiert, dass sie sich als Schüler einer Zauberschule die Wirkungen der folgenden Zaubersprüche einprägen sollen. Danach startet der erste Block des Experiments. Hier wurden nacheinander je ein Zauberspruch und ein valenter Stimuli durch eine Relation verknüpft präsentiert. Je nach Bedingung des Faktors Reihenfolge war das relationale Wortpaar symmetrisch bzw. asymmetrisch. Nach dieser Lernphase wurden die Versuchspersonen dazu aufgefordert die präsentierten Zaubersprüche auf einer Skala von -10 bis +10



zu bewerten. Danach begann der zweite Block mit dem jeweils anderen Wortpaar, welcher im Aufbau gleich ist.

### 3. Ergebnisse

Um die erhobenen evaluativen Ratings zu analysieren wurde eine vierfaktorielle ANOVA im  $2 \times 2 \times 2 \times 2$  - Design mit den oben genannten Faktoren durchgeführt. Hierbei ergab sich ein marginal signifikanter Haupteffekt des Faktors US-Valenz,  $F(1,128) = 2.798, p = 0.097, \eta^2p = 0.021, 95\% \text{-KI} [0; 0.084]$ . Der Interaktionseffekt der Faktoren Valenz und Relation wurde hochsignifikant,  $F(1,128) = 197.440, p < 0.001, \eta^2p = 0.607, 95\% \text{-KI} [0.475; 0.657]$ . Die Dreifachinteraktion von Valenz, Relation und Blockreihenfolge wurde signifikant,  $F(1,128) = 197.440, p < 0.023, \eta^2p = 0.040, 95\% \text{-KI} [0.000; 0.113]$ , sowie auch die Vierfachinteraktion von Valenz, Relation, Block und Block-Reihenfolge marginal signifikant wurde,  $F(1,128) = 197.440, p < 0.059, \eta^2p = 0.028, 95\% \text{-KI} [0; 0.094]$ . Alle weiteren Effekte sind nicht signifikant mit  $p \geq 0,126$ . So auch die Interaktion der Faktoren Valenz und Block, was bedeutet, dass die Asymmetrie der EC Effekte nicht durch die Symmetrie/Asymmetrie der Relationspaare moderiert wurde. Unsere Hypothese konnte somit nicht bestätigt werden.

Der marginal signifikante Haupteffekt des Faktors Valenz zeigt an, dass Paarungen eines CS mit einem US mit negativer Valenz im Mittel negativere Bewertungen induzierte als die Paarung eines CS mit einem positiven US. Durch den signifikanten Interaktionseffekt der Faktoren Valenz und Relation konnte das typische Befundmuster (disordinale Interaktion) des relationalen EC-Paradigmas repliziert werden (z.B. Zanon et al., 2014; Hu et al., 2017). Die marginal signifikante Dreifachinteraktion der Faktoren Valenz, Relation und Block-Reihenfolge zeigt, dass die Reihenfolge der Blöcke einen Einfluss auf die Symmetrie der EC Effekte hatte. Dies kann in der Vierfachinteraktion aufgeschlüsselt werden. So zeigt sich hier, dass Proband\*innen, die den symmetrischen Block zuerst bearbeiteten, in diesem auch ein symmetrisches Bewertungsmuster produzierten und im asymmetrischen Block ein asymmetrisches Muster. Dieser Effekt entspricht der Logik unserer Hypothese. Jedoch zeigte sich dieser nicht, wenn Proband\*innen zuerst den asymmetrischen Block bearbeiteten.

### 4. Diskussion

Die Hypothese konnte nicht bestätigt werden. Jedoch gibt der Befund der Vierfachinteraktion Hoffnung auf Bestätigung der Hypothese in modifizierten Folgestudien. Er gibt außerdem Hinweise auf eine von mehreren Limitationen dieser Studie: Das gewählte *Block-Design* zur Manipulation der Symmetrie des Relationspaares scheint ungeeignet, da sich nur in Abhängigkeit der Blockreihenfolge das entsprechende Muster fand. Eine mögliche Interpretation dieses Befundes ist, dass Proband\*innen intuitiv zunächst symmetrische Bewertungen vergeben. Folgt auf den ersten, symmetrisch bewerteten Block der asymmetrische, kann nun auch asymmetrischer bewertet werden. Folgt hingegen der symmetrische Block haben die Proband\*innen nicht die Möglichkeit *noch* symmetrischer zu bewerten. Im Umgang mit diesem Effekt sollten in zukünftigen Studien verschieden-symmetrische Wortpaare im selben Block präsentiert werden. Auch könnten Wortpaare verwendet werden, die sich in ihrer (A-)Symmetrie noch stärker unterscheiden. Weitere Limitationen der Studie stellen die sehr kleine Stichprobe sowie technische Probleme bei der Durchführung des Experiments dar.

### 5. Literatur

- Hu, X., Gawronski, B., & Balas, R. (2017). Propositional Versus Dual-Process Accounts of Evaluative Conditioning: I. The Effects of Co-Occurrence and Relational Information on Implicit and Explicit Evaluations. *Personality & social psychology bulletin*, 43(1), S. 17-32. doi:10.1177/0146167216673351
- Zanon R., De Houwer J., Gast A., & Smith CT. (2014). When does relational information influence evaluative conditioning? *Quarterly Journal of Experimental Psychology*, 67(11), S. 2105-2122. doi:10.1080/17470218.2014.907324

# Social Networks and Stress During the Covid-19 Pandemic

Laura Maas, Moritz Ratzke, Hannes Thomas

Advisor: Dr. Jennifer A. Bellingtier

## 1. Introduction

Transactional theory views stress as an interaction of personal and situation variables that influence perceptions of stress (Lazarus & Folkman, 1987). One such variable, social networks (e.g., friends, family, colleagues), typically helps to ameliorate perceived stress and anxiety. Larger networks enable access to formal and informal social support that can help individuals cope with stressors directly or indirectly by bolstering self-esteem (Harpham et al., 2002; Fiorillo & Sabatini, 2011). For example, Gerich (2013) found that larger close social networks were associated with reduced levels of perceived stress. On the other hand, such a relationship could not be found between the size of non-close networks (i.e., less emotionally-close friends) and perceived stress. The Covid-19 pandemic has restricted relationships via lockdowns and social distancing, as well as heightening concern for our own and other's health and well-being. Thus the benefits of social networks in these times of trouble could be mixed. Indeed, larger networks could increase the odds an individual knows someone adversely affected by Covid-19 whilst still providing a source of social support for coping with stressors. This leads us to our main hypothesis: larger social networks will be associated with more corona-related stressors, yet will buffer the relationship of corona-related stressors and perceived stress. We expected this buffering relationship to be greater for close (vs. non-close) social networks.

## 2. Method

*Participants and Procedure.* 130 participants ranging in age from 19 to 40 years ( $M = 21.73$ ,  $SD = 3.57$ ; 22% men; 98% native German speakers) completed our survey online via SoSci Survey. Participants were recruited via emailed survey-links, in compliance with data protection regulations (DGSVO). Before the start of the survey, we asked for informed consent and verified that every participant was 18 or older. As an incentive, participants could choose to enter a gift card raffle or receive experiment subject hours if they were psychology students at FSU Jena.

*Measures.* A translated and adapted version of the Coronavirus Impact Scale (Stoddard & Kaufman, 2020) measured the extent of corona-related stressors participants faced. It consisted of 12 corona-related stressors such as changed routines and concern for the well-being/health of others. Participants rated how much the pandemic has changed these aspects of their lives on a 4-point scale ranging from *no changes* (0) to *severe changes* (3). Mean scores were created ( $\alpha = .77$ )

Social network size was measured following van den Berg and Timmermans (2015) by asking for the number of people with whom one feels emotionally connected (close network) and those who are more than casual acquaintances but not very close (non-close network) in six categories (e.g., relatives, neighbors).

The German version of the Perceived Stress Scale-10 (Klein et al., 2016) measured participants' stress appraisals during the past month of the pandemic. Mean scores were created ( $\alpha = .84$ ).

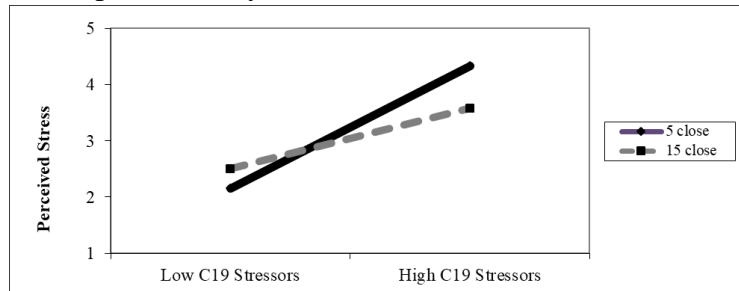
## 3. Results

Network size was not correlated with total corona-related stressors ( $r = .16$ , 95% BCa CI [-.06, .35],  $p = .068$ ), but did predict greater stress related to the single-item tapping concern for the well-being and health of others ( $r = .21$ , 95% BCa CI [.05, .38],  $p = .015$ ).

We performed bias-corrected and accelerated bootstrapped hierarchical regression analyses predicting perceived stress by entering the predictors in 3 steps: (1) corona stressors, (2) network

sizes, (3) interactions of stressors by network. This was performed separately for close networks, non-close networks, and in a combined model. In all models, there was a main effect of corona stressors ( $\beta = .49$  to  $.63$ ,  $ps < .001$ ) indicating that individuals who reported more stressors related to Covid-19 perceived greater stress in the preceding month. Covid-19 stressors accounted for 24% of the variance in perceived stress. When modeled separately, larger close networks ( $\beta = -.19$ ,  $p = .015$ ) and non-close networks ( $\beta = -.24$ ,  $p = .002$ ) both predicted lower perceived stress scores.

When modeled simultaneously, the effect of close networks became insignificant ( $\beta = -.06$ ,  $p = .514$ ), whereas the effect of non-close networks remained ( $\beta = -.20$ ,  $p = .042$ ). There were no interactions between C19 stressors and network size in the separated analyses, but there was an interaction with close networks in the combined model ( $\beta = -.33$ ,  $p = .039$ ). This interaction indicated that individuals with smaller close networks reported more perceived stress when C19 stressors were higher than those with larger close networks. In total, social networks explained an additional 8% of the variance in perceived stress.



#### 4. Discussion

Overall, our findings support the transactional theory of stress (Lazarus & Folkman, 1987) by indicating that the experience and perception of stress during the pandemic varied based on social network size. Although larger networks did not predict more overall stressors, they did predict more stress related to others. In addition, larger networks were associated with less perceived stress, and this was especially true for non-close networks. This stronger main effect for non-close networks was in contrast to our hypothesis. Potentially, people share more positive information/emotions with non-close networks compared to sharing stressors/negative emotions with close networks. It is also possible that non-close networks became "closer" through Zoom. Moreover, it is important to keep in mind that the variance in non-close networks was greater than in close networks.

In line with our hypothesis, the buffering effect of close networks on the relationship between stressors and perceived stress was significant and larger than the effect of non-close others. This finding supports the view that despite lockdowns and social distancing, close networks continued to provide social support and reduce stress perceptions during the pandemic (cf. Gerich, 2013).

#### 5. Literature

- Fiorillo, D., & Sabatini, F. (2011). Quality and quantity: The role of social interactions in self-reported individual health. *Social Science & Medicine*, 73(11), 1644–1652. <https://doi.org/10.1016/j.socscimed.2011.09.007>
- Gerich, J. (2013). Effects of social networks on health from a stress theoretical perspective. *Social Indicators Research*, 118(1), 349–364. <https://doi.org/10.1007/s11205-013-0423-7>
- Harpham, T., Grant, E., & Thomas, E. (2002). Measuring social capital within health surveys: Key issues. *Health Policy and Planning*, 17(1), 106–111. <https://doi.org/10.1093/heapol/17.1.106>
- Klein, E. M., Brähler, E., Dreier, M., Reinecke, L., Müller, K. W., Schmutzer, G., Wölfling, K., & Beutel, M. E. (2016). The German version of the Perceived Stress Scale—Psychometric characteristics in a representative German community sample. *BMC Psychiatry*, 16. <https://doi.org/10.1186/s12888-016-0875-9>
- Lazarus, R. S., & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, 1(3), 141–169. <https://doi.org/10.1002/per.2410010304>
- Stoddard, J., & Kaufman, J. (2020). Coronavirus Impact Scale. Retrieved from [https://www.phenxtoolkit.org/toolkit\\_content/PDF/CIS\\_Stoddard.pdf](https://www.phenxtoolkit.org/toolkit_content/PDF/CIS_Stoddard.pdf)
- Van den Berg, P., & Timmermans, H. (2015). A multilevel path analysis of social networks and social interaction in the neighbourhood. *Region*, 2(1), 55–66. <https://doi.org/10.18335/region.v2i1.32>

# Sind Polizisten die besseren Augenzeugen? Der Einfluss des Waffen Fokus Effekts auf die Zuverlässigkeit von Augenzeugenaussagen

Pauline Ender, Anne Fischer, Maximilian Francke, Lisa Hensel, Sümeyye Kocak

Leitung: Dr. Jürgen Kaufmann und B.Sc. Pauline Gürtler

## 1. Einleitung

Bisherige Studien zum *Weapon Focus Effect* zeigten, dass dieser einen besonderen Einfluss auf das Erinnerungsvermögen von Augenzeugen hat. Der Effekt ist dadurch charakterisiert, dass ein bedrohlicher Stimulus (z.B. eine Waffe) in der Hand eines Täters das Erinnerungsvermögen von Augenzeugen beeinträchtigt, da der Fokus während der Tat auf der Waffe liegt und nicht auf der Umgebung (Loftus et al., 1987). Die Einschränkung des Erinnerungsvermögens bei Zeugen geht scheinbar aus der Ungewöhnlichkeit des besonderen Stimulus hervor und resultiert in einer Fokussierung auf die zentralen Details statt der peripheren. Demnach sei der *Weapon Focus Effect* hoch, wenn der Stimulus nicht vorhersehbar ist (Pickel, 1998). Das ließ sich besonders an Waffen untersuchen, die stereotypisch für das jeweilige Geschlecht sind (Pickel, 2009). Es scheint sich dabei um einen moderaten Effekt auf das Augenzeugengedächtnis zu handeln (Fawcett et al., 2013). Allerdings sind bisherige Befunde zum *Weapon Focus Effect* eher uneindeutig (Baddeley, 2015). Die vorliegende Studie zielte darauf ab zunächst einen solchen *Weapon Focus Effect* nachzuweisen. Demnach wurde ein eingeschränktes Erinnerungsvermögen an periphere Details erwartet, wenn Probanden ein Bild mit einer Waffe sahen im Gegensatz dazu, wenn auf dem Bild keine Waffe zu sehen war. Darüber hinaus sollte überprüft werden, ob der *Weapon Focus Effect* bei Polizisten abgeschwächt war, da diese aufgrund des regelmäßigen Kontaktes mit Waffen eher an diese gewöhnt sind.

## 2. Methode

*Stichprobe.* Eine zu Beginn der Studie durchgeführte Poweranalyse ergab eine Stichprobengröße von  $N = 128$ . Die tatsächliche Stichprobe umfasste 193 Probanden (111 Männer, 82 Frauen) mit einem Altersdurchschnitt von 35,01 ( $SD = 14,74$ ). Es wurden 91 Polizisten und 101 Nicht-Polizisten (davon 26 Psychologiestudierende) miteinander verglichen. Die Teilnahme wurde bei Psychologiestudierenden, die durch den VPN-Verteiler der Uni Jena rekrutiert wurden, mit einer halben Versuchspersonenstunde vergütet. Ebenso wurden Probanden, wie etwa Polizisten, durch persönliche Anfragen im Bekanntenkreis rekrutiert und durch zusätzliche Anfragen an die Hochschule der Sächsischen Polizei (FH) in Rothenburg, sowie das Polizeirevier in Plauen.

*Material, Design und Vorgehen.* In diesem Online-Experiment, welches mit PsyToolKit erstellt wurde, wurden den Probanden zehn Bilder mit unterschiedlichen Szenarien und Personen gezeigt, die mit Photoshop (Version: Adobe Photoshop CS2) noch zusätzlich bearbeitet wurden. Jedes Bild existierte in zweifacher Ausprägung, d.h. die Person hielt entweder eine Waffe oder einen neutralen Gegenstand in der Hand. Gemäß eines gemischten Designs wurde jedem Probanden eine von zwei Reihenfolgen randomisiert zugewiesen. Somit wurde sichergestellt, dass jeder Proband ein Bild in nur in einer der beiden Ausführungen sah. Auf die Bilder folgten jeweils sechs Fragen (offen/geschlossen) zu peripheren sowie zu zentralen Details, wie auch zur Waffe bzw. zum Gegenstand selbst.

### 3. Ergebnisse

Für diese Studie wurde eine ANOVA mit Messwiederholung für die Faktoren Gegenstand (Waffe vs. Nicht-Waffe) und Details (peripher vs. zentral) durchgeführt, wobei ein Vergleich zwischen dem Zwischengruppenfaktor Polizisten und Nicht-Polizisten angestellt wurde.

Da finale Ergebnisse zu diesem Zeitpunkt nicht vorliegen, können diese noch nicht berichtet werden. Erwartet wird eine Interaktion zwischen den Faktoren Gegenstand und Details, wobei mehr zentrale als periphere Details bei dem Gegenstand Waffe erinnert werden. Generell werden schlechtere Erinnerungen an Details im Kontext der präsentierten Waffe erwartet. Dabei wird ein Gruppenunterschied zwischen Polizisten und „Nicht-Polizisten“ vermutet, bei dem ein abgeschwächter *Weapon Focus Effect* für die Gruppe der Polizisten erwartet wird.

### 4. Diskussion

Da die Ergebnisse noch nicht vorliegen, kann ein *Weapon Focus Effect* noch nicht bestätigt werden. Sollte ein Gruppenunterschied zwischen Polizisten vs. Nicht-Polizisten vorliegen, kann dies ein Hinweis darauf sein, dass der häufige Kontakt mit Waffen den *Weapon Focus Effect* verringert. Aufgrund der besseren Erinnerung an Details könnten Polizisten somit durchaus bessere Augenzeugen sein.

Die Studie leistet einen bedeutsamen Beitrag für die Wissenschaft und die Untersuchung des *Weapon Focus Effect*. Sie zeichnet sich vor allem durch den Gruppenvergleich mit Polizisten aus, der aufdecken kann, inwieweit der Kontakt zu Waffen eine Rolle für die Erinnerung von Details spielt. Eine mögliche Implikation dieser Studie könnte sein, dass bei Verbrechen, in denen eine Waffe genutzt wurde, das Erinnerungsvermögen der Augenzeugen und somit ihre Aussagen verzerrt sind. Bei einer Befragung von Augenzeugen zu einem Verbrechen könnte somit die Ermittlung des Kontakts mit Waffen ratsam sein, um so zu überprüfen, ob die Person als zuverlässiger Augenzeuge in Frage kommt.

Um die Hypothese hinsichtlich des Kontaktes zu bestärken, könnten künftige Untersuchungen mit weiteren Gruppen durchgeführt werden, die ebenfalls erhöhten Kontakt zu Waffen haben. Beispielsweise könnten Gamer, die virtuellen Kontakt zu Waffen haben, mit Nicht-Gamern verglichen werden.

In Zukunft könnte auf eine bessere Vergleichbarkeit zwischen den Gruppen geachtet werden, um gezieltere Rückschlüsse auf den Einfluss des Kontaktes mit Waffen zu treffen. Im Falle einer Replikation dieser Studie könnte die Einheitlichkeit der Berufstätigkeit der Gruppe der Nicht-Polizisten fokussiert werden.

### 5. Literatur

- Baddeley, A., Eysenck, M., & Anderson, M.C. (2015). *Memory*.
- Fawcett, J.M., Russell, E.J., Peace, K.A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the “weapon focus” literature. *Psychology, Crime & Law*, 19, 35-66.
- Loftus, E.F., Loftus, G.R., & Messo, J. (1987). Some facts about “weapon focus”. *Law and Human Behavior*, 11, 55-62.
- Pickel, K.L. (2009). The weapon focus effect on memory for female versus male perpetrators. *Memory*, 17(6), 664-678.
- Pickel, K.L. (1998). Unusualness and threat as possible causes of “weapon focus”. *Memory*, 6(3), 277-95.
- PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. Stoet, G. (2017). *Teaching of Psychology*, 44(1), 24-31.
- PsyToolkit -A software package for programming psychological experiments using Linux. Stoet, G. (2010). *Behavior Research Methods*, 42(4), 1096-1104.

# Implizites Beobachtungslernen unter Konformitätseinfluss

Niclas Dittmann, Nicole Klubkova, Marco Klug, Karl Schramm

Leitung: Dr. Carina G. Giesen

## 1. Einleitung

In unserem Empirischen Praktikum untersuchten wir das sogenannte *Observational Stimulus-Response (SR)-Binding*. Darunter versteht man die kognitive Verbindung eines auftretenden Stimulus und einer dazu beobachteten Reaktion eines Anderen (Giesen et al., 2014). Wiederholt sich der Stimulus und entspricht die geforderte Reaktion der beobachteten, zeigt sich dies in verkürzten Reaktionszeiten. Ist eine andere Reaktion gefordert, wird die Reaktionszeit verlängert, da die abgerufene beobachtete Reaktion inkompatibel ist. In einer Studie von Giesen et al. (2014) zum Observational SR-Binding stellte sich heraus, dass observational SR-Binding auftrat, wenn Interaktionspartner in einer Beziehung zueinander standen (Wettkampf oder Kooperation), jedoch nicht bei Unabhängigkeit voneinander. Somit warf sich uns die Frage nach weiteren sozialen Moderatoren des observational SR-Binding Effekts auf. Forschungen zur Konformität konnten zeigen, dass individuelle Urteile über die Gesichtsattraktivität im Einklang mit einer normativen Gruppenmeinung angepasst wurden (Shestakova et al. 2013). Wir vermuten, dass Konformität den sozialen Fokus und damit die Aufmerksamkeit auf die Reaktion des Interaktionspartners steigert. Daraus generiert sich unsere Forschungshypothese, die besagt, dass Konformität das SR-Binding Effekt verstärkt, im Vergleich zu Nichtkonformität.

## 2. Methode

Unsere Studie wurde vor der Erhebung am 24.05.2021 präregistriert (<https://aspredicted.org/blind.php?x=fr448x>) und hat ein positives Ethikvotum erhalten.

*Stichprobe.* Von 21 vollständigen Datensätzen mussten zwei wegen zu hoher Fehlerquote (>20%) im Gedächtnistest ausgeschlossen werden, sodass 19 Versuchspersonen (VP) in die Auswertung einfließen (9 Männer, 10 Frauen,  $M_{Alter} = 23.9$ ,  $SD = 7.1$ ).

*Ablauf.* Den VP wurde vor Beginn des Experiments suggeriert, sie würden das Experiment mit einer wissenschaftlichen Hilfskraft durchführen. Nach den Instruktionen folgte die erste Phase, in welcher die VP Gesichter nach Attraktivität auf einer Skala (1 = gar nicht attraktiv, 8 = sehr attraktiv) bewerten sollten. Es gab 29 Trials, ohne Feedback durch die vermeintliche Referenzgruppe. Anschließend folgte in Phase 2 die Observational SR-Binding Aufgabe (64 Prime-Probe Sequenzen), bei der jeweils zunächst die Reaktion des „Partners“ auf ein weißes Wort beobachtet (Prime Durchgang) und dann auf ein Wort in roter oder grüner Schrift durch Tastendruck (A (rot) / L (grün)) reagiert (Probe Durchgang) werden musste. Nach 25% Probe-Durchgänge erfolgte ein Gedächtnistest, bei dem die beobachtete Reaktion des „Partners“ durch Tastendruck wiederholt werden sollte. In Phase 3 wurden erneut Gesichter nach ihrer Attraktivität auf einer Skala von 1-8 beurteilt, diesmal gab es jedoch Rückmeldung, wie attraktiv eine vermeintliche studentische Referenzgruppe dieselben Gesichter bewertet hat. VP der Gruppe „mehrheitlich konformes Feedback“ erhielten in 75% der Durchgänge Feedback, das mit dem eigenen Rating übereinstimmte und 25% Rückmeldungen, die von der eigenen Meinung abwichen ( $\pm 3$  Punkte). Die VP der Gruppe „mehrheitlich nicht-konformes Feedback“ erhielten in 75% der Durchgänge Rückmeldungen, die von der eigenen Meinung abwichen ( $\pm 3$  Punkte), und 25% Feedback, das mit der eigenen Meinung übereinstimmte. In Phase 4 wurde die Observational SR-Binding Aufgabe wiederholt (64 Prime-Probe Sequenzen). Schlussendlich folgte erneut ein Attraktivitätsrating, das als Manipulationscheck diente.

*Design.* In dieser Onlinestudie haben wir ein 2x2x2 Design mit den unabhängigen Variablen „Gruppe“ (mehrheitlich konformes vs. non-konformes Feedback; between subjects), „Stimulusrelation“ (gleiches Wort vs. wechselndes Wort in Prime und Probe) und

„Reaktionskompatibilität“ (Probe Reaktion ist kompatibel vs. inkompatibel mit beobachteter Prime-Reaktion) gewählt.

### 3. Ergebnisse

Manipulation Checks: Attraktivitätsrating. Die mittlere Veränderung zwischen erstem Rating und Manipulation-Check-Rating für jede VP haben wir in einer 2 (Gruppe: konform vs. non-konform) x 3 (Gruppenmeinung: negativ, gleich, positiv) faktoriellen Varianzanalyse (ANOVA) getestet. Hier zeigte sich entgegen unserer Hypothese kein signifikanter Haupteffekt der Gruppe (konform vs. non-konform) auf die Veränderung der Attraktivitätsbeurteilung ( $F[1,11]=0.08$ ,  $p=.789$ ,  $\eta^2=.01$ ). Alle anderen Effekte wurden ebenfalls nicht signifikant (alle  $F_s < 0.71$ , alle  $p_s > .504$ ).

Performanz in der Observational SR-Binding Aufgabe. Für jede Versuchsperson wurden die mittleren Probe-Reaktionszeiten für das faktorielle Design berechnet. Danach wurden diese in eine 2 (Stimulusrelation: Stimulus Wiederholung vs. Wechsel) x 2 (Reaktionskompatibilität: kompatibel vs. inkompatibel) x 2 (Gruppe: konform vs. non-konform) gemischt-faktorielle ANOVA eingefügt. Die Analyse ergab ausschließlich nicht-signifikante Ergebnisse (alle  $F_s < 2.4$ , alle  $p_s > .146$ ). Die erwartete Dreifachinteraktion zwischen Stimulusrelation, Reaktionskompatibilität und Gruppe erreichte kein konventionelles Signifikanzniveau ( $F[1,17]=0.18$ ,  $p=.675$ ,  $\eta^2=.01$ ). Die Interaktion zwischen Stimulusrelation und Reaktionskompatibilität blieb ebenfalls aus ( $F[1,17]=0.30$ ,  $p=.593$ ,  $\eta^2=.02$ ), was bedeutet, dass sich keinerlei observational SR-Binding Effekte in unser Studie zeigten.

### 4. Diskussion

Unsere Hypothese, die besagt, dass Konformität mit einer Gruppe den Interaktionseffekt bei der Observational SR-Binding Aufgabe im Gegensatz zu Non-Konformität verstärkt, lässt sich nicht bestätigen. Es gibt keine signifikanten Unterschiede in der Reaktionszeit zwischen Probanden der konformen und non-konformen Bedingung. Der Effekt tendiert in Richtung eines stärkeren observational SR Binding Effekts in der non-konformen Bedingung. Entgegen unserer Hypothese lässt das vermuten, dass Non-Konformität den Interaktionseffekt zwischen Stimulusrelation und Reaktionskompatibilität verstärkt. Möglicherweise liegt das an einer Ausschlusserfahrung, die der Mensch durch Non-Konformität erlebt, welche den sozialen Fokus und damit Imitationsverhalten verstärkt. Eine weitere Ursache für die nicht signifikanten Ergebnisse ist unter anderem die geringe Power, die sich aus der geringen Probandenzahl ergibt (post-hoc Poweranalyse:  $1-\beta=0.20$ ). Auf die Frage, wie realistisch die Probanden die Interaktion im Experiment empfanden, wurde im Durchschnitt mit 4.7 (1 = gar nicht realistisch, 9 = sehr realistisch) geantwortet. Das lässt bezweifeln, dass die Suggestion eines sozialen Interaktionspartners erfolgreich war. Auch die Manipulation der Konformität scheint nicht gelungen, da sich die Gruppen in ihren Attraktivitätsratings nach Erhalt des Gruppenfeedbacks nicht signifikant unterschieden haben. Zusammenfassend generiert unsere Untersuchung jedoch weitere Ideen für Forschungsfragen im Bereich des Beobachtungslernens und dessen Beeinflussung durch soziale Moderatoren wie Konformität oder auch Ausschlusserfahrung.

### 5. Literatur

- Giesen, C., Hermann, J., & Rothermund, K. (2014). Copying Competitors? Interdependency Modulates Stimulus-Based Retrieval of Observed Responses. *Journal of Experimental Psychology: Human Perception & Performance*, 40(5), 1978-1991. doi: 10.1037/a0037614
- Shestakova A, Rieskamp J, Tugin S, Ossadtchi A, Krutitskaya J, Klucharev V. Electrophysiological precursors of social conformity. *Soc Cogn Affect Neurosci*. 2013 Oct;8(7):756-63. doi: 10.1093/scan/nss064. Epub 2012 Jun 7. PMID:22683703; PMCID: PMC3791064

# „Singen auf der Cocktail-Party“: Effekte von Stimmenklangfarbe und -karikierung auf die auditive Szenenanalyse

Anna Knospe, Natalie Freiboth, Jonas Dragendorf

Leitung: PD Dr. Romi Zäske

## 1. Einleitung

Einer der wohl wichtigsten Kommunikationskanäle ist unsere Stimme, deren Klangfarbe so individuell wie ein Fingerabdruck ist. Klangfarbe umfasst alle akustischen Parameter, die Stimmen voneinander unterscheidbar machen, wenn Tonhöhe, Lautheit, Schallort und Dauer gleich sind. Wie wir simultan präsentierte Stimmen anhand der Klangfarbe unterscheiden, ist jedoch weitgehend unerforscht. Während die auditive Szenenanalyse [1], also die Segregation mehrerer simultaner Schallquellen, eine wichtige Funktion des auditiven Systems darstellt, soll dieser Effekt im Chorgesang möglichst reduziert werden. Hier sollten sich Stimmen eines Registers ähneln und vereinen zu einem homogenen Klang. Aus Zählaufgaben wissen wir, dass in einer simulierten Cocktail-Party Situation 3-5 simultane Sprechstimmen unterschieden werden können [2], jedoch ist bisher unklar, ob hierfür die Klangfarbe oder andere Parameter verantwortlich sind. In einer ähnlichen Zähl Aufgabe manipulierten wir daher ausschließlich die Klangfarbe in „Multisänger-Szenen“, indem wir alte vs. junge Stimmen miteinander kombinierten unter Kontrolle aller anderen akustischen Parameter (WiSe). Da ältere im Vergleich zu jüngeren Stimmen als distinkter wahrgenommen werden [3] und damit leichter unterscheidbar und inhomogener klingen sollten [4], sagten wir für ältere Stimmen eine höhere Performanz vorher. Das Herausstechen einzelner Stimmen aus einer „Multisänger-Szene“, das in Chören eher unerwünscht ist, untersuchten wir im SoSe erstmals mit der sogenannten „disappearance-task“ [5]. Hier postulierten wir, dass das Verschwinden einer Stimme aus einer Szene häufiger bemerkt wird, wenn diese karikiert ist im Vergleich zu ihrer Originalversion bzw. Antikarikatur.

## 2. Methode

Am WiSe-Experiment nahmen 47 Personen (35 weiblich; Alter:  $M=23,4$ ; Range: 18 – 59 Jahre) teil. Im SoSe konnten wir nur 14 Personen (6 weiblich; Alter:  $M=22,6$ ; Range: 19 – 30 Jahre) rekrutieren, obwohl eine a priori Poweranalyse einen Stichprobenumfang von  $N=28$  ergeben hatte. Alle Probanden bestätigten über ein normales Hörvermögen zu verfügen, deutsche Muttersprachler zu sein und das jeweilige Online-Experiment [6,7] mit Kopfhörern durchzuführen. Die Teilnahme wurde auf Wunsch mit Vph vergütet. Als Stimuli dienten Vokale [a, e, i, o, u] aus der JESS-Datenbank [8], wobei für das Zählparadigma (WiSe) junge (18-25 Jahre) und ältere (60-73 Jahre) Stimmen und für das Disappearance-Paradigma (SoSe) nur junge Stimmen verwendet wurden.

**Zähl Aufgabe:** Pro Vokal (5), Geschlecht (2) und Altersgruppe (2) wurden je 5 Einzelstimmen ausgewählt, die eines von 20 Stimulusset ergaben. Daraus wurden je 2-4 Einzelstimmen zu „Multisinger-Stimuli“ kombiniert, wobei sich die Stimmen innerhalb eines Sets und zw. den Altersgruppen max. um 1 Hz in der Grundfrequenz unterschieden und auch hinsichtlich Intensität und Dauer (1,4 s) standardisiert waren. Die Probanden schätzten per Tastendruck (1-9) die Anzahl der Sänger für 240 Trials ein, die randomisiert aus insgesamt 800 Stimuli gezogen wurden und 1-4 Stimmen enthielten. Somit erhielten wir ein 2 x 4 Design mit Stimmenalter (jung vs. alt) und Stimmenanzahl (1,2,3,4) als Innersubjektfaktoren. **Disappearance-Aufgabe:** Pro Trial präsentierten wir nacheinander je zwei Multisinger-Stimuli (S1 und S2), wobei per Tastendruck entschieden wurde, ob diese identisch waren oder in S2 eine Stimme fehlte. S1 enthielt stets 3 Stimmen, d.h. zwei Originale und eine Targetstimme, welche ebenfalls als Original, Karikatur oder Antikarikatur vorkam (Innersubjektfaktor Karikaturgrad). Fehlte in S2 eine Stimme, war dies die Targetstimme. Aus insgesamt 900 Multisinger-Stimuli wurden randomisiert 120 Trials gezogen.



### 3. Ergebnisse

**Zähltaufgabe:** Eine ANOVA mit MW auf Stimmenalter (2) und Stimmenanzahl (4) ergab eine sign. Abnahme der Akkuratheit (ACC) mit steigender Stimmenanzahl,  $F(3, 138) = 240.35, p < .001, \eta_p^2 = .84$ . Außerdem war die ACC bei älteren vs. jüngeren Sängern höher,  $F(1, 46) = 10.96, p < .002, \eta_p^2 = .19$ . Eine Interaktion zw. beiden Faktoren,  $F(3, 138) = 14.57, p < .001, \eta_p^2 = .24$ , spiegelte eine höhere ACC für alte im Vergleich zu jungen Sängern wider, wenn 1, 3 und 4 Stimmen präsentiert wurden – und umgekehrt bei 2 Stimmen. Außerdem stieg die mittlere Schätzung der Stimmenanzahl sign. mit zunehmender Stimmenanzahl,  $F(3, 138) = 279.6, p < .001, \eta_p^2 = .86$ . Alte Stimmen erhielten sign. höhere Schätzungen als junge  $F(1, 46) = 21.35, p < .001, \eta_p^2 = .32$ . Eine sign. Interaktion ( $F[3, 138] = 14.57, p < .001, \eta_p^2 = .24$ ) spiegelte wider, dass die mittlere Schätzung bei alten vs. jungen Sängern höher war für 2, 3 und 4 Stimmen, umgekehrt bei einer Stimme. **Disappearance-Aufgabe:** Eine ANOVA mit MW auf Karikaturgrad (3) ergab einen sign. HE,  $F(2, 26) = 19.89, p < .001, \eta_p^2 = .60$ . Paarweise Vergleiche aller Stufen von Karikaturgrad ergaben sign. höhere ACCs für Karikaturen vs. Originale ( $t[13] = -4.21, p = .001$ ), Karikaturen vs. Antikarikaturen ( $t[13] = -4.21, p = .001$ ), aber keinen sign. Unterschied von Antikarikaturen und Originalen ( $t[13] = 2.02, p = .06$ ). RTs zeigten keinen sign. HE von Karikaturgrad,  $F(2, 26) < 1$ .

### 4. Diskussion

Wir konnten erstmals zeigen, dass sich sowohl das Zählparadigma [4] als auch das Disappearance-Paradigma [5] für die Untersuchung der Wahrnehmung von Multi-Sänger-Szenen eignet. Im Einklang mit [4] zeigte sich eine Verschlechterung der Performanz und damit der Segregationsfähigkeit mit steigender Sängeranzahl. Wie vermutet, wurden ältere Stimmen besser segregiert als junge, was mit einer distinkteren Klangfarbe älterer Stimmen zusammenhängen könnte [3]. Für ältere Chöre dürfte es daher schwieriger sein einen homogenen Klang zu erzeugen. Im Disappearance-Paradigma wurden Änderungen in der Größe einer Sängerszene gut, jedoch nicht perfekt wahrgenommen, analog zur „change deafness“ bei Änderungen in Umweltgeräuschen [5]. Erwartungsgemäß erhöhte die Karikierung einer einzelnen Stimme die Segregierbarkeit von 3 simultanen Sängern. Auch natürlich vorkommende Stimmen, welche wie Karikaturen akustisch vom Durchschnitt abweichen, könnten demnach möglicherweise leichter in Cocktail-Party-Situationen herausgefiltert und besser verstanden werden. Überraschenderweise erhöhten auch Antikarikaturen die Segregierbarkeit von Stimmen im Trend. Eine Erhöhung der Stichprobengröße ist nötig, um zu beurteilen, inwiefern es sich hierbei um einen tatsächlich existierenden Effekt handelt. Es bleibt zu untersuchen, ob unsere Befunde für einfache Vokale auch auf komplexere Äußerungen übertragbar sind und auf welche spezifischen Parameter der Klangfarbe ein homogener Chorklang beziehungsweise das Herausstechen von einzelnen Stimmen zurückzuführen ist.

### 5. Literatur

1. Bregman, A. S. (1990). Auditory scene analysis: The perceptual organization of sound. *The MIT Press*.
2. Kawashima, T., Sato, T. (2015). Perceptual limits in a simulated "Cocktail party". *Attention Perception Psychophysics*, 77(6), 2108-2120.
3. Zäske, R., et al. (2018). Electrophysiological correlates of voice memory for young and old speakers in young and old listeners. *Neuropsychologia*, 116, 215-227.
4. Latinus, M., et al. (2013). Norm-Based Coding of Voice Identity in Human Auditory Cortex. *Current Biology*, 23(12), 1075-1080.
5. Eramudugolla, R., et al. (2005). Directed attention eliminates 'Change deafness' in complex auditory scenes. *Current Biology*, 15(12), 1108-1113.
6. Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096-1104.
7. Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, 44(1), 24-31.
8. Zäske, R., et al. (2020). The Jena Speaker Set (JESS) - A database of voice stimuli from unfamiliar young and old adult speakers. *Behavior Research Methods*, 52(3), 990-1007.

# Anderthalb Meter Gesellschaft: Wie verändert COVID-19 unsere Gesichtswahrnehmung?

Erik Eyring, Dorothea Eichentopf, Lydia Krause, J.S., Jana Bodenbinder

Leitung: Dr. Géza Ambrus und B.Sc. Alexia Dalski

## 1. Einleitung

Menschen nehmen Gesichter subjektiv wahr - aber welche Faktoren haben einen Einfluss auf diesen Prozess? Häufig betrachtete Dimensionen der subjektiven Wahrnehmung von Gesichtern sind Attraktivität, Dominanz und Vertrauenswürdigkeit, die wiederum von einer Vielzahl von Faktoren beeinflusst werden. Die vorliegende Studie beschäftigt sich mit dem Einfluss des sozialen Kontextes auf Gesichtswahrnehmung, eine Thematik, die besonders in Hinblick auf die Pandemie und den damit einhergehenden sozialen Veränderungen von großer Bedeutung ist. So ist der Forschung zu entnehmen, dass bspw. die Einsamkeit während COVID-19 tendenziell anstieg, während die Lebenszufriedenheit und das Wohlbefinden unverändert blieben (Entringer & Kröger., 2020). Diese Studie beschäftigt sich mit der Frage: Wie beeinflusst das soziale Umfeld (Wohnsituation, Menge/Qualität der sozialen Kontakte, Einsamkeit, Beziehungsstatus etc.) die Wahrnehmung von Attraktivität, Vertrauenswürdigkeit und Dominanz von Gesichtern?

## 2. Methode

*Stichprobe und Verlauf.* Zu Beginn des Fragebogens sahen die Teilnehmer (N = 155, davon 115 weiblich, 36 männlich, 4 divers,  $Mdn=21$ ) insgesamt 20 kaukasisch aussehende Gesichter von erwachsenen Männern und Frauen. Diese Gesichter sollten auf einer Skala von 1 bis 7 auf den Eigenschaften „Vertrauenswürdigkeit“, „Attraktivität“ und „Dominanz“ bewertet werden. Folgende Kovariaten wurden erhoben: Alter, Geschlecht, Wohnsituation und der Beziehungsstatus. Weiterhin wurden, um das soziale Umfeld zu erfassen, folgende Variablen erhoben: Anzahl persönlicher und virtueller Kontakte in den letzten 7 Tagen und die soziale Unterstützung (erfasst durch die deutsche Version des "Enriched Social Support Inventory"). Zudem wurde die Zufriedenheit mit der Anzahl der angegebenen Kontakte miterhoben. Weiterhin wurden die Einsamkeit (UCLA-Loneliness-Scale) und die Depressivität (ADS-L) erfasst. Die Erhebung fand während der COVID-Pandemie statt. Die Daten wurden anschließend mit einem Rating-Algorithmus von vor der Pandemie verglichen.

## 3. Ergebnisse

Einige signifikante Ergebnisse ( $p \leq .05$ ), die in der durchgeführten Analyse gefunden wurden, werden im Folgenden vorgestellt.

*3.1. Zusammenhänge von Einsamkeit & depressiver Stimmung mit der Wahrnehmung von Gesichtern*  
Lediglich zwischen Depressivität und der Gesichtsbewertung in Bezug auf die Attraktivität wurde ein signifikanter negativer Zusammenhang gefunden ( $r = -.184, p = .024$ ).

*3.2. Zusammenhänge zwischen der Quantität der sozialen Kontakte und der Wahrnehmung von Gesichtern*

Personen mit im Mittel mehr realen Kontakten bewerten die Gesichter signifikant dominanter als Personen mit im Mittel weniger realen Kontakten ( $R^2 = 0.354, F(147) = 5.394, p = .022$ ).

Personen mit im Mittel mehr virtuellen als realen Sozialkontakten, haben die Gesichter als signifikant dominanter eingeschätzt, als Personen mit im Mittel mehr realen als virtuellen Kontakten ( $r = -.144$ ,  $p = .048$ ).

### 3.2. Zusammenhänge zwischen der Qualität der sozialen Kontakte und der Wahrnehmung von Gesichtern

Die Zufriedenheit der Sozialkontakten korreliert positiv mit der Wahrnehmung der Proband\*innen von Dominanz, Vertrautheit und auch Attraktivität.

(Dominanz ( $r = .249$ ,  $p = .002$ ); Vertrautheit ( $r = .343$ ,  $p = .000$ ); Attraktivität ( $r = .276$ ,  $p = .001$ ))

### 3.4. Unterschiede der Gesichtswahrnehmung als Mittelwertsdifferenzen zwischen Daten während Covid-19 und Daten vor der Pandemie (ermittelt durch den Algorithmus)

#### 3.4.1. Dominanz

Es wurden keine signifikanten Unterschiede in der Dominanz Wahrnehmung der Gesichtsstimuli gefunden.

#### 3.4.2. Attraktivität

Die weiblichen Gesichtsstimuli wurden von den Proband\*innen signifikant weniger attraktiv eingeschätzt, als vor der Pandemie (.493, 95% KI [.085, .901],  $t(157) = 2,385$ ,  $p = .018$ ).

Der Unterschied in Bezug auf die männlichen Gesichtsstimuli ist nicht signifikant.

#### 3.4.3. Vertrauenswürdigkeit

Sowohl die weiblichen Gesichtsstimuli, als auch die männlichen Gesichtsstimuli wurden von den Proband\*innen signifikant weniger vertrauenswürdig eingeschätzt, als vor der Pandemie

(weibliche Stimuli: .502, 95% KI [.107, .898],  $t(157) = 2,511$ ,  $p = .013$ ; männliche Stimuli: .582, 95% KI [.182, .983],  $t(156) = 2,871$ ,  $p = .005$ ).

## 4. Diskussion

Die vorliegenden Daten zeigen sich abzeichnende Tendenzen für Zusammenhänge von der Art unserer sozialen Kontakte und der Einschätzung von Gesichtern. Es fand sich jedoch nur in geringerem Maße ein Zusammenhang in Bezug auf die Kovariaten der Depressivität und Einsamkeit. Ein ausbalanciertes Design hätte die notwendige Power hervorgerufen, um den Aussagen bezüglich der Geschlechtseffekte Bedeutsamkeit zu verleihen. Dennoch hat uns die Pandemie und die damit einhergehenden starken Kontaktbeschränkungen gezeigt, welche Auswirkungen das soziale Umfeld auf Menschen und deren Wahrnehmung haben kann.

## 5. Literatur

Entringer, T.M. & Kröger, Hannes (2020). *Einsam, aber resilient: Die Menschen haben den Lockdown besser verkraftet als vermutet*. DIW aktuell, No.46, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

# Hat die unterschiedliche Symmetrie relationaler Wortpaare einen Einfluss auf EC-Effekte?

Friederike Brabetz, Jan Goebel, Lydia Hoffmeier, Clara Ladig

Leitung: M.Sc. Karoline Bading

## 1. Einleitung

Das Forschungsfeld der Evaluativen Konditionierung diskutiert derzeit die Kontroverse zwischen dem Ein-Prozess-Modell, welches sich allein auf Propositionsbildung bezieht und dem Zwei-Prozess-Modell, welches durch Assoziations- und Propositionsbildung das evaluative Lernen zu erklären versucht (Bading, Stahl & Rothermund, 2020). Das Relationale EC-Paradigma (Hughes, Ye, Van Dessel, De Houwer, 2018) dient der Untersuchung des Einflusses von relationalen Informationen (assimilative vs. kontrastive Beziehung) auf die Bewertung eines konditionierten Stimulus (CS) und ist als 2x2 Forschungsdesign (US-Valenz x CS-US-Relation) aufgebaut. Die Paarung von konditionierten und unkonditionierten Stimuli (CS-US-Paarung) kennzeichnet sich durch eine Aufteilung in assimilative vs. kontrastive Relationen. Bei assimilativen Relationen verschiebt sich die Bewertung des CS in Richtung US-Valenz, während kontrastive Relationen gegenteilige Beobachtungen herbeiführen. Laut des Ansatzes des Main-Effect-Approach wurde ein Haupteffekt des Faktors Valenz gefunden, welcher darauf hindeutet, dass es in der assimilativen und kontrastiven Bedingung unterschiedliche Bewertungen gibt. Die CS-Bewertungen zeigen ein asymmetrisches Muster in der Form auf, dass kontrastive Relationen in beiden US-Valenz-Bedingungen schwächere Bewertungen induzieren als die assimilativen. Dies lässt durch ein Zusammenspiel von propositionalen und assoziativen Prozessen (Zwei-Prozess-Modell) erklären, welche in der kontrastiven Bedingung entgegengesetzt zueinander wirken und eine abgeschwächte CS-Bewertung produzieren. Eine alternative *propositionale* Hypothese besagt, dass die Wortpaare selbst jene Asymmetrien verursachen. Um diese zu überprüfen, sollte in dieser Studie untersucht werden, ob Wortpaare mit niedriger Symmetrie auch asymmetrischere EC-Effekte hervorrufen.

## 2. Methode

*Design:* Das Design dieser Studie ist ein 2 (Symmetrie: niedrige Symmetrie vs. Hohe Symmetrie) x 2 (Relation: assimilativ vs. kontrastiv) x 2 (Valenz: positiv vs. negativ) x 2 (Reihenfolge Blöcke: zuerst symmetrische Paarung vs. zuerst asymmetrische Paarung) – Design. Die Faktoren „Symmetrie“, „Relation“, „Valenz“ wurden in einem *within-subject-Design* und der Faktor „Reihenfolge Blöcke“ (asymmetrisch/symmetrisch) in einem *between-subject-Design* manipuliert. Durch Verknüpfung der *within-Faktoren* wurden die Versuchspersonen mit folgenden vier Kombinationen konfrontiert:

Der CS wurde durch eine assimilative Beziehung mit einem positiven US gepaart

Der CS wurde durch eine kontrastive Beziehung mit einem negativen US gepaart

Der CS wurde durch eine assimilative Beziehung mit einem negativen US gepaart

Der CS wurde durch eine kontrastive Beziehung mit einem positiven US gepaart

In einem Block wurde dabei ein Wortpaar mit einer hohen Symmetrie und im zweiten Block eines mit einer geringen Symmetrie verwendet.

*Teilnehmer:* An der Studie nahmen 21 Versuchspersonen teil, von deren Beobachtungen keine Werte ausgeschlossen werden mussten und vollständig verwendet werden konnten.

*Material:* Als neutrale Stimuli (CS) wurden die Gesichter von Politiker\*innen verwendet, welche je nach Bedingung etwas Positives oder Negatives „verursachen“ oder „verhindern“ (symmetrisches Wortpaar) beziehungsweise „befürworten“ oder „ablehnen“ (asymmetrisches Wortpaar). Diese Wortpaare wurden aus einer von den Autor\*innen durchgeführten Vorstudie generiert. Als Stimuli dienten Bilder zu den Themen Tierhaltung, Umweltpolitik und Energieerzeugung. In jedem Block

wurden vier neutrale Stimuli (CS) präsentiert (ein CS pro Kombination – s.o.), wobei die neutralen Stimuli zufällig zu den Bedingungen zugeordnet wurden (*between-subject-Design*). Die neutralen Stimuli (CS) wurden mit valenten Stimuli (US) von gleicher Valenz (positiv/negativ), gepaart.

*Prozedur:* Die Versuchspersonen bekamen als Teilnehmer eines Kongresses für junge Politiker\*innen den Auftrag, sich deren politischen Programme einzuprägen, bevor der erste Block begann. Ein CS (Politiker\*in) wurde immer mit einem Verb und einem Stimulus (positiv/negativ) präsentiert. Zum Schluss wurden die Proband\*innen gebeten, die Politiker\*innen auf einer Skala (-10 bis +10) zu bewerten. Nach erneuter Instruktion folgte der zweite Block, welcher sich nur durch das Wortpaar (symmetrisch/asymmetrisch) für die CS-US-Paarung, vom ersten Block unterschied und wieder mit der Bewertung der Politiker\*innen endete.

### 3. Ergebnisse

Die evaluativen Ratings wurden mit einer mehrfaktoriellen Anova getestet ( $2 \times 2 \times 2 \times 2$ ; siehe oben). Aus der Analyse ergab sich ein signifikanter Interaktionseffekt der Faktoren Valenz und Relation,  $F(1,152)=231.520$ ,  $p<2e-16$ ,  $Eta^2p=0.604$ , 95%-KI (0.486;0.654), sowie ein marginal signifikanter Effekt der Vierfach-Interaktion von Valenz, Symmetrie, Relation und Reihenfolge,  $F(1,152)=3.076$ ,  $p<0.082$ ,  $Eta^2p=0.02$ , 95%-KI(0.111;0.314). Alle anderen Effekte wurden nicht signifikant (alle p-Werte  $\geq 0.265$ ). Darunter fällt auch der Interaktionseffekt der Faktoren Valenz und Asymmetrie,  $F(1,152)=0.451$ ,  $p<0.503$ ,  $Eta^2p=0.003$ , 95%-KI(0;0.039), was bedeutet, dass sich die Asymmetrie der EC-Effekte nicht zwischen den Blöcken mit je unterschiedlich symmetrischen Wortpaaren unterscheidet. Unsere Hypothese wurde damit nicht bestätigt. Der signifikant gewordene Interaktionseffekt Valenz-Relation zeigt wie erwartet einen regulären EC-Effekt in der assimilativen und einen gegenteiligen Effekt in der kontrastiven Bedingung. Die geringfügig signifikante Vierfach-Interaktion zeigt, dass die Asymmetrie der EC-Effekte durch die Relation beeinflusst wurde, kann aber nicht eindeutig interpretiert werden.

### 4. Diskussion

Unsere Hypothese, dass weniger symmetrische Wortpaare in einem relationalen EC-Paradigma asymmetrischere Bewertungen erzeugen, wurde nicht bestätigt. Unser Forschungsansatz konnte somit keine eindeutige empirische Evidenz im Sinne des propositionalen Ein-Prozess-Modells liefern. Diese Studie unterliegt jedoch einigen Limitationen (sehr kleine Stichprobe, teilweise technische Probleme bei der Durchführung des Experiments), weshalb die Ergebnisse lediglich als vorläufig interpretiert werden können. Für zukünftige Studien könnten einige Hinweise abgeleitet werden. Zum einen sollte in Zukunft auf das *between-subject-Design* verzichtet werden, da die unterschiedliche Asymmetrie der verwendeten Wortpaare durch die Reihenfolge-Manipulation und die damit verbundenen Möglichkeiten der Korrektur des eigenen Bewertungsmusters nicht richtig zur Geltung kommen kann. Zum anderen vermuten wir, dass sich die Relationen nicht absolut, sondern kontextabhängig auf die EC-Effekte auswirken. Möglicherweise sind in unserem Fall durch die sehr alltagsrelevante Politiker-Thematik auch persönliche Einstellungen mit in die Bewertung eingeflossen.

### 5. Literatur

- Hughes, S., Ye, Y., Van Dessel, P., & De Houwer, J. (2019). When People Co-occur With Good or Bad Events: Graded Effects of Relational Qualifiers on Evaluative Conditioning. *Personality & social psychology bulletin*, 45(2), S. 196-208. doi:10.1177/0146167218781
- Bading, K., Stahl, C., & Rothermund, K. (2020). Why a standard IAT effect cannot provide evidence for association formation: The role of similarity construction. *Cognition and emotion*, 34(1), S. 128-143. doi:10.1080/02699931.2019.1604322

	<u>Stream 1</u>	<u>Stream 2</u>	<u>Stream 3</u>	<u>Stream 4</u>
Link:	<a href="https://uni-koeln.zoom.us/j/99687219499">https://uni-koeln.zoom.us/j/99687219499</a>	<a href="https://uni-jena-de.zoom.us/j/61482916657">https://uni-jena-de.zoom.us/j/61482916657</a>	<a href="https://uni-jena-de.zoom.us/j/61936208810">https://uni-jena-de.zoom.us/j/61936208810</a>	<a href="https://uni-jena-de.zoom.us/j/66238150189">https://uni-jena-de.zoom.us/j/66238150189</a>
Meeting-ID:	996 8721 9499	614 8291 6657	619 3620 8810	662 3815 0189
Passwort:	stream1	stream2	stream3	stream4
Host:	Karoline Bading	Tingting Huang	Anne Grigutsch	Clara de Paula Couto
Jury:	Tobias Koch, Antje Rauers, Jenny Bauer	Stefan Schweinberger, Christian Bloszies, Scania Steger	Ilona Croy, Helene Kreysa, Alina Suelmann	Thomas Kessler, Birk Hagemeyer, Lena Schenker
	Gruppe 1	Gruppe 6	Gruppe 15*	Gruppe 16
14:00 – 14:20	<i>Nur gut gemeint!? – Sexismus in unseren Komplimenten</i>	<i>Wenn Stimmen hervorstechen - Effekte von Karikierung und selektiver Aufmerksamkeit auf die Wahrnehmung von Gesangsgruppen</i>	<i>Social Networks and Stress During the Covid-19 Pandemic</i>	<i>Sind Polizisten die besseren Augenzeugen? Der Einfluss des Waffen Fokus Effekts auf die Zuverlässigkeit von Augenzeugenaussagen</i>
	Gruppe 2	Gruppe 7	Gruppe 12	Gruppe 17
14:25 – 14:45	<i>Bitte steh mir bei! – Bystander-Effekt bei Online-Mobbing</i>	<i>Gesellschaftliche Normen – Die Mutter aller Vorurteile?</i>	<i>Let's face it! Wie die unidirektionale Manipulation konfigurationaler Information das Lernen von Gesichtern beeinträchtigt.</i>	<i>Implizites Beobachtungslernen unter Konformitätseinfluss</i>
	Gruppe 3	Gruppe 8	Gruppe 13	Gruppe 18
14:50 – 15:10	<i>Propositionale Maße im Duell - Effekte von Leistungsfeedback auf indirekt und direkt gemessenen Selbstwert</i>	<i>Traumurlaubsziel vs. Mordgelüste: Wie verändern sich automatische Gedächtnisprozesse des Lügens bei intimen Themen im Vergleich zu oberflächlichen Themen?</i>	<i>Die Schwierigkeit und die Vorhersagekraft im IAT</i>	<i>„Singen auf der Cocktail-Party“: Effekte von Stimmenklangfarbe und -karikierung auf die auditive Szenenanalyse</i>
	Gruppe 4	Gruppe 9	Gruppe 14	Gruppe 19
15:15 – 15:35	<i>Komponenten von Bestrafung und deren Wirkung - Wie nehmen wir Täter*innen wahr?</i>	<i>Effects of Personality on Perceived Stress during the COVID-19 Pandemic</i>	<i>Die magische evaluative Konditionierung - Produzieren asymmetrische Relationspaare asymmetrische EC-Effekte?</i>	<i>Anderthalb Meter Gesellschaft: Wie verändert COVID-19 unsere Gesichtswahrnehmung?</i>
	Gruppe 5	Gruppe 10	Gruppe 11*	Gruppe 20
15:40 – 16:00	<i>Wer ist der Täter? Wie Attraktivität die Auswahl in Lineups beeinflusst</i>	<i>Effects of punishment on offenders' self-perception</i>	<i>Observational SR-Binding – „Mein, Dein oder Unser?“ – wie beeinflusst unser Mindset das Beobachtungslernen</i>	<i>Hat die unterschiedliche Symmetrie relationaler Wortpaare einen Einfluss auf EC-Effekte?</i>

\* Änderung zum ursprünglichen Programm: Gruppe 11 und 15 haben die Präsentationszeiten getauscht.