



Programm 2024

19. Jenaer
Empiriepraktikumskongress

Psychologie
FSU Jena

Foyer, EAP
05.07.2024

Danksagung

Es ist wieder so weit: Heute, am 05. Juli 2024., findet bereits zum 19. Mal der Jenaer Empiriepraktikumskongress statt. Dieser Kongress kommt jedes Jahr nur durch die Unterstützung und das Engagement zahlreicher Personen zustande, denen an dieser Stelle Dank gebührt.

Zuallererst möchten wir den Juroren und Jurorinnen aus der Professorenschaft, dem Mittelbau und der Studierendenschaft danken, die sich die Zeit nehmen, die einzelnen Arbeiten mit viel Sorgfalt, Expertise und Eifer zu begutachten.

Darüber hinaus gilt unser Dank allen teilnehmenden Gruppen und ihren 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 den Gastvortrag zum Thema *Let it Flow? The Challenges in Manipulating Flow (Optimal) Experience in Experiments*“ haben wir Dr. Dwight Tse (Lecturer Psychology, University of Strathclyde Glasgow) gewinnen können, der außerdem als Juror die Begutachtung der Arbeiten unterstützt. Für diese Bereicherung des Kongressprogramms danken wir ihm herzlich.

Einen besonderen Programmpunkt des EmPra-Kongresses bildet jedes Jahr die Auszeichnung und Prämierung der besten Arbeiten. Für die finanzielle Unterstützung durch Bereitstellung der Preisgelder bedanken wir uns beim Institut für Psychologie. Ganz besonders freuen wir uns, auch in diesem Jahr wieder Gutscheine für die Jenaer Universitätsbuchhandlung Thalia vergeben zu können und danken Herrn Christopher Schnell für die langjährige Unterstützung.

Des Weiteren danken wir den Studierenden aus der Fachschaft Psychologie für die Hilfe bei der Suche studentischer Juroren. Ebenso danken wir den Mitgliedern der CampusCouch, die den Sektverkauf organisieren, welcher einen maßgeblichen Beitrag zu einer angenehmen Kongressatmosphäre leistet.

Ein letztes großes Dankeschön geht an die Hilfskräfte des Lehrstuhls für Allgemeine Psychologie II, die sich stets in vielfältiger Weise an der Organisation beteiligen und damit einen reibungslosen Ablauf des Kongresses erst ermöglichen.

Für die Abteilung Allgemeine Psychologie II,

Clara de Paula Couto, Kira Franke, Matthäus Rudolph

Inhaltsverzeichnis

Gruppe 1	6
<i>Einfluss der wahrgenommenen Vertrauenswürdigkeit durch verschiedene Gesichterstimuli auf beobachtungsbasierte Stimulus-Reaktions-Bindung</i>	
Jule Grahmann, Ali Hwajeh, Alena Schmidt, Ilka Schöttke, Laura Weber	
Gruppe 2	8
<i>Das Bewusstsein beim Kontingenzlernen</i>	
Marléne Beck, Johann Drilltzsch, Leon Lochau, Jakob Menkens, Till Streifler	
Gruppe 3	10
<i>Effekte geteilter und ungeteilter Vorteile auf Gruppendynamiken</i>	
Kerstin Hupach, Lotta Ullrich, Neele Rosenthal & Leon Göldner	
Gruppe 4	12
<i>Think outside the box – look outside the window Zuschreibung (nicht-)geteilter Vorurteile und Verhalten in Sozialen Medien im Gruppenkontext innerhalb und außerhalb des normativen Fensters</i>	
J. Ochs, S. Pohl, A. Budnitskiy, A.-M. Spengler-Ulbrich, P. Schultz	
Gruppe 5	14
<i>„You proved me right!“ – Effects of age-based stereotype threat on older adults’ performance in technology</i>	
Lilli Scherf, Wanda Denecke, Amelie Große, Abd-u-Rahman Siddiqi, Elias Chryssikakis, Louis Gresser	
Gruppe 6	16
<i>Young and distracted: Is it stereotype threat or stereotype challenge?</i>	
Elena Luna Schwieger; Julia Kemp; Maren Schneider; Luise Göttlich; Lilli Westphal	
Gruppe 7	18
<i>Can you feel the music or are you just listening? Differences in the emotions induced by music as a function of musicality and engagement</i>	
Albrecht Blochmann, Michal Duda, Lisa Fischer, Antonia Kaiser, Mirjam Schlinke, Amy Joy Zacharias	
Gruppe 8	20
<i>“When music doesn’t matter” – Die deutsche Version des Barcelona Music Reward Questionnaire (dtBMRQ): Eine Validierungsstudie.</i>	
Madita Jodes, Charlotte Jolmes, Kim Meyer, Willi Moxter, Lena Rölz, Merle Roth	
Gruppe 9	22
<i>Bist du umweltfreundlich oder tust du nur so? Untersuchung der prädiktiven Power des impliziten Assoziationstests</i>	
Simone Dahmen, Jonas Krellwitz, Zoe Mack, Judy Schubert	
Gruppe 10	24
<i>Lässt sich die Testschwierigkeit und somit die prädiktive Power des impliziten Assoziationstests über die Target-Stimuli beeinflussen?</i>	
Barbara Gräser, Magdalena Klodt, Conrad Opitz, Annika Thiel	
Gruppe 11	26
<i>Worte, die verbinden: Kann Selbstoffenbarung Beziehungen vertiefen?</i>	
Anna Dührsen, Fenja König, Clara Elia Makrinus, Kira Wagner, Melina Zäh	

Gruppe 12	28
<i>Do we feel worse together? Zusammenhänge zwischen negativem Affekt und dem Teilen von Emotionen in Partnerschaften</i>	
Nina Brünkmann, Fynn Kornmann, Sandra Büttner, Julia Nassua, Friederike Peters	
Gruppe 13	30
<i>Schatz, erzähl mir mehr von deinen Emotionen! Soziales Teilen von Emotionen in Partnerschaften und die Rolle von Gesprächszufriedenheit</i>	
Helene Merten, Felicitas Schneider-Rausch, Lisa Rabe, Kristin Stöhr	
Gruppe 14	32
<i>#staypositive - Der Einfluss von Alter und Valenz berichteter Erfahrung auf das Soziale Teilen von Emotionen bei Frauen</i>	
Antonia Burzer, Pauline Götze, Wiebke Klages, Yolenn Mindt, Una Minten	
Gruppe 15	34
<i>Sonnenbrille oder Bart: Nettes Accessoire oder erfolgreiche Tarnung?</i>	
Monique Becker, Katharina Bösendorfer, Lena Breidt, Emma Fieler, Friederike Gehre, Lilly Hesse, Lena-Sophie Raspe, Elisa Ruehlmann, Alexandra Zimmer, Tabea Neumann	
Gruppe 16	36
<i>Wer war's? Forschungsansätze zur Verbesserung von Zeugenaussagen mittels der Stimme und virtueller Realität</i>	
Xenia Elenschleger, Johannes Fester, Camille Louise Fischer, Rebecca Fischer, Helin Hersek, Lena Konradi, Elina Lippmann, Celine Lindig, Filiz Scholz, Denis Pantschenko, Maximilia Schinner	
Gruppe 17	38
<i>Zwischen Bildschirmen und Begegnungen – die Rolle von parasozialen und sozialen Interaktionen bei der Wahrnehmung von Persönlichkeiten</i>	
Lara Hornischer, Linda Lehmann, Emma Lippold, Sophie Mintert, Laura Press, Annika Schmidt	
Gruppe 18	40
<i>Binge watching für kulturelle Sensibilität: Beeinflusst das Anschauen koreanischer Serien den Other-Race Effekt?</i>	
Pauline Fahle, Elham Kazemi, Samuel Melle, Katharina Riewe, Emily Schulz.	
Gruppe 19	42
<i>Listen to my voice – do you trust me?</i>	
Demiray, Canel; Huang, Ying; Kemmler, Lea; Märkl, Marlene; Nedziella, Angelina; Sahin, Beyza	
Gruppe 20	44
<i>Trust me, you know me! Familiarität und Glaubwürdigkeit von Stimmen</i>	
Daniel Bruins, Leonie Friedrich, Jordana Knappe, Elisa Krebs, Carina Marit Smit	
Gruppe 21	46
<i>Money, Money, Money! Does your political orientation determine whether you explore and invest in a local or foreign bank?</i>	
Manzke, Lisa; Möller, Mareike; Scheipers, Merle; Thiele, Victoria; Weihert, Marlene	
Gruppe 22	48
<i>Would you eat this fruit? Effect of political orientation on exploration and attitudes towards foreign and local foods</i>	
Torben Dolgner, Sophia Hering, Isabelle Scheibler, Lina Bulmahn	

Einfluss der wahrgenommenen Vertrauenswürdigkeit durch verschiedene Gesichterstimuli auf beobachtungsbasierte Stimulus-Reaktions-Bindung

Jule Grahmann, Ali Hwajeh, Alena Schmidt, Ilka Schöttke, Laura Weber

Leitung: Kira Franke

1. Einleitung

Tagtäglich treffen wir auf viele Menschen mit verschiedensten Gesichtern, von denen einige deutlich mehr vertrauenswürdig erscheinen als andere. Es besteht generelle Einigkeit darüber, dass Vertrauen fundamental für soziale Beziehungen ist (Lewis & Weigert, 1985) sowie auf impliziten Einschätzungen beruht und dynamisch auf der Grundlage von Erfahrungen aktualisiert wird (Chang et al., 2010). Frühere Studien zeigen, dass soziale Relevanz soziales Lernen und die Bildung von Stimulus-Reaktions-Verbindungen modulieren kann. Durch die Wiederholung des Stimulus wird eine Verknüpfung zur vorherigen Beobachtung aufgebaut, welches anschließend die eigene Reaktion beeinflusst (Giesen et al., 2014). Deshalb haben wir uns in der vorliegenden Studie die Frage gestellt, welchen Einfluss die Vertrauenswürdigkeit von Gesichtern auf beobachtungsbasierte S-R-Bindungs- und Abruffeffekte in einem Online-Video-Paradigma hat. Da durch Vertrauen die soziale Relevanz steigt, werden Handlungen einer als vertrauenswürdig wahrgenommenen Person mehr Aufmerksamkeit geschenkt, was wiederum das eigene Verhalten beeinflusst. Wir sind davon ausgegangen, dass die Abruffeffekte größer sind, wenn die Vertrauenswürdigkeit des Gesichtes des Interaktionspartners hoch ist im Gegensatz zu niedriger Vertrauenswürdigkeit.

2. Methode

Stichprobe. Für die Stichprobe wurden 163 Personen (91 weibliche und 72 männliche) für die online-Studie rekrutiert, von denen 82 in der Gruppe mit dem sehr vertrauenswürdig und 81 in der Gruppe mit dem wenig vertrauenswürdig aussehenden Interaktionspartner waren.

Design. Die Studie umfasste den Zwischensubjekt-Faktor Vertrauenswürdigkeit (hohe vs. niedrige Vertrauenswürdigkeit) sowie den Innersubjekt-Faktor Antwortkompatibilität (kompatibel vs. inkompatibel) und Stimulusrelation (Wortwiederholung vs. Wortwechsel) mit jeweils zwei Stufen und ist damit ein 2x2x2 Design. Zur Manipulation der wahrgenommenen Vertrauenswürdigkeit interagierten die Versuchspersonen mit einem scheinbaren Interaktionspartner mit einem entweder sehr vertrauenswürdigem oder einem wenig vertrauenswürdigem Avatar. Die abhängige Variable war die Reaktionszeit in den Probe-Sequenzen.

Ablauf. Den Versuchspersonen wurde suggeriert, dass sie an einer Farbklassifizierungsaufgabe gemeinsam mit einer anderen Person teilnahmen, der Interaktionspartner war jedoch programmiert. Während der Durchführung des Experiments, in dem farbige Worte als „rot“ oder „grün“ klassifiziert werden sollten, wurde vor dem Erscheinen des Wortes jeweils der Avatar der Person gezeigt, die reagierte. Während die Teilnehmenden die scheinbare Reaktion ihres Interaktionspartners auf dem Bildschirm beobachteten, erschien ihnen das Stimulus-Wort in weiß. Die Versuchspersonen reagierten immer in der Probe-Sequenz, wobei die auszuführende Reaktion entweder der im Prime beobachteten entsprach (kompatibel) oder nicht (inkompatibel) und der Stimulus entweder wiederholt oder verändert wurde. In 25% aller Trials erfolgte am Ende ein Gedächtnistest für die Versuchsperson, in dem sie die letzte Reaktion ihres Interaktionspartners anzeigen sollten. Der Manipulationscheck erfolgte durch die Abfrage der Vertrauenswürdigkeit aller im Experiment genutzten Stimuli auf einer siebenstufigen bipolaren Skala.

3. Ergebnisse

Anhand eines t -Test ließe sich die Funktionalität des Gedächtnistest beweisen $t(159.6) = 0.2922$, $p = .7705$. Zudem zeigte ein t -Test keinen signifikanten Unterschied in den Vertrauenswürdigkeitsbewertungen zwischen den Gruppen "Vertrauenswürdig" und "Nicht vertrauenswürdig" $t(161) = .3409$, $p = .7336$. Die Bewertung der Vertrauenswürdigkeit für alle Stimuli (Partner und Kontrollstimuli) wurde durchgeführt und die Ergebnisse zeigen einen signifikanten Effekt der Vertrauenswürdigkeit auf die Beurteilungen $t(161) = 13.41$, $p < 0.001$. Zudem zeigte t -Test keinen signifikanten Unterschied in den Realitätsbewertungen zwischen den Gruppen $t(161) = -.1708$, $p = .8646$.

Test der Hypothesen. Die Überprüfung der Hypothesen erfolgte anhand einer dreifaktorielle ANOVA mit den Faktoren Stimulusrelation, Antwortkompatibilität und Vertrauenswürdigkeit. Die Analyse ergab keinen signifikanten zweifaktoriellen Interaktionen zwischen Stimulusrelation und Antwortkompatibilität, was darauf hindeutet, dass keine Abruffeffekte vorliegen $F(1,161) = 1.01$, $p = .32$. Ebenso wurde keine signifikante dreifach-Interaktion zwischen Stimulusrelation, Antwortkompatibilität und Vertrauenswürdigkeit festgestellt $F(1,161) = .14$, $p = 0.71$. deutet darauf hin, dass die Vertrauenswürdigkeitsbewertung des Gesichts keinen unterschiedlichen Einfluss auf die Wechselwirkung zwischen Antwortkompatibilität und Stimulusrelation hatte, was der ursprünglichen Hypothese widerspricht.

Zudem ergab die explorative Analyse einen signifikanten Einfluss der Interaktion mit dem Partner auf die wahrgenommene Vertrauenswürdigkeit $F(1,8) = 23.73$, $p < .001$.

4. Diskussion

Die Ergebnisse der Studie zeigen, dass die Manipulation der Vertrauenswürdigkeit zwar erfolgreich war, aber nicht zu signifikanten Unterschieden in den Abruffeffekten geführt hat, das heißt, die Studienteilnehmenden aus beiden Gruppen haben die beobachteten Reaktionen nicht zur eigenen Handlungsregulation genutzt. Unsere Analysen ergaben auch, dass die alleinige Interaktion mit dem Partner die wahrgenommene Vertrauenswürdigkeit positiv beeinflusst, was gegen eine ausreichend starke soziale Relevanz der Vertrauenswürdigkeitsmanipulation spricht. Ausschließen kann man allerdings mangelnden Fokus, da allerdings durch den Gedächtnistest die Aufmerksamkeit wirksam auf die Handlungen des Interaktionspartners gerichtet werden konnte. Die signifikante zweifach-Interaktion im ersten Block zwischen niedriger Vertrauenswürdigkeit und Stimulusrelation könnte damit erklärt werden, dass bedrohliche Gesichter mehr Aufmerksamkeit auf sich ziehen und sich Teilnehmende dadurch eher von mit ihnen assoziierter Handlungen beeinflussen lassen. Der Effekt verfällt möglicherweise im zweiten Block, weil mit häufiger Interaktion ein Gewöhnungseffekt eintritt.

5. Literatur

- Chang, L. J., Doll, B. B., van 't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87–105
<https://doi.org/10.1016/j.cogpsych.2010.03.001>
- Giesen, C, Herrmann, J., & Rothermund, K. (2014). Copying competitors? Interdependency modulates stimulus-based retrieval of observed responses. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(5), 1978–1991.
<https://doi.org/10.1037/a0037614>
- Lewis, J. D., & Weigert, A. (1985). Trust as a Social Reality. *Social Forces*, *63*(4), 967.
<https://doi.org/10.2307/2578601>

Das Bewusstsein beim Kontingenzlernen

Marléne Beck, Johann Drilltzsch, Leon Lochau, Jakob Menkens, Till Streifler

Leitung: Matthäus Rudolph

1. Einleitung

Kontingenzlernen ist die Fähigkeit des kognitiven Systems, Regelmäßigkeiten in der Umwelt zu erkennen und auf diese zu reagieren. Die aktuelle Grundlagenforschung beschäftigt sich mit der Frage, ob dem Kontingenzlernen assoziatives Lernen, propositionales Lernen oder episodischer Abruf zu Grunde liegt. In der aktuellen Studie wurde untersucht, inwiefern das Kontingenzlernen von der verfügbaren kognitiver Kapazität abhängt: Dazu sollte parallel zur einer Farbkategorisierungsaufgabe eine Zweitaufgabe bearbeitet werden (Mitchell et al., 2009).

Im Experiment wurde (a) die kognitive Belastung durch eine zweite Aufgabe (leicht vs. schwer) manipuliert, während (b) gleichzeitig für den episodischen Abruf der letzten Reaktion kontrolliert wird, die einen Großteil des Kontingenzlernerffekts erklärt (Giesen et al., 2020; Rudolph & Rothermund, 2024).

Zunächst wird dabei überprüft, ob ein Kontingenzlernerffekt (schnellere Reaktion bei häufigen Wort-Farb Kombinationen; CL) auftritt. Anschließend wurde untersucht, ob der CL Effekt vollständig durch den Abruf der Reaktion beim letzten Auftreten desselben Wortes erklärt werden kann oder ob noch ein residualer CL Effekt zu finden ist, der den Einfluss globaler Kontingenzen auf das Verhalten widerspiegelt (residualer CL-Effekt). Des Weiteren prüfen wir, ob die Schwierigkeit der Zweitaufgabe den CL-Effekt moderiert (d.h., ob der residuale CL-Effekt geringer ist, wenn die zweite Aufgabe schwieriger ist). Wir postulieren, dass (1) der residuale CL-Effekt von der Schwierigkeit der Aufgabe abhängt (d.h., dass kein residuale CL-Effekt auftritt, wenn die Zweitaufgabe zu schwer ist) und das (2) episodischer Abruf unabhängig von der Zweitaufgabe erfolgt (keine Interaktion zwischen Reaktionsabruf und Aufgabenschwierigkeit).

2. Methode

Bei dieser Studie wurden Daten von 71 Versuchspersonen erhoben. Das Experiment dauerte etwa 30 Minuten. Dabei wurde ein Farb-Wort-Kontingenzlernparadigma mit einer zusätzlichen Zweitaufgabe verwendet. Als abhängige Variablen wurden die Reaktionszeit und die Genauigkeit der Reaktionen gemessen.

Es wurden vier neutrale deutsche Wörter (fast, ganz, kurz, klein) angezeigt, die in der Kategorisierungsaufgabe mit den Tasten „K“ und „D“, abhängig von der Darstellungsfarbe zugeordnet werden mussten. Je zwei der vier Farben (rot oder grün bzw. blau oder gelb) waren einer Taste zugeordnet. Jedes Wort wurde häufig in einer Farbe (hohe Kontingenz) und selten in der anderen Farbe angezeigt (geringe Kontingenz). Bevor das Wort in einer Farbe erschien, kam es in weiß. Die Wörter wurden angezeigt, bis die Versuchspersonen reagierten.

Zusätzlich sollten in einigen Blöcken zwischen den Farbkategorisierungen Rechenaufgaben gelöst werden. Auf einen Startwert mussten präsentierte Werte aufaddiert oder subtrahiert werden. Dabei gab es leichte (Startwert: 0, immer +1) und schwere Rechenaufgaben (Startwert: 240-260, zwischen -9 und +9).

Insgesamt durchliefen die Versuchspersonen 8 Experimentalblöcke. Die ersten 6 Blöcke bestanden abwechselnd aus schweren und leichten Zweitaufgaben. In jedem Block erfolgten etwa 6 Lösungsabfragen der Rechenaufgaben. Zwei der Wort-Farb Kombinationen waren immer der einfachen Zweitaufgabe zugeordnet, während die anderen zwei Wort-Farb Kombinationen immer mit der schweren Zweitaufgabe kamen. Nach den Experimentalblöcken wurde das Bewusstsein über die Kontingenzen abgefragt.

3. Ergebnisse

Für die Analyse wird eine Multilevel-Regression (Trials genestet in Versuchspersonen) genutzt. Als Prädiktor wird zuerst die Kontingenz (hoch vs. niedrig) verwendet. Es ergibt sich ein starker Kontingenzlernereffekt, das heißt Versuchspersonen reagieren schneller in Trials mit hoher Kontingenz als in Trials mit geringer Kontingenz ($M = -11.51$ ms, $t[25589] = -6.79$, $p < 0.001$). Bei der Hinzunahme des Prädiktors für episodischen Abruf (gleiche Reaktion vs. andere Reaktion) verschwindet der reine Kontingenzlernereffekt vollständig ($M = -0.37$ ms, $t[25587] = -0.2$, $p = 0.844$) und ein starker Effekt des episodischen Abrufs ist feststellbar. Demnach reagieren die Teilnehmer schneller auf Wörter, welche beim letzten Vorkommen desselben Wortes die gleiche Reaktion erfordern ($M = -22.37$ ms, $t[25587] = -13.29$, $p < 0.001$).

Zuletzt wird der Prädiktor für die Zweitaufgabe (leicht vs. schwer) aufgenommen. Der residuale Kontingenzlernereffekt ist weiterhin nicht signifikant ($M = -0.37$ ms, $t[25584] = -0.19$, $p = 0.846$) und episodischer Abruf hat einen starken Effekt auf die Reaktionszeit ($M = -22.43$ ms, $t[25584] = -13.36$, $p < 0.001$). Der neue Prädiktor für die Aufgabenschwierigkeit zeigt ebenfalls einen signifikanten Haupteffekt, das heißt Versuchspersonen reagieren schneller in Trials mit einer leichten Zweitaufgabe im Vergleich zu denen mit schwerer Zweitaufgabe ($M = -16.03$ ms, $t[25583] = -11.06$, $p < 0.001$). Die Interaktion zwischen Zweitaufgabe und residualem Kontingenzlernereffekt ($M = 5.53$ ms, $t[25584] = 1.47$, $p = 0.142$), sowie die Interaktion zwischen Zweitaufgabe und episodischem Abruf ($M = -4.45$ ms, $t[25585] = -1.33$, $p = 0.185$) zeigen beide keinen signifikanten Effekt.

Zur Ermittlung des allgemeinen Kontingenzbewusstseins der Teilnehmer, wird ein t -Test genutzt, wobei der prozentuale Anteil der korrekten Antworten in der Abfrage des Kontingenzbewusstseins gegen die Wahrscheinlichkeit, zufällig richtig zu liegen (25%), getestet wurde ($M = 0.34$, $t[283] = 3.25$, $p < 0.001$).

4. Diskussion

Mit unserer Studie konnten wir zeigen, dass die Manipulation mit einer Zweitaufgabe das Antwortverhalten der Versuchspersonen beeinflusst. Das bedeutet, dass die Teilnehmer im Mittel schneller in Blöcken mit einfacher Zweitaufgabe reagierten. Es konnte jedoch nicht beobachtet werden, dass der Kontingenzlernereffekt durch die Zweitaufgabe moduliert wird. Die Hypothese, dass episodischer Abruf unabhängig von der Zweitaufgabe ist, konnte bestätigt werden.

Es ist aber anzumerken, dass durch Hinzunahme des Prädiktors für episodischen Abruf kein residueller CL-Effekt mehr auftrat. Dies steht im Einklang damit, dass der residuale CL Effekt nur auftritt, wenn sich die Versuchspersonen der Kontingenzen bewusst sind (Rudolph & Rothermund, 2024). Demzufolge konnte der Einfluss der Zweitaufgabe auf den residualen Kontingenzlernereffekt nicht untersucht werden. In Zukunft könnte man die Kontingenz im Experiment erhöhen und die Häufigkeit der Wort-Farb-Paarungen damit salienter machen. Da auch das Bewusstsein über die Kontingenzen sehr gering ausfiel, könnte man zukünftig allgemeine Instruktionen zu den vorhandenen Kontingenzen geben.

5. Literatur

- Giesen, C. G., Schmidt, J. R., & Rothermund, K. (2020). The Law of Recency: An Episodic Stimulus-Response Retrieval Account of Habit Acquisition. *Front Psychol*, 10, 2927. <https://doi.org/10.3389/fpsyg.2019.02927>
- Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behav Brain Sci*, 32(2), 183-198; discussion 198-246. <https://doi.org/10.1017/s0140525x09000855>
- Rudolph, M., & Rothermund, K. (2024). Two sources of color-word contingency learning: Episodic retrieval of stimulus-response bindings and propositional knowledge. *J Exp Psychol Learn Mem Cogn*. <https://doi.org/10.1037/xlm0001353>

Effekte geteilter und ungeteilter Vorteile auf Gruppendynamiken

Kerstin Hupach, Lotta Ullrich, Neele Rosenthal & Leon Göldner

Leitung: Pascal Gelfort

1. Einleitung

Im Rahmen der Europawahl, ging die populistische und als gesichert rechtsextreme Partei „Alternative für Deutschland“ in Thüringen als deutlich stärkste Kraft hervor (vgl. Bundeswahlleiterin.de, 2024). Eine zentrale Rolle im Wahlkampf spielt dabei das Verhältnis der Parteien zu verschiedenen sozialen Gruppen, wie beispielsweise Geflüchteten. Darauf aufbauend stellt sich die Frage, unter welchen Bedingungen Vorurteile zur Mobilisierung von Wähler*Innen beitragen können.

Deswegen wird in der aktuellen Studie untersucht, wie geteilte und nicht geteilte Vorurteile die Beziehung von Individuen zu einer Eigengruppe beeinflussen. Basierend auf Heiders Balancetheorie (1958), Turners Selbstkategorisierungstheorie (1987) und den Befunden von Peters und Kashima (2007) nehmen wir an, dass Individuen ihre Beziehungen zu anderen Personen und Gruppen auf der Grundlage gemeinsamer Vorurteile anpassen.

Wir erwarten, dass Teilnehmende, die ein Vorurteil mit einem Erzähler teilen, die Gruppe stärker unterstützen und sich dieser verstärkt zugehörig fühlen, als wenn sie das Vorurteil nicht teilen. Darüber hinaus vermuten wir, dass dieser Zusammenhang durch die Art des Erzählers moderiert wird: Ein*e individuelle*r Repräsentant*In der Gruppe hat weniger Einfluss auf die Beziehung zur geteilten Eigengruppe als eine Aussage der Gruppe an sich. Letzterer ist wahrscheinlicher, die Beziehung des Individuums zur gesamten Gruppe zu ändern. Ziel dieser Untersuchung ist es, die Mechanismen zu verstehen, durch die Vorurteile und die Darstellung des Erzählers die Gruppendynamik beeinflussen.

2. Methode

Die Studie verwendet ein 2x2 Between-Subjects-Design mit den unabhängigen Variablen Vorurteilkongruenz (geteilte vs. nicht geteilte Vorurteile) und Narrationsweise (individuelles Gruppenmitglied vs. Gruppe als Ganzes)

Im Zuge wurden Wahlplakate von verschiedenen Parteien (CDU, AFD, B90/Grüne, SPD, Die Linke, FDP), welche zum Erhebungszeitpunkt in thüringer Kommunalwahlen antraten, und anschließend den Versuchspersonen ($n=207$) präsentiert. Jedes Wahlplakat zeigt dabei eine von zwei möglichen, von uns ausgedachten, Aussagen („Flüchtlinge integrieren und tolerieren!“ und „Feste Grenzen für Deutschland. Vor Überfremdung schützen!“), wobei die erste Aussage als kongruentes Vorurteil für Wähler*Innen der Parteien AFD und CDU, die zweite als kongruentes Vorurteil für Wähler*Innen der Parteien SPD, B90/Grüne, Die Linke und FDP präsentiert wurden, umgekehrt als inkongruentes Vorurteil.

Der Faktor der Erzählart wurde manipuliert, indem die Aussage zusammen mit einem/einer Politiker*In auf dem Wahlplakat (individuelles Gruppenmitglied) oder alleinstehend und mit dem Logo der Partei versehen gezeigt wird (Gruppe als Ganzes). Um mögliche Bias´ und Einflüsse der Beliebtheit von lokalen Politiker*Innen auf Versuchspersonen entgegenzuwirken wurden für alle Parteien die gleichen fiktiven Kandidat*Innen gezeigt. Jede Partei wurde durch den/die gleiche/n Politiker*In vertreten, mit gleichen fiktionalen Namen und einem fiktionalen KI-generierten Portrait. Die Prompts für die KI-Bildererstellung waren dabei *Mann/Frau*, im *mittelhohen Alter* und in *formaler Business Kleidung*.

Um die Auswirkungen des Geschlechts des Erzählers auf die abhängigen Variablen (AVs) ausschließen zu können, wurde in der Bedingung mit individuellem Erzähler eine zufällige Zuweisung eines Mannes oder einer Frau vorgenommen. Zudem haben wir überprüft, ob die

Versuchsteilnehmenden, die von uns den Parteien zugeschriebenen kongruenten oder inkongruenten Meinungen teilen. Versuchspersonen (17), deren persönliche Vorurteile nicht kongruent zu den der Parteien zugeschriebenen waren, wurden aus der Analyse ausgeschlossen.

3. Ergebnisse

Unsere Hypothese, dass inkongruente Vorurteile bei einem gruppenbasierten Erzähler (Individuum vs. Logo) zu geringerer Sympathie, Identifikation und Einsatzbereitschaft (im Rahmen einer zukünftigen Wahl-, Beteiligungs- und Spendenbereitschaft) führen, wurde bestätigt.

Nach einer zweifaktoriellen ANOVA konnten wir einen statistisch signifikanten Haupteffekt der Vorurteilkongruenz auf eingeschätzte Sympathie der Partei, ($F(1,188) = 69.621$, $p < 0.001$, $\eta^2_p = 0.232$) und eine signifikante Interaktion zwischen Art des Erzählers und der Kongruenz ($F(1,188) = 6.028$, $p = 0.015$, $\eta^2_p = 0.02$) beobachten.

In einer post-hoc t-Test Analyse zeigt sich, dass die Art des Erzählers nur in der inkongruenten Bedingung einen Einfluss auf die Sympathie hat ($t(83) = -2.538$, $p = .013$, $d = 0.534$).

Das gleiche Muster findet sich auch bei Identifikation (t-Test inkongruent: $t(93) = -2.137$, $p = 0.035$, $d = 0.449$) und bei zukünftiger Wahlbereitschaft ($t(85) = -3.246$, $p = 0.002$, $d = 0.678$). Für die Spendenbereitschaft findet man lediglich einen geringen Haupteffekt der Kongruenz ($F(1,188) = 6.138$, $p = 0.014$, $\eta^2_p = 0.397$).

Bei der Bereitschaft, sich für die Partei einzusetzen, bspw. an Sitzungen teilzunehmen oder der Partei beizutreten fanden sich keine signifikanten Ergebnisse ($p > 0.05$).

Im Rahmen eines Robustheitstest haben wir unsere Ergebnisse bei Konstanthaltung verschiedener Drittvariablen wie Alter, Bildungsgrad oder Geschlecht bestätigen lassen, welche ebenfalls keine signifikanten Effekte gezeigt haben.

4. Diskussion

Die Ergebnisse unserer Studie bestätigen die Hypothese, dass kongruente Vorurteile zu einer gesteigerten Sympathie, Identifikation und zukünftigen Wahlbereitschaft für die Partei führen. Eine zweifaktorielle ANOVA zeigt einen statistisch signifikanten Haupteffekt der Vorurteilkongruenz auf die wahrgenommene Sympathie der Partei. Signifikant ist die Interaktion zwischen der Narrationsweise und der Kongruenz, die in der inkongruenten Bedingung stark ausgeprägt ist.

Unsere Ergebnisse legen nahe, dass Vorurteile, die durch die Gruppe als Ganzes vermittelt werden, das Verhältnis zu dieser Gruppe stärker beeinflussen als solche, die durch ein individuelles Gruppenmitglied geäußert werden. Dies unterstreicht die Bedeutung kollektiver Repräsentation in der politischen Kommunikation. In kongruenten Vorurteilsbedingungen hingegen zeigt die Art des Erzählers keinen signifikanten Einfluss auf Sympathie oder Wahlbereitschaft.

Unsere Erkenntnisse unterstützen die Annahme, dass kollektive Narrative in der politischen Kommunikation eine herausragende Rolle spielen und die Gruppenzugehörigkeit stärker beeinflussen als individuelle Aussagen.

Diese Ergebnisse betonen die Notwendigkeit, die kollektive Dynamik von Vorurteilen und Erzählungen in der politischen Kommunikation stärker zu berücksichtigen. Sie verdeutlichen die Macht kollektiver Identitäten und ihre Rolle in der Stabilität politischer Gruppen.

5. Literatur

- Bundeswahlleiterin (2024). Europawahlergebnisse Thüringen
<https://bundeswahlleiterin.de/europawahlen/2024/ergebnisse/bund-99/land-16.html>
- Heider, F. (1958). *The psychology of interpersonal relations*. John Wiley & Sons Inc. <https://doi.org/10.1037/10628-000>
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.
- Peters, K., & Kashima, Y. (2007). From social talk to social action: Shaping the social triad with emotion sharing. *Journal of Personality and Social Psychology*, 93(5), 780–797. <https://doi.org/10.1037/0022-3514.93.5.780>

Think outside the box – look outside the window

Zuschreibung (nicht-)geteilter Vorurteile und Verhalten in Sozialen Medien im Gruppenkontext innerhalb und außerhalb des normativen Fensters

J. Ochs, S. Pohl, A. Budnitskiy, A.-M. Spengler-Ulbrich, P. Schultz

Leitung: Pascal Gelfort

1. Einleitung

Du stolperst über ein Social-Media Profil der AfD – Was erwartest du? Lobende Worte über Ehrenamtliche, Hasskommentare über Geflüchtete? Und wie reagierst du? Ob wir Personen bestimmte Vorurteile zuschreiben, ist nicht nur Resultat deren Gruppenmitgliedschaft sondern auch des sog. „normative window“ (Crandall et al. 2013). Dieses besagt, dass es Unterschiede in der gesellschaftlichen Wertung von Einstellungen in Bezug auf soziale Gruppen gibt, abhängig davon, ob sie kontrovers betrachtet werden oder es einen Bewertungs-Konsens gibt. Kontrovers diskutierte Gruppen befinden sich demnach im “normativen Fenster von Vorurteilen” und werden in der Vorurteilsforschung vielseitig beleuchtet. Darauf bezugnehmend und in Anlehnung an Heiders Balance-Theorie (1958) und die Selbstkategorisierungstheorie von Turner (1987) sollten Individuen erwarten, dass Mitglieder der Eigengruppe ihre individuellen Vorurteile teilen. Dieser Zusammenhang sollte durch das Vorurteilstarget (d.h. die bewertete Gruppe) weiter eingeschränkt werden. Gemäß der Konzeptualisierung des “normative window” ist es unwahrscheinlich, dass die Versuchspersonen (folgend VPN) Fremdgruppenmitgliedern Einstellungen zuschreiben, die ihren eigenen widersprechen, wenn das Vorurteilstarget gesamtgesellschaftlich einvernehmlich bewertet wird. Befindet sich das Target jedoch in kontroversen Debatten, erwarten wir hier eine Zuschreibung von Vorurteilen entgegen derer der VPN. Wir untersuchen außerdem, wie die Wahrnehmung der Gruppenzugehörigkeit die Reaktion auf wertende Kommentare in Sozialen Medien verändert. Weiterhin erwarten wir in Bezug auf Targets innerhalb des normativen Fensters (umstrittene) eine höhere Handlungstendenz, also das Verhalten in den Sozialen Medien der VPN (z.B. auf Kommentare, welche sich auf das Target beziehen, zu reagieren) als bei Targets außerhalb des normativen Fensters (unumstrittene).

2. Methode

Unsere Stichprobe besteht aus $n = 245$ Versuchspersonen (184W/56M/3D), die durchschnittlich 24.89 Jahre alt, größtenteils studentisch und nach eigenen Angaben eher links liberal orientiert sind. Unsere Unabhängigen Variablen umfassen Individuelle Vorurteile, Eigen- und Fremdgruppenzugehörigkeit sowie Umstrittene/ Unumstrittene Vorurteilstargets. Als Abhängige Variablen wurden die attribuierten Vorurteile, attribuierte Stereotype, die Erwartete Valenz von Kommentaren und das Verhalten in Sozialen Medien erhoben. Für die Zuweisung von Eigen-/Fremdgruppe wurde ein forced choice Format genutzt, bei dem Parteien nach Präferenz geordnet werden mussten. Anhand der am stärksten präferierten/ abgelehnten Partei erfolgte die experimentelle Gruppenzuteilung. Den VPN wurde daraufhin ein fiktiver Instagram-Post präsentiert, aus dem das Vorurteilstarget und die Gruppe, nicht aber die Bewertung des Targets durch die Gruppe, ersichtlich wurde. Targets sind dabei Gruppen, die sich entweder außerhalb (Ehrenamtliche, Terrorist*innen) oder innerhalb (Arbeitslose, Menschen muslimischen Glaubens) des normativen Fensters befinden. Es wurde erfragt, welche Vorurteile, Stereotype und Kommentare in sozialen Medien von Mitgliedern der präsentierten Gruppe erwartet werden und welche Reaktionen (likes, Kommentare,

etc.) auf diese gezeigt werden würden. Außerdem sollten die Proband*innen selbst einen Kommentar verfassen.

3. Ergebnisse

In einer multiplen Regressionsanalyse wurde die erwartete Kommentarvalenz auf die Prädiktoren Individuelle Vorurteile, Gruppenzugehörigkeit und die Art des Vorurteilstargets regrediert. Es zeigten sich signifikante Interaktionen zwischen Individuellen Vorurteilen und der Gruppenzugehörigkeit, $t(219) = -2.79, p < .01$, zwischen Individuellen Vorurteilen und Art des Targets, $t(219) = 2.79, p < .01$, sowie Target und Gruppenzugehörigkeit, $t(219) = 2.89, p < .01$.

Simple Slope Analysen zeigten, dass VPN nur in der unumstrittenen Bedingung erwarteten, dass die Bewertung der Targets mit der ihren übereinstimmt, $b = 9.01, SE = 0.94, t(219) = 9.53, p < .01$. Wir konnten zeigen, dass VPN von der Fremdgruppe Bewertungen erwarteten, die ihren eigenen widersprachen, $b = -6.38, SE = 1.97, t(219) = -3.24, p < .01$.

Post-Hoc T-Test zeigen signifikante Mittelwertsunterschiede zwischen der Eigen- und Fremdgruppe bezüglich eines umstrittenen Targets. Von der Eigengruppe werden demnach positivere Kommentare bzgl. eines umstrittenen Targets erwartet als von der Fremdgruppe, $d = 0.87$.

Zudem fanden wir signifikante Mittelwertsunterschiede zwischen den Fremdgruppenbedingungen mit umstrittenen vs. unumstrittenen Targets. Von der Fremdgruppe werden demnach negativere Kommentare über ein umstrittenes Target erwartet, im Vergleich zu solchen die sich auf ein unumstrittenes Target beziehen, $d = -0.66$.

Mittels einer logistischen Regression wurde die Wahrscheinlichkeit, einen Kommentar zu " liken " getestet, wobei sich eine signifikante Interaktion aus Art des Kommentars und Gruppenzugehörigkeit ergab, $z = -3.01, p = >.01$. Jedoch stieg die Wahrscheinlichkeit für ein like nur bei Eigengruppenmitgliedern mit der Positivität des Kommentars, $OR = 2.21$.

4. Diskussion

Entgegen unserer Hypothese fanden wir keine 3-fach Interaktion der Faktoren Individuelle Vorurteile, Gruppe und Target. Dennoch zeigte sich, dass die erwarteten Kommentare abhängig von der Art der Gruppe und der des Targets sind. Kommentare, die mit der eigenen Haltung übereinstimmen, werden zum einen von Eigengruppen-Mitgliedern und generell bei nicht-kontroversen Targets erwartet. Konfliktlinien gibt es hauptsächlich um kontroverse Targets. So finden sich mehr positive Interaktionen mit der Eigengruppe, auch wenn die Fremdgruppe die gleiche Einstellung teilt.

Das Ausbleiben der 3-fach-Interaktion könnte auf eine zu geringe Power und eine Kolinearität der Faktoren individuelle Vorurteile und Targets zurückzuführen sein. Darauf aufbauende Studien sollten dementsprechend ein größeres Sample an Targets testen, um die Effekte für Targets mit höherer Varianz in der Bewertung zu finden. Zudem sollten für mehr Varianz die Anzahl der Targets erhöht werden, um einer Konfundierung entgegenzuwirken.

5. Literatur

- Crandall, C. S., Ferguson, M. A., & Bahns, A. J. (2013). When we see prejudice: The normative window and social change. In C. Stangor & C. S. Crandall (Eds.), *Stereotyping and prejudice*, Psychology Press, (3) 53–69. <https://doi.org/10.4324/9780203567708>
- Heider, F. (1958). *The psychology of interpersonal relations*. Psychology Press. 2013 John Wiley & Sons Inc. <https://doi.org/10.1037/10628-000>
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.

„You proved me right!” – Effects of age-based stereotype threat on older adults’ performance in technology

Lilli Scherf, Wanda Denecke, Amelie Große, Abd-u-Rahman Siddiqi, Elias Chrysikakis, Louis Gresser

Advisor: M. Clara P. de Paula Couto

1. Introduction

Can exposure to age-stereotypes undermine performance in stereotype-relevant tasks? Previous research has shown that stereotypes, especially age stereotypes, have serious negative effects on older adults and their performance (Lamont et al. 2015). Exposure to relevant stereotypes can lead individuals to conform to these stereotypes (Steele & Aronson 1995), leading to a self-fulfilling prophecy. Age is the only social category where group belonging changes over life. Stereotype embodiment theory suggests that age stereotypes, internalized over life, become self-relevant as people age, influencing aging in accordance with these stereotypes (Levy 2009). Previous research has suggested that technology is an important stereotype domain in which older adults are negatively perceived (Lamont et al. 2021). Despite the potential negative consequences of age-based stereotype threat (ABST) regarding technology this area is still understudied. Hence, in our pre-registered study we examined the influence of ABST on older adults’ performance in technology tasks. Our hypotheses were: (1) Under ABST, older adults perform worse in technology tasks, and (2) self-efficacy, importance of the stereotyped domain, and intergenerational contact moderate the relationship between stereotype threat and task performance.

2. Methods

Participants. An a priori power analysis indicated a minimum of 102 participants to discover a medium effect ($d = .50$). We recruited 139 participants aged 50-68 via email. After applying the exclusion criteria, the final sample was 106 split into Control ($n = 56$, 66.1% female, $M_{age} = 66.23$, $SD = 8.86$) and ABST ($n = 50$, 66.0% female, $M_{age} = 65.76$, $SD = 8.25$).

Design. A 2 (Group: Control vs. ABST, between) x 2 (Performance assessment: before vs. after technology tasks, within) mixed-design was used.

Materials and Procedure. The study was conducted online using SoSci Survey. Participants provided informed consent, answered demographic questions, and completed baseline behavior items and importance of the assessed domain. They were then randomly assigned to either the Control group or ABST group. ABST was manipulated via a Social Consensus Feedback (SCF) paradigm, where participants reflected on others’ agreement with ABST. The ABST group was informed that most people believe older adults are worse with technology, while the Control group was told performance depends on individual factors. Participants completed 10 self-developed pretested technology tasks (e.g., social media, messaging apps, Office, online shopping) as quickly as possible to assess the performance. Intergenerational contact and self-efficacy were measured post-task. The mean accuracy in the technology tasks was the dependent variable.

3. Results

We entered baseline behavior, age, domain importance, Group, self-efficacy, and the interaction of self-efficacy by Group as one block in a regression model. Being older was associated with worse performance, $B = -1.09$, $p < .001$. Against our hypothesis, the main effect of Group was not significant, $B = 2.09$, $p = .280$. However, we found a cross-over self-efficacy x group interaction ($B = 1.71$, $p = .045$) This means that the interaction was indicated by the difference in the mean accuracy between the Control and ABST groups at low and high levels of self-efficacy rather than by the

significance of the slopes. Even though the difference between groups was not significant for both low ($M = 69\%$, $SD = 21.5$ and $M = 63\%$, $SD = 21$ for the Control and ABST groups, respectively, $t[46] = -1.04$, $p = .304$) and high levels of self-efficacy ($M = 67\%$, $SD = 21$, and $M = 72\%$, $SD = 19$, for the Control and ABST groups, respectively, $t[54] = .870$, $p = .388$), it was less accentuated for high levels of self-efficacy, thus indicating that as compared to the pattern found for low levels of self-efficacy, the ABST group showed a tendency to surpass the Control group when self-efficacy was high.

4. Discussion

In this study, we found no significant differences between the ABST and the Control group regarding their performance in technology. This lack of the hypothesized difference could be attributed to several factors: One possible explanation is that the ABST group members, being highly educated and clearly interested in science, were motivated to perform better. Their desire to prove competence may have nullified the negative effects of our ABST manipulation. Another plausible explanation is that the ABST group engaged in effective self-regulation strategies to manage stress. By remaining calm, they may have maintained their performance levels despite the ABST manipulation (Vohs & Baumeister, 2004). We should not disregard the possibility that the manipulation did not work as intended. The stereotype that “older people have less technological ability” is particularly salient (Mariano et al., 2022). This could mean that simply participating in a study on this topic, even without an ABST manipulation, could lead to an ABST effect in the control group as well. However, we have evidence of a moderation in our data, which suggests that the manipulation was, in some respects, effective. Accordingly, the self-efficacy by group interaction showed a significant result. The cross-over interaction pattern for this interaction indicated a smaller difference between the ABST and the Control group for high levels of self-efficacy. This suggests that relative to the pattern observed for low self-efficacy levels, the ABST group tended to outperform the Control group when self-efficacy was high. This variability in performance across different levels of self-efficacy highlights the potential impact of individual self-efficacy beliefs on task performance outcomes. It suggests that individuals with higher self-efficacy might be more resilient to stereotype threat effects, potentially leading to better performance compared to those with lower self-efficacy levels.

5. Literature

- Lamont, R. A., Swift, H. J., & Abrams, D. (2015). A review and meta-analysis of age-based stereotype threat: Negative stereotypes, not facts, do the damage. *Psychology and Aging*, *30*(1), 180–193. <https://doi.org/10.1037/a0038586>
- Lamont, R. A., Swift, H. J. & Drury, L. (2021). Understanding perceived age-based judgement as a precursor to age-based stereotype threat in everyday settings. *Frontiers in Psychology*, *12*. <https://doi.org/10.3389/fpsyg.2021.640567>
- Levy, B. (2009). Stereotype embodiment. *Current Directions in Psychological Science*, *18*(6), 332–336. <https://doi.org/10.1111/j.1467-8721.2009.01662.x>
- Mariano, J., Marques, S., Ramos, M. R., Gerardo, F., da Cunha, C. L., Girenko, A., Alexandersson, J., Stree, B., Lamanna, M., Lorenzatto, M., Mikkelsen, L. P., Bundgård-Jørgensen, U., Rêgo, S., & de Vries, H. (2022). Too old for technology? Stereotype threat and technology use by older adults, *Behaviour & Information Technology*, *41*(7), 1503-1514, <https://doi.org/10.1080/0144929X.2021.1882577>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797-811. <https://doi.org/10.1037/0022-3514.69.5.797>
- Vohs, K. D., & Baumeister, R. F. (2004). Understanding self-regulation: *Handbook of self-regulation: Research, theory, and applications* (pp. 1-9). New York, NY: Guilford Press.

Young and distracted: Is it stereotype threat or stereotype challenge?

Elena Luna Schwieger; Julia Kemp; Maren Schneider; Luise Göttlich; Lilli Westphal

Advisor: M. Clara P. de Paula Couto

1. Introduction

Stereotypes are schemas that can influence our thinking, our perception, and our cognitive representation of social groups. For example, they can relate to age, being predominantly negative, like older adults are forgetful and young people are inexperienced. Age stereotypes can threaten individuals of different ages, affecting their performance. According to Steele and Aronson (1995), stereotype threat is the concern to confirm a negative stereotype when a situation arises in which a person has to complete a task in a stereotyped domain. This concern results in a performance that tends to be worse than would normally be the case. Age-Based Stereotype Threat (ABST) has been extensively studied in older adults, with only a few studies examining the effects of ABST in young people (Lamont et al., 2021). To address this gap, in this pre-registered study we investigated ABST among young people in the academic context. This is an understudied domain, which is however, especially important for young people while they pursue their educational path. We experimentally manipulated the stereotype that young people have a short attention span and examined the effect of making this stereotype salient on the cognitive performance of young adults. We hypothesized that young adults under ABST would show worse cognitive performance in an attention test compared to a control group. This ABST effect would further depend on self-efficacy, age identification, and domain importance (i.e., moderators) such that as compared to the control group, young people under ABST who reported higher self-efficacy, lower levels of identification with young people, and greater levels of domain importance would perform better in the attention test.

2. Methods

Participants. An a priori power analysis indicated a minimum of 102 participants to discover a small to medium effect ($d = .50$). We recruited 111 Psychology students of the FSU-Jena who were able to receive 0,5 VPN hours in return for their participation. We excluded seven participants who failed the attention check after the manipulation and another participant who did not comply with the attention test instructions. Hence, the final sample comprised 103 participants split into Control ($n = 53$, 87% female, $M_{\text{age}} = 21.1$, $SD = 1.97$) and ABST ($n = 50$, 84% female, $M_{\text{age}} = 21.4$, $SD = 1.78$).

Design. The design of this study was an experimental, cross-sectional design with groups varying between subjects. Accordingly, we randomly assigned participants to two groups: either ABST or Control.

Materials and Procedure. We used SoSci Survey to collect, manage, and extract data online. After providing informed consent and answering demographic and baseline behaviour questions (i.e., technology consumption during lectures and technology interference), participants were randomly assigned the either the Control or the ABST group. Using the Social Consensus Feedback (SCF) paradigm, we asked participants to reflect on others' agreement with ABST. The SCF is based on the idea that individuals typically feel more confident in their beliefs when they perceive that a majority shares these. Thus, presenting consensus information should lead individuals to align their beliefs with those of the majority to gain acceptance, attention, and support from others. Given previous evidence of changes in beliefs (Stagnor et al., 2001), the SCF is ideally suited to manipulate ABST. In the ABST group, the SCF indicated that most people believed that young people have difficulties concentrating for long periods due to technology use, whereas in the Control group, the SCF indicated that most people believed that the ability to concentrate does not necessarily depend on age but also on individual factors. We measured the cognitive performance through an attention test (a self-programmed online version of the *Aufmerksamkeits- und Konzentrationstest d2-R*). The adapted d2

included 14 blocks with 60 trials each. Participants had 20 seconds to complete each block with a 5-second pause in between blocks. We excluded the first and the last blocks to avoid warm-up and end-spurt (final sprint) effects. Performance was assessed by speed (i.e., number of items worked through) and corrected by omission and commission errors.

3. Results

To investigate our hypotheses of an effect of ABST on the cognitive performance of young adults, and whether this effect was moderated, we conducted a moderated multiple regression analysis in which we controlled for baseline behavior, and entered group, self-efficacy, age identification, domain importance, and their interaction with group. This analysis revealed the expected main effect of group, $B = -3.51$, $p = .016$, and a significant interaction between group and domain importance, $B = -5.11$, $p = .035$. To follow-up on the main effect of group, we carried out an independent t-test. A significant effect of ABST was found, $t(99) = 2.3$, $p = .023$, as the ABST group ($M = 34.6$, $SD = 7.41$) showed better performance than the Control group ($M = 31.1$, $SD = 8.01$). To examine the group \times domain importance interaction, we carried out simple slope analyses that revealed that domain importance was positively related to the performance on the d2 test for the ABST ($B = 3.18$, $p = .040$) but not for the control group ($B = -1.93$, $p = .135$).

4. Discussion

In this study, we found a significant ABST effect that indicated that young people under ABST performed better than those in the Control group did. Even though this finding is opposed to our prediction, it corresponds with the commonly found result that young people feel rather challenged than threatened by negative age-based stereotypes, which leads them to disconfirm the stereotype, hence showing a boosted performance compared to the control group (Hehman & Bugental, 2013; von Hippel et al., 2019). We also found a moderation by domain importance confirming previous findings that ABST effects are influenced by the value that individuals place on their performance in the stereotyped domain (Hess et al., 2003). Our study has limitations. First, the selected sample mainly consisted of Psychology students who may already know about the methods this study used or about age stereotypes and their effects. Thus, our sample may have a more reflected view of age-based stereotypes. This suggests that our manipulation may affect a more variable sample more than it affected our current sample. That is why a more representative sample of young people is needed to achieve less limited results within further research.

5. Literature

- Hehman, J. A. & Bugental, D. B. (2013). "Life stage-specific" variations in performance in response to age stereotypes. *Developmental Psychology*, 49(7), 1396–1406. <https://doi.org/10.1037/a0029559>
- Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. *The Journals of Gerontology. Psychological Sciences*, 58(1), P3–P11. <https://doi.org/10.1093/geronb/58.1.p3>
- Lamont, R. A., Swift, H. J., & Drury, L. (2021). Understanding perceived age-based judgement as a precursor to age-based stereotype threat in everyday settings. *Frontiers in Psychology*, 12, 640567. <https://doi.org/10.3389/fpsyg.2021.640567>
- Stangor, C., Sechrist, G. B., & Jost, J. T. (2001). Changing racial beliefs by providing consensus information. *Personality and Social Psychology Bulletin*, 27(4), 486–496. <https://doi.org/10.1177/0146167201274009>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811. <https://doi.org/10.1037//0022-3514.69.5.797>
- Von Hippel, C., Kalokerinos, E. K., Haantera, K. & Zacher, H. (2019). Age-based stereotype threat and work outcomes: Stress appraisals and rumination as mediators. *Psychology and Aging*, 34(1), 68–84. <https://doi.org/10.1037/pag0000308>

Can you feel the music or are you just listening? Differences in the emotions induced by music as a function of musicality and engagement

Albrecht Blochmann, Michal Duda, Lisa Fischer, Antonia Kaiser, Mirjam Schlinke, Amy Joy Zacharias

Supervision: Lars Rogenmoser

1. Introduction

For centuries, music has been an important means of expressing and perceiving emotions. Until today, the impact of both musicality and the engagement in music on the experienced emotions remains unexplored. With this study, we intended to investigate the effects of musical education and participation in music on the experienced emotions. For this, two groups, one with low musicality (i.e., non-musicians) and the other with high musicality (i.e., musicians), were recruited. Both groups were exposed to two different experimental conditions – listening and singing along to music – to gain insight into the effect of participation. We hypothesized that:

1. Both musicians and non-musicians experience a stronger emotional response when singing along to music compared to just listening to music.
2. Musicians have a stronger emotional response than non-musicians, both when listening to music and when singing along to music.
3. The difference in emotional responses between musicians and non-musicians is greater when singing along to music than when just listening to music.

2. Method

Random Sample: An a priori power analysis estimated that a mixed factorial ANOVA, with a sample size of $n \geq 52$ (26 per group) was sufficiently powered to detect effect sizes of $f \geq 0.2$. In this study, 55 participants (26 non-musicians & 29 musicians; gender: 46 female, 6 male, 2 diverse; age: $M = 21.89$, range 18 - 28) were randomly sampled. The inclusion criteria for the musicians were an active participation in music with a regularity of at least once a month for the past six months and a formal musical education with a minimum duration of five years. The inclusion criterion for the non-musicians was a lack of active participation in music for the last five years. Furthermore, it was compulsory for all participants to be native German speakers or possess at least a C1 level of German language proficiency. All participants were required to have no medical diagnosis of hearing difficulties. These criteria ensured a clear distinction between musicians and non-musicians and guaranteed that all participants could comprehend the test material. All participants were offered 0.75 VPH. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the FSU.

Design: To assign participants to the two groups, a questionnaire based on the Goldsmith Musical Sophistication Index was used (Müllensiefen et al., 2014). The two groups underwent two conditions, in which the emotional state while “listening to music” (condition 1: listening) and “participating in music” (condition 2: singing along) was captured. These conditions were presented in a randomized order and both consisted of the same four songs, with each condition lasting exactly two minutes. In order to ensure a variety in valence among our stimuli, we included two songs that are well-known in Germany (“99 Luftballons” by Nena (Offizielle Deutsche Charts, n.d.; NENA, 2021) & “Call me Maybe” by Carly Rae Jepsen (Offizielle Deutsche Charts, n.d.; CarlyRaeJepsenVEVO, 2012) and two self-selected favorites. After each musical stimulus the participants rated their emotional experience during the previous condition based on 9-point scales, according to the dimensions of

valence (1 = very negative, 9 = very positive), and arousal (1 = very calm, 9 = very arousing). Furthermore, familiarity was controlled using a binary scale (yes/no) to account for the possibility that some participants might be unfamiliar with the given songs.

3. Results

The repeated measures ANOVAs with both valence and arousal as dependent variables revealed significant main effects. Valence and arousal differed significantly as a function of engagement (valence: $F(1, 53) = 4.045$, $p = 0.049$; arousal, $F(1, 53) = 23.279$, $p < 0.001$), with the musicians revealing higher ratings for both. The interaction between the groups (musicality) and engagement was neither significant for valence, $F(1, 53) = 0.423$, $p = 0.518$, nor for arousal, $F(1, 53) = 2.939$, $p = 0.092$.

Independent sample t-tests revealed a significant difference in arousal between musicians and non-musicians in the singing condition, $t(53) = -2.015$, $p = 0.024$, 95%-CI[-1.28, -0.00], with musicians showing higher ratings. There were no significant group differences in the other three contrasts: singing – valence, $t(53) = -0.448$, $p = 0.328$, 95%-CI[0.78, 0.50], listening – valence, $t(53) = 0.194$, $p = 0.423$, 95%-CI[-0.55, 0.67] and listening – arousal, $t(53) = -0.081$, $p = 0.468$, 95%-CI[-0.75, 0.70].

Paired t-tests revealed no statistical difference in valence when comparing singing over listening among non-musicians, $t(25) = 0.852$, $p = 0.201$, whereas they exhibited significantly higher ratings in arousal in the singing condition, $t(25) = 1.981$, $p = 0.029$. Among musicians, the paired t-tests revealed significant engagement differences in both valence, $t(28) = 2.147$, $p = 0.020$ and arousal, $t(28) = 5.162$, $p < 0.001$, with higher ratings in the singing condition.

4. Discussion

The first hypothesis was supported, indicating that participants experience a stronger emotional response when singing along to music compared to just listening to music. However, the ANOVAs did not confirm the second and third hypotheses regarding the effects of musicality and an interaction with engagement. Nonetheless, post hoc analyses indicate a trend suggesting that emotional experience is enhanced by both the level of musical education and the level of engagement during music activities. Further research could benefit from implementing stricter criteria to distinguish between participants, such as including only professional musicians, which might reveal significant interactions. Additionally, aligning musical education more closely with the stimulus – considering that participants in this study primarily chose pop songs despite likely having backgrounds in classical music – could potentially strengthen observed effects. It would also be valuable to investigate whether musicians or non-musicians diverge more from a third reference group consisting of participants with moderate levels of musicality. This could be explored through regression analysis to better understand these distinctions.

5. Literature

- CarlyRaeJepsenVEVO. (2012, 1. März). *Carly Rae Jepsen - Call me maybe* [Video]. YouTube. <https://www.youtube.com/watch?v=fWNaR-rxAic>
- Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PloS One*, 9, e89642. <https://doi.org/10.1371/journal.pone.0089642>
- NENA. (2021, 24. April). *NENA | 99 Luftballons [1983] [Offizielles HD musikvideo]* [Video]. YouTube. <https://www.youtube.com/watch?v=Fpu5a0B18eY>
- Offizielle Deutsche Charts. (n.d.). *Nena 99 Luftballons*. Retrieved June 16, 2024, from <https://www.offiziellecharts.de/titel-details-922>
- Offizielle Deutsche Charts. (n.d.). *Carly Rae Jepsen Call me maybe*. Retrieved June 16, 2024, from <https://www.offiziellecharts.de/titel-details-933091>

“When music doesn’t matter” – Die deutsche Version des Barcelona Music Reward Questionnaire (dtBMRQ): Eine Validierungsstudie.

Madita Jodes, Charlotte Jolmes, Kim Meyer, Willi Moxter, Lena Rölz, Merle Roth

Leitung: Lars Roggenmoser

1. Einleitung

Musik ist für viele Menschen ein wichtiger Bestandteil ihres Alltags und erfüllt aufgrund ihrer belohnenden Wirkung vielfältige kognitive, soziale und emotionale Funktionen. Doch was genau empfinden Sie beim Hören von Musik? Können Sie sich vorstellen, dabei nichts zu fühlen?

Es gibt tatsächlich Menschen, die trotz normaler hedonistischer Reaktionen auf andere Sinnesreize unempfindlich gegenüber musikinduzierter Belohnung sind und dementsprechend unter musikalischer Anhedonie leiden (Belfi & Loui, 2019). Eine etablierte Methode zur Erfassung des Belohnungsempfindens gegenüber Musik ist der Barcelona Music Reward Questionnaire (BMRQ), ein Fragebogen von Mas-Herrero et al. (2013) mit 20 Items, die fünf Faktoren zugeordnet werden können. Dieser erfasst das Maß an Belohnungsempfinden in Verbindung mit Musik. Das Ziel dieser Studie ist es, eine deutsche Version des BMRQ zu entwickeln und zu validieren, um ein klinisches Instrument zur Erforschung der Belohnungssensitivität bereitzustellen.

Dafür überprüfen wir vorrangig die Faktorenstruktur des originalen BMRQ für den deutschen Sprachraum und integrieren außerdem die Skalen des behavioralen Hemmungs- und Aktivierungssystems (BIS/BAS) (Strobel et al., 2006) sowie einen Kurzfragebogen zur Erfassung der Big Five (Gosling et al., 2003) in die Studie, um Assoziationen zwischen musikalischer Anhedonie und weiteren Variablen exploratorisch zu untersuchen.

2. Methode

Der originale Fragebogen von Mas-Herrero et al. (2013) wurde mittels Vorwärts- und Rückwärts-Übersetzung von deutschen und englischen Muttersprachler:innen ins Deutsche überführt. Dabei zielten wir auf eine möglichst große Kongruenz beider Versionen ab. Die Erhebung erfolgte online über SoSci-Survey, wobei die Teilnehmenden insgesamt 54 Items beantworteten.

An unserer Studie nahmen $n = 471$ Proband:innen im Alter von 18 bis 71 Jahren ($M = 24.5$, $SD = 9.3$) teil, wobei 81.08% weiblich waren. Die erreichte Stichprobengröße überschritt das angestrebte Ziel, wodurch wir laut Bühner (2011) einen angemessenen bis hohen Effekt erreichen konnten. Alle Teilnehmenden sprachen Deutsch mindestens auf C2-Niveau und lebten in einem deutschsprachigen Land. Mehr als die Hälfte betrachtete sich im Bereich musikalischer Bildung als Amateur:in. Es gab keine zeitliche Begrenzung für das Ausfüllen der Fragebögen.

3. Ergebnisse

CFA für BMRQ. Die Daten wurden mit R-Studio (Version 2024.04.2+764) analysiert. Dabei wurden fehlende Werte entfernt, um die Datenqualität und Durchführbarkeit der konfirmatorischen Faktorenanalyse (CFA) sicherzustellen. Der Kaiser-Meyer-Olkin-Test lieferte uns einen akzeptablen Wert von 0.88. Für die CFA nahmen wir vorerst das Originalmodell von Mas-Herrero et al. (2013) an und erhielten folgende Fit-Indizes: $CFI = 0.847$, $RMSEA = 0.082$ und $SRMR = 0.072$. Da diese die Grenzwerte nicht vollständig erreichten, verbesserten wir das Modell mithilfe der von R vorgeschlagenen Modifikationsindizes und führten zusätzlich eine explorative Faktorenanalyse durch. Wir entschieden uns für die ökonomischste Modellanpassung und entfernten folglich vier Items mit unzureichenden Faktorladungen unter 0.45 (Hair et al., 2009). Aufgrund signifikanter Kreuzladung auf zwei Faktoren wurde Item 2 ebenso entfernt.

Modellanpassung. Nach erneuter CFA wurden verbesserte Fit-Indizes gefunden: CFI = 0.926, RMSEA = 0.067 und SRMR = 0.052. Die Varianzaufklärung mit unserem angepassten Modell beträgt 44,622 %. Die relevanten Faktorladungen der fünf Faktoren Emotionsevokation (EE), Sensomotorik (SM), Stimmungsregulation (SR), Musiksuche (MS) und Soziale Belohnung (SB) liegen zwischen 0.4 und 0.88.

Faktorenwerte und Gesamtwert. Die Berechnung der Faktorenwerte erfolgte gemäß dem von Ten Berge et al. (1999) vorgeschlagenen Verfahren. Die Mittelwerte der Faktorwerte liegen zwischen 21.56 (MS) und 41.49 (EE). Der gewichtete Gesamtwert beträgt $M = 29.33$ mit $SD = 4.40$. Mithilfe dieser Ergebnisse kann eine Kategorisierung der Daten in hedonisch und anhedonisch vorgenommen werden, wonach alle Werte, die mehr als eine Standardabweichung unterhalb des Gesamtscores liegen, als anhedonisch klassifiziert werden. Die absolute Anzahl der als musikalisch anhedonisch kategorisierten Proband:innen beträgt 74 von insgesamt 471 (15,71%).

Gütekriterien und Validierung. Wir ermittelten Cronbachs Alpha, die durchschnittliche extrahierte Varianz (AVE) und die zusammengesetzte Zuverlässigkeit (CR), wobei wir akzeptable Ergebnisse erhielten.

CFA für BIS/BAS. Für die zweite CFA übernahmen wir das Modell von Strobel et al. (2006). Die Berechnung der internen Konsistenz, der AVE-Werte und der CR-Werte der BISBAS-Faktoren lieferte insgesamt weniger zufriedenstellende Ergebnisse und konnte somit nicht für die Integration überzeugen.

4. Diskussion

Unser Ziel, eine deutsche Version des BMRQ zu erstellen, zu validieren und damit ein klinisches Instrument zu entwickeln, haben wir erreicht. Unsere Daten der BMRQ-Faktoren konnten die ursprüngliche Faktorstruktur von Mas-Herrero et al. (2013) nicht bestätigen, ähneln jedoch den Ergebnissen der chinesischen Validierung von Wang et al. (2023). Das angepasste Modell der BMRQ-Faktoren weist trotz hoher Selektivität und damit minderer Repräsentativität der Stichprobe akzeptable Gütekriterien auf. Es existieren bereits mehrere weitere Validationsstudien zum BMRQ für andere Sprachen, die jedoch mit denselben Problemen zu kämpfen hatten. Dennoch konnten wir mit unserem Modell eine höhere Varianzaufklärung als in den vergangenen Validierungen des BMRQ erzielt (vgl. Wang et al., 2023). Im Zusammenhang dazu konnten Unterschiede über verschiedene Kulturen hinweg, sowie interessante Unterschiede in den Geschlechtern festgestellt werden, die einer umfassenderen Untersuchung dienen könnten. Mit der Studie konnten wir die bestehende Forschung weiterbringen und einen Grundstein für die Untersuchung musikalischer Anhedonie im deutschsprachigen Raum legen. Weiterhin können auf unserer Studie aufbauende Erkenntnisse zur Erforschung der Belohnungssensitivität und damit dem Verständnis bestimmter Störungsbilder sowie deren Behandlungsmethoden beitragen.

5. Literatur

- Belfi AM, Loui P. Musical anhedonia and rewards of music listening: current advances and a proposed model. *Ann N Y Acad Sci.* 2020 Mar;1464(1):99-114. doi: 10.1111/nyas.14241. Epub 2019 Sep 23. PMID: 31549425.
- Bühner, M. (2011). Testtheoretische Grundlagen. In: M. Bühner (Hrsg.), Einführung in die Test- und Fragebogenkonstruktion, 3., aktualisierte und erw. Auflage, S. 20-44. Pearson Studium.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528. [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1)
- Hair, J. F. (2009). *Multivariate data analysis.* (7th ed.). Prentice Hall.
- Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R. J., & Rodriguez-Fornells, A. (2013). Barcelona Music Reward Questionnaire (BMRQ) [Database record]. *APA PsycTests.* <https://doi.org/10.1037/t31533-000>
- Strobel, A., Beauducel, A., Debener, S., & Brocke, B. (2006). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differentielle und Diagnostische Psychologie, 22*(3), 216. <https://doi.org/10.1024//0170-1789.22.3.216>
- Wang, J., Xu, M., Jin, Z., Xia, L., Lian, Q., Huyang, S., & Wu, D. (2023). The Chinese version of the Barcelona Music Reward Questionnaire (BMRQ): Associations with personality traits and gender. *Musicae Scientiae, 27*(1), 218–232. <https://doi.org/10.1177/10298649211034547>

Bist du umweltfreundlich oder tust du nur so?

Untersuchung der prädiktiven Power des impliziten Assoziationstests

Simone Dahmen, Jonas Krellwitz, Zoe Mack, Judy Schubert

Leitung: Merlin Urban

1. Einleitung

Neuere IAT-Forschung versucht anhand von Moderatoren zu identifizieren, unter welchen Umständen sich die prädiktive Power des impliziten Assoziationstests (IAT; Greenwald et al., 1998) erhöhen lässt. Urban et al. (2024) zeigten zum Beispiel, dass IATs mit mittlerer Testschwierigkeit (definiert als mittlere IAT-Effekte von null) mehr Varianz und somit eine bessere prädiktive Power haben als IATs mit extremer Testschwierigkeit (definiert als mittlere IAT-Effekte, die stark von null abweichen). IAT-Forschung zeigt außerdem, dass der mittlere IAT-Effekt (i.e., IAT-Testschwierigkeit) über die Valenz der Target-Stimuli beeinflusst werden kann. Wenn ein IAT aus einer positiven (z.B. Blumen) und einer negativen (z.B. Insekten) Target-Kategorie besteht und die Target-Stimuli in ihrer Valenz typisch sind, d.h. der Valenz der Target-Kategorien entsprechen (z.B. *Rose* für Blumen und *Wespe* für Insekten), dann ist der IAT Effekt stärker von null abweichend (i.e., extremere IAT-Testschwierigkeit) als wenn die Target-Stimuli in ihrer Valenz atypisch sind, d.h. nicht der Valenz der Target-Kategorien entsprechen (z.B. *Giftefeu* für Blumen und *Schmetterling* für Insekten; vgl. Govan & Williams, 2004). Wir wollen diese Befunde zum einen anhand anderer Target-Kategorien replizieren und zum anderen schauen, ob mit der Beeinflussung der mittleren IAT-Effekte, die von Urban et al. (2024) gefundenen Zusammenhänge einhergehen. Es lassen sich folgende zentrale Hypothesen ableiten: *H1*) Die IAT-Testschwierigkeit lässt sich über die Manipulation der Valenz der Target-Stimuli auf die beschriebene Weise beeinflussen. *H2*) Je extremer die IAT-Testschwierigkeit desto geringer die IAT-Varianz. *H3*) Je extremer die IAT-Testschwierigkeit und je geringer die IAT-Varianz, desto geringer die prädiktive Power des IATs.

2. Methode

Stichprobe. Nach Anwendung unserer Ausschlusskriterien betrug die finale Stichprobengröße 95 Proband*innen (Alter: $M = 22.18$; $SD = 3.19$; Studienfach: 94.74% Psychologiestudierende).

Design. Wir nutzten ein 2x2 between-Design mit den Faktoren *IAT-Typ* (IAT mit atypisch vs. IAT mit typisch valenten Target-Stimuli) und *Blockreihenfolge* (kompatibler vs. inkompatibler Block zuerst). Die Zuordnung zu den Bedingungen wurde randomisiert.

Ablauf. Die Proband*innen nahmen online teil. Nach einem Fragebogen zur Überprüfung der Repräsentativität der Stimuli für die Target-Kategorien folgte einer der beiden IATs. Im Anschluss folgte ein sogenannter Category-IAT, auf den im Rahmen dieses Kurzberichts aus Platzgründen nicht weiter eingegangen wird. Zuletzt wurden die vorherzusagenden Kriterien, zwei explizite Einstellungsmaße und ein Verhaltensmaß mit Fragebögen, erhoben.

Messinstrumente. Sowohl der atypische ($\alpha = .57$) als auch der typische IAT ($\alpha = .52$) wurden in Anlehnung an Greenwald et al. (1998) entworfen. Beide IATs hatten dieselben Targetkategorien (Umweltschutz/Umweltschädigung), dieselben Attributkategorien (gut/schlecht), dieselbe Anzahl an Stimuli (5 Stimuli pro Kategorie) sowie dieselben Attributstimuli (z.B. *toll* für gut und *mies* für schlecht). Sie unterschieden sich lediglich in den verwendeten Target-Stimuli (typischer IAT: z.B. *Artenvielfalt* für Umweltschutz und *Abholzung* für Umweltschädigung; atypischer IAT: z.B. *Konsumverzicht* für Umweltschutz und *Globalisierung* für Umweltschädigung). Die IAT-Effekte wurden anhand des D-Score Algorithmus berechnet. Das erste Einstellungsmaß (Stimulusvalenzmaß; $\alpha = .8$) basierte auf den Bewertungen der in beiden IATs verwendeten Target-Stimuli. Das zweite Einstellungsmaß (Gefühlsmaß; $\alpha = .63$) basierte auf Bauch- und tatsächlichen

Gefühlen gegenüber den Target-Kategorien. Das Verhaltensmaß ($\alpha = .39$) erfasste die Häufigkeit von umweltschützendem oder -schädigendem Verhalten.

3. Ergebnisse

H1: Wie angenommen wurde der Haupteffekt IAT-Typ in einer zweifaktoriellen ANOVA signifikant, $F(1, 91) = 27.13, p < .0001, \eta_{\text{part}}^2 = .23$, während weder der Haupteffekt Blockreihenfolge signifikant wurde, $F(1, 91) = 2.83, p = .09, \eta_{\text{part}}^2 = .03$, noch die Interaktion der beiden Faktoren, $F(1, 91) = 0.88, p = .35, \eta_{\text{part}}^2 = .09$. Unabhängig davon, ob mit dem kompatiblen oder dem inkompatiblen Block begonnen wurde, war der mittlere D-Score des atypischen IATs näher an null ($M_{\text{komp}} = 0.40, M_{\text{inkomp}} = 0.53$) als der typische IAT ($M_{\text{komp}} = 0.69, M_{\text{inkomp}} = 0.72$) und hatte somit eine weniger extreme Testschwierigkeit. *H2*: Entgegen *H2* unterschieden sich die IAT-Varianzen des atypischen ($\text{Var} = 0.054$) und typischen IATs ($\text{Var} = 0.047$) im Levene-Test nicht signifikant, $F(1, 93) = .21, p = .64$. *H3*: Entgegen *H3* waren weder die Korrelationen des atypischen noch die des typischen IATs mit den Kriteriumsmaßen signifikant von null verschieden (siehe Tabelle 2). Z-Tests für unabhängige Korrelationen zeigten zudem, dass sich der atypische und der typische IAT in ihren Korrelationen mit den Kriteriumsmaßen nicht signifikant voneinander unterschieden (Valenzmaß: $z = 0.03, p = .49$; Gefühlsmaß: $z = 0.01, p = .5$; Verhaltensmaß: $z = 0.9, p = .18$).

Tabelle 1

Korrelationen der IATs mit den expliziten Maßen.

	Valenzmaß	Gefühlsmaß	Verhaltensmaß
Atypischer IAT	.06	.12	-.15
Typischer IAT	.07	.12	.04

Anmerkung. * Signifikanz auf α -Fehler-Niveau 5% gekennzeichnet.

4. Diskussion

Bzgl. der *H1* zeigte sich, dass der atypische IAT einen kleineren IAT-Effekt hatte als der typische IAT und somit wie angenommen eine weniger extreme Testschwierigkeit aufwies. Unser Ergebnis bzgl. *H2*, dass sich die IAT-Varianzen des atypischen und typischen IATs nicht signifikant unterschieden, steht im Gegensatz zu Urban et al. (2024) und der Fragebogenforschung, laut welcher Fragebögen mit extremerer Testschwierigkeit eine geringere Varianz haben (Schmidt-Atzert & Amelang, 2012). Das könnte an der kontinuierlichen Skala der D-Scores liegen, da extremere Mittelwerte auf der Skala im Vergleich zu diskreten Fragebogenskalen mit eingeschränktem Antwortbereich nicht mit einer schiefen Verteilung einhergehen müssen, sondern sich die Verteilung lediglich verschieben kann. Entgegen *H3* unterschieden sich die Korrelationen des atypischen und typischen IATs mit den Kriteriumsmaßen nicht signifikant, was auf die fehlende Beeinflussung der IAT-Varianzen zurückgeführt werden kann. Die geringe Reliabilität unserer IATs sowie die Homogenität der Stichprobe und die damit einhergehende Varianzeinschränkung all unserer Maße könnten für die insgesamt überraschend niedrigen Korrelationen verantwortlich sein.

5. Literatur

- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Govan, C. L. & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. *Journal Of Experimental Social Psychology*, 40(3), 357–365. <https://doi.org/10.1016/j.jesp.2003.07.002>
- Schmidt-Atzert, L. & Amelang, M. (2012). Psychologische Diagnostik (Lehrbuch mit Online-Materialien). Springer Science & Business Media.
- Urban, M., Koch T., & Rothermund K. (in press). The Implicit Association Test and its Difficulty(ies): Introducing the Test Difficulty Concept to Increase the True-Score Variance and, Consequently, the Predictive Power of Implicit Association Tests. *Journal of Personality and Social Psychology*.

Lässt sich die Testschwierigkeit und somit die prädiktive Power des impliziten Assoziationstests über die Target-Stimuli beeinflussen?

Barbara Gräser, Magdalena Klodt, Conrad Opitz, Annika Thiel

Leitung: Merlin Urban

1. Einleitung

In den letzten Jahren haben sich viele Forschungsarbeiten zum impliziten Assoziationstest (IAT; Greenwald et al., 1998) damit beschäftigt, Moderatoren zu identifizieren, um die prädiktive Power des IATs besser zu verstehen. So fanden beispielsweise Urban et al. (2024) heraus, dass IATs mit mittlerer Testschwierigkeit (definiert als mittlere IAT-Effekte von null) mehr Varianz und somit eine bessere prädiktive Power haben als IATs mit extremer Testschwierigkeit (definiert als mittlere IAT-Effekte, die stark von null abweichen). Außerdem zeigten Forschungsarbeiten zum IAT, dass der mittlere IAT-Effekt (i.e., IAT-Testschwierigkeit) über die Valenz der Target-Stimuli beeinflusst werden kann. Besteht ein IAT aus einer positiven (z.B. Blumen) und einer negativen (z.B. Insekten) Target-Kategorie und die Target-Stimuli sind in ihrer Valenz typisch, d.h. der Valenz der Target-Kategorien entsprechend (z.B. positiver Stimulus *Tulpe* für Blumen und negativer Stimulus *Raupe* für Insekten), dann ist der IAT-Effekt stärker von null abweichend (i.e., extremere IAT-Testschwierigkeit) als wenn die Target-Stimuli in ihrer Valenz atypisch sind, d.h. nicht der Valenz der Target-Kategorien entsprechen (z.B. negativer Stimulus *Giftefeu* für Blumen und positiver Stimulus *Libelle* für Insekten; vgl. Govan & Williams, 2004). Einerseits wollen wir diesen Befund anhand anderer Target-Kategorien replizieren, andererseits wollen wir schauen, ob mit der Beeinflussung der mittleren IAT-Effekte, die von Urban et al. (2024) gefundenen Zusammenhänge einhergehen. Daraus leiten sich folgende Hypothesen ab: *H1*) Die IAT-Testschwierigkeit lässt sich über die Manipulation der Valenz der Target-Stimuli auf die beschriebene Weise beeinflussen. *H2*) Je extremer die IAT-Testschwierigkeit desto geringer die IAT-Varianz. *H3*) Je extremer die IAT-Testschwierigkeit und je geringer die IAT-Varianz, desto geringer die prädiktive Power des IATs.

2. Methode

Stichprobe. An unserer Studie nahmen 86 Proband*innen teil (Alter: $M = 21.76$, $SD = 2.73$; Studienfach: 95.35% Psychologiestudierende).

Design. Die Studie hatte ein 2x2 mixed Design mit dem within Faktor *Stimulusset* (Stimulusset mit atypisch vs. Stimulusset mit typisch valenten Target-Stimuli) und dem between Faktor *Blockreihenfolge* (Beginn mit kompatiblen vs. inkompatiblen Block). Die Zuordnung zu den Bedingungen wurde randomisiert.

Ablauf. Die Proband*innen nahmen online teil. Nach einem Fragebogen zur Überprüfung der Repräsentativität der Stimuli für die Target-Kategorien folgte der IAT. Zuletzt wurden die vorherzusagenden Kriterien, zwei explizite Einstellungsmaße und ein Verhaltensmaß erhoben.

Messinstrumente. Der IAT wurde in Anlehnung an Greenwald et al. (1998) entworfen. Er hatte die Target-Kategorien *Umweltschutz/Umweltschädigung* und die Attributkategorien *gut/schlecht*. Es gab sechs Stimuli pro Attributkategorie (z.B. *herrlich* für gut und *schrecklich* für schlecht) und sechs Stimuli pro Target-Kategorie, wobei jeweils drei Stimuli typisch und drei Stimuli atypisch in ihrer Valenz waren (typisches Stimulusset: z.B. *Solarenergie* für Umweltschutz und *Vermüllung* für Umweltschädigung; atypisches Stimulusset: z.B. *Flugzeugverbot* für Umweltschutz und *Klimaanlage* für Umweltschädigung). Die IAT-Effekte wurden anhand des D-Score Algorithmus berechnet, wobei lediglich die Target-Stimuli in die Berechnung einfließen. Das atypische Stimulusset hatte eine interne Konsistenz von $\alpha = .46$ und das typische Stimulusset von $\alpha = .45$. Das erste Einstellungsmaß (Stimulusvalenzmaß; $\alpha = .77$) basierte auf den Bewertungen aller verwendeten Target-Stimuli. Das zweite Einstellungsmaß (Gefühlsmaß; $\alpha = .57$) basierte auf Bauch-

und tatsächlichen Gefühlen gegenüber den Target-Kategorien. Das Verhaltensmaß ($\alpha = .22$) erfasste, die Häufigkeit von umweltschützendem oder -schädigendem Verhalten.

3. Ergebnisse

H1: Wie angenommen wurde der Haupteffekt Stimulusset in einer zweifaktoriellen mixed ANOVA signifikant, $F(1, 84) = 49.49$, $p < .0001$, $\eta_{part}^2 = .37$, während weder der Haupteffekt Blockreihenfolge signifikant wurde, $F(1, 84) = 1.44$, $p = .23$, $\eta_{part}^2 = .02$, noch die Interaktion der beiden Faktoren, $F(1, 84) = 0.29$, $p = .59$, $\eta_{part}^2 = .00$. Unabhängig davon, ob mit dem kompatiblen oder dem inkompatiblen Block begonnen wurde, war der mittlere D-Score des atypischen Stimulussets näher an null ($M_{komp} = 0.41$, $M_{inkomp} = 0.51$) als der des typischen Stimulussets ($M_{komp} = 0.66$, $M_{inkomp} = 0.73$), sodass das atypische Stimulusset eine weniger extreme Testschwierigkeit hatte als das typische Stimulusset. *H2*: Entgegen *H2* zeigte ein Levene-Test, dass die IAT-Varianz des atypischen Stimulussets ($\hat{\sigma} = 0.118$) nicht signifikant höher war als die des typischen Stimulussets ($\hat{\sigma} = 0.123$), $F(1, 170) = .07$, $p = .80$. *H3*: Entgegen *H3* waren weder die Korrelationen des atypischen noch die des typischen Stimulussets mit den Kriteriumsmaßen signifikant von null verschieden (siehe

Tabelle 1

Korrelationen der Stimulussets mit den Kriteriumsmaßen

	Stimulusvalenzmaß	Gefühlsmaß	Verhaltensmaß
Atypisches Stimulusset	.03	.07	-.05
Typisches Stimulusset	.02	-.11	.01

Anmerkung. * $p < .05$.

Tabelle 1). Z-Tests für abhängige Korrelationen zeigten, dass sich das typische und atypische Stimulusset in ihren Korrelationen mit dem Stimulusvalenzmaß ($t = 0.08$, $p = .47$) und dem Verhaltensmaß ($t = 0.65$, $p =$

.26) nicht signifikant, aber mit dem Gefühlsmaß ($t = 1.86$, $p = .03$) signifikant unterschieden.

4. Diskussion

Entsprechend *H1* und in Einklang mit bisheriger Literatur (vgl. Govan & Williams, 2004) ließ sich die Testschwierigkeit insofern verändern, als dass das atypische Stimulusset einen geringeren mittleren IAT-Effekt hatte als das typische Stimulusset, und somit eine weniger extreme IAT-Testschwierigkeit aufwies. Entgegen *H2* konnten wir keine signifikanten Varianzunterschiede in den Stimulussets finden. Das hängt vermutlich mit der kontinuierlichen Skala der D-Scores zusammen, da extremere Mittelwerte auf der Skala im Vergleich zu z.B. diskreten Fragebogenskalen mit eingeschränktem Antwortbereich nicht mit einer schiefen Verteilung einhergehen müssen, sondern sich die Verteilung, wie in unserem Fall, lediglich verschieben kann. In Bezug auf *H3* hatten wir eine erhöhte prädiktive Power für das atypische Stimulusset erwartet. Aufgrund dessen, dass die IAT-Varianz aber nicht erhöht war, sind fehlende Korrelationsunterschiede nicht überraschend. In Bezug auf das Gefühlsmaß zeigten sich zwar signifikante Unterschiede, aber die Korrelation des atypischen Stimulussets war dennoch nicht signifikant von null verschieden und die prädiktive Validität somit weiterhin eingeschränkt. Außerdem waren die Korrelationen generell sehr niedrig, was auf die niedrigen IAT-Reliabilitäten oder die Homogenität der Stichprobe und eine damit verbundene Varianzeinschränkung in allen Maßen zurückzuführen sein könnte.

5. Literatur

- Govan, C. L. & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. *Journal Of Experimental Social Psychology*, 40(3), 357–365. <https://doi.org/10.1016/j.jesp.2003.07.002>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037//0022-3514.74.6.1464>
- Urban, M., Koch T., & Rothermund K. (in press). The Implicit Association Test and its Difficulty(ies): Introducing the Test Difficulty Concept to Increase the True-Score Variance and, Consequently, the Predictive Power of Implicit Association Tests. *Journal of Personality and Social Psychology*.

Worte, die verbinden: Kann Selbstoffenbarung Beziehungen vertiefen?

Anna Dührsen, Fenja König, Clara Elia Makrinus, Kira Wagner, Melina Zäh

Leitung: Jenny Jaquet

1. Einleitung

Zwischenmenschliche Beziehungen und soziale Anbindung sind essentielle Bestandteile eines erfüllten Lebens. Besonders enge und vertrauensvolle Beziehungen begleiten Menschen meist sogar über die gesamte Lebensspanne hinweg (Kahn & Antonucci, 1980). Doch wodurch entstehen nahe und vertrauensvolle Beziehungen, insbesondere wenn sich zwei Personen zuvor fremd sind? Als eine wichtige Komponente wird z.B. ein wechselseitiger emotionaler Austausch genannt (Rimé et al. 2020). Wenn wir uns öffnen, unser Erlebtes anderen Menschen mitteilen und uns ihnen anvertrauen, scheint es fast so, als würden wir ein unsichtbares Band der Vertrautheit weben. Schlichtweg wäre ein Leben ohne Kommunikation und ohne emotionalen Austausch nur schwer vorstellbar.

Aufgrund der aufgezeigten Signifikanz und Allgegenwärtigkeit dieser Thematik haben wir die Zusammenhänge zwischen dem sozialen Teilen von Emotionen (Social Sharing of Emotions, kurz: SSE) einer Person und der durch ihre*n Gesprächspartner*in wahrgenommene Sympathie genauer betrachtet. Basierend auf dem „Interpersonal Process Model of Intimacy“ (Reis & Shaver, 1988), welches annimmt, dass die Preisgabe selbstrelevanter Informationen (z.B. zu emotionalen Zuständen) ein wichtiger Bestandteil für die Entstehung von Nähe ist, haben wir die Hypothese abgeleitet, dass diese Korrelation positiv ist, d.h. je häufiger eine Person Emotionen teilt, umso sympathischer wird diese durch ihr Gegenüber wahrgenommen (H1). Des Weiteren legten wir den Fokus bei der Beantwortung der Fragestellung auf eine entwicklungspsychologische Perspektive und interessierten uns zudem dafür, inwieweit dieser Zusammenhang altersdifferentiell ist, d.h. sich abhängig vom Alter der Interaktionspartner*innen unterscheidet. Diesbezüglich stützen wir uns auf die Überlegungen von Carstensen (2006), welche in der von ihr aufgestellten *Sozioemotionalen Selektivitätstheorie* annimmt, dass sich die Gestaltung sozialer Beziehungen und Motive über die Lebensspanne hinweg verändert.

So lässt sich aus der Theorie ableiten, dass jüngere Erwachsene ein höheres Interesse an neuen Beziehungen haben als ältere Erwachsene, welche auf Grund einer subjektiv kürzer wahrgenommenen Zeitperspektive eher auf bereits bestehende und vertrauensvolle Beziehungen fokussieren. In Kontexten, in denen fremde Personen miteinander interagieren, erwarten wir, dass ältere Erwachsene weniger empfänglich für die Effekte des sozialen Teilens (als eine Form von Investition in die Beziehung) sind, da sie weniger motiviert sind, neue Beziehungen zu knüpfen. Unsere zweite Hypothese beinhaltet deshalb, dass der Effekt des sozialen Teilens von Emotionen auf die vom Gesprächspartner wahrgenommene Sympathie größer ist, wenn die zuhörende Person jung ist (H2).

2. Methode

Unsere Untersuchung beruht auf Daten der Gesprächsstudie von Blanke et al. (2015). Die Stichprobe besteht aus 208 Studienteilnehmerinnen, welche sich in 102 jüngere (20 - 31 Jahre) und 106 ältere (69 - 80 Jahre) Frauen unterteilen lassen. Daraus ergeben sich drei Gruppen von Paaren, 52 jung-alte, 27 alt-alte und 25 jung-junge Dyaden.

Ein zentraler Bestandteil der Studie war es, dass die Dyaden ein Gespräch führten, in dem sie sich abwechselnd zuerst von einem negativen, dann von einem positiven Thema berichteten. Diese Gespräche wurden auf Video aufgezeichnet und transkribiert. Jene Transkripte bilden die Datengrundlage für diese Studie. Innerhalb der Gesprächstranskripte wurde jeder Hauptsatz danach untersucht, ob darin Emotionen geteilt wurden (ja = 1) oder nicht (nein = 0) (basierend auf bestehendem Kodiermanual; Rauers, 2020). Je nach Valenz des Emotionsausdrucks wurden

unterschiedliche Codes vergeben: „SSE+“ für positive emotionale Ausdrücke, „SSE-“ sofern negative emotionale Ausdrücke vorlagen und „SSE?“ bei gemischten Emotionen. Nach Vergabe der Codes berechneten wir einen Summenscore aller SSE unabhängig der Valenz für jeden der beiden Interakteure. Die abhängige Variable, wahrgenommene Sympathie, wurde mittels Selbstbericht erfasst. Die zugrundeliegende Skala besteht aus 6 Items, die auf einer 7-Punkte-Likert-Skala von 0 (= stimme überhaupt nicht zu) bis 6 (= stimme voll und ganz zu) bewertet wurden. Ein beispielhaftes Item ist „Ich habe meine Gesprächspartnerin angenommen und respektiert“. Abschließend wurde der Mittelwert der einzelnen Items errechnet.

3. Ergebnisse

Um unsere erste Hypothese zu prüfen, nutzten wir eine einfache lineare Regression. Dabei zeigte sich, dass die Häufigkeit der geteilten Emotionen kein signifikanter Prädiktor für die vom Gesprächspartner berichtete Sympathie war, $b < 0.01$, $t(206) = 0.02$, $p = .985$, 95% BCa KI [-0.02, 0.02], $R^2 = .000$.

Eine weiterführende Moderationsanalyse mittels multipler Regression zeigte zudem, dass sich der Effekt nicht danach unterschied, ob die zuhörende Person jung versus alt war; Interaktionseffekt: $b < 0.01$, $t(204) = 0.22$, $p = .825$, BCa KI [-0.03, 0.04].

Anschließend wurde das multiple Regressionmodell unter Ausschluss des Interaktionseffekt erneut aufgesetzt. Hierbei zeigte sich ein signifikanter Haupteffekt der Altersgruppe; ältere Frauen berichteten signifikant mehr Sympathie ihrer Gesprächspartnerin gegenüber verglichen mit jüngeren Frauen ($b = 0.39$, $t(205) = 3.48$, $p = .001$, BCa KI [-0.18, 0.60], $M_{alt} = 5.20$, $M_{jung} = 4.81$).

4. Diskussion

Beide Hypothesen müssen verworfen werden. Wir schließen daraus, dass, zumindest in unserer Stichprobe, kein Zusammenhang zwischen der Häufigkeit geteilter Emotionen und der eingeschätzten Sympathie besteht und dieser Effekt auch nicht vom Alter des Gegenübers abhängig ist. Es lässt sich vermuten, dass die Befunde darauf zurückgehen, dass recht wenig Varianz in der empfundenen Nähe vorlag, denn fast alle Teilnehmerinnen gaben an, sich ihrem Gegenüber nah gefühlt zu haben. Kaum jemand äußerte, das Gefühl gehabt zu haben, sich seiner Partnerin nicht nah zu fühlen. Somit kann die eben genannte Varianzeinschränkung eine mögliche Ursache für die gezeigten Ergebnisse darstellen.

Trotz allem konnten wir in unseren Daten jedoch einen Effekt des Alters auf die wahrgenommene Sympathie feststellen, welcher sich dadurch äußerte, dass sich ältere Frauen näher fühlten, als es junge taten. Möglicherweise ist dieser Effekt auch durch das Studiendesign und die Rekrutierungsstrategie bedingt. So ist die finanzielle Vergütung der Teilnahme vor allem ansprechend für jüngere Erwachsene unabhängig von der sozialen Motivation, während die älteren Erwachsenen, die an solchen Studien teilnehmen, dies explizit tun, um andere Menschen kennen zu lernen. Dies könnte sich dann auch in Bewertungen der Interaktionen widerspiegeln.

5. Literatur

- Blanke, E. S., Rauers, A., & Riediger, M. (2015). Nice to meet you—adult age differences in empathic accuracy for strangers. *Psychology and Aging*, *30*(1), 149–159. <https://doi.org/10.1037/a0038459>
- Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science*, *312*(5782), 1913–1915. <https://doi.org/10.1126/science.1127488>
- Kahn, R. L., & Antonucci, T. C. (1980). Convoys over the life course: Attachment, roles, and social support. In P. B. Baltes & O. G. Grim (Eds.), *Life span development and behavior* (Vol. 3, pp. 253–286). Academic Press.
- Reis, H. T., & Shaver, P. R. (2018). Intimacy as an interpersonal process. In Routledge eBooks (pp. 113–143). <https://doi.org/10.4324/9780203732496-5>
- Rimé, B., Bouchat, P., Paquot, L., & Giglio, L. (2020). Intrapersonal, interpersonal, and social outcomes of the social sharing of emotion. *Current Opinion in Psychology*, *31*, 127–134. <https://doi.org/10.1016/j.copsyc.2019.08.024>

Do we feel worse together? Zusammenhänge zwischen negativem Affekt und dem Teilen von Emotionen in Partnerschaften

Nina Brünkmann, Fynn Kornmann, Sandra Büttner, Julia Nassua, Friederike Peters

Leitung: Jenny Jaquet

1. Einleitung

Emotionen teilen – eine universelle Praxis, die tief in der menschlichen Erfahrung verwurzelt ist. Das soziale Teilen von Emotionen stellt eine häufig genutzte Strategie zur Bewältigung emotionaler Erlebnisse dar (Rimé et al., 1991). Obwohl dieses Teilen eine weitverbreitete Praxis ist (Rimé et al., 1991), zeigen frühere Forschungen, dass das Teilen negativer Emotionen die Stimmung der beteiligten Personen negativ beeinflussen kann. Man nimmt an, dass entsprechende negative Emotionen durch das Teilen reaktiviert (d.h. erneut erlebt) werden (Rimé, 2020).

Vor diesem Hintergrund wurde im empirischen Forschungsseminar untersucht, ob und wie das soziale Teilen negativer Emotionen mit dem momentanen Affekt der beteiligten Personen zusammenhängt und ob es dabei Unterschiede zwischen verschiedenen Altersgruppen gibt. Dabei wurde sich auf die folgenden Hypothesen konzentriert: Mehr soziales Teilen negativer Emotionen ist sowohl mit einem höheren momentanen negativen Affekt der Sprechenden (H1a) als auch der Zuhörenden Person (H1b) unmittelbar nach dem Gespräch assoziiert. So kann argumentiert werden, dass das soziale Teilen durch Mechanismen wie emotionale Ansteckung und Empathie möglicherweise nicht nur Konsequenzen für den Affekt der Person selbst, sondern auch für deren Gegenüber hat (Rauers & Riediger, 2023). Weiterhin nehmen wir an, dass der Zusammenhang zwischen dem Teilen negativer Emotionen und negativem Affekt bei älteren Sprechenden (H2a) und älteren Zuhörenden (2b) schwächer ausgeprägt ist als bei jüngeren.

Für die Hypothesen 2a und 2b lässt sich aus der Emotionsregulationstheorie von McRae und Gross (2020) ableiten, dass ältere Erwachsene effektivere Strategien zur Emotionsregulation anwenden als jüngere. Beispielhaft ist die Strategie der kognitiven Neubewertung, bei der ältere Erwachsene negative Ereignisse eher umdeuten und positivere Aspekte betonen. Dies geht auch d'accord mit den Annahmen des Strength and Vulnerability Integration Models (SAVI; Charles, 2010), welches besagt, dass ältere Erwachsenen Laufe ihres Lebens bessere Emotionsregulationsstrategien entwickelt haben, die sie befähigen, positive Emotionen zu maximieren und negative zu minimieren.

2. Methode

Unsere Untersuchung basiert auf den Daten von Rohr et al. (2019), die 156 Gesprächstranskripte von 78 heterosexuellen Paaren umfassen, bestehend aus 37 jüngeren Paare ($M = 24$ Jahre, $SD = 3.93$) und 41 älteren Paaren ($M = 41$ Jahre, $SD = 5.44$). Jede Dyade führte zwei Gespräche, in denen sie gebeten wurden, emotionale Erlebnisse zu rekapitulieren. Beide waren einmal in der Rolle der Sprechenden und der Zuhörenden Person. Alle Gespräche der Paare wurden aufgezeichnet und anschließend transkribiert. Das soziale Teilen von Emotionen (social sharing of emotion (SSE)) ist die unabhängige Variable, welche basierend auf einem bestehenden Kodiermanual (Rauers, 2020) neu kodiert wurde. Hauptsätze bildeten dabei die Analyseeinheit. Pro Einheit wurde bewertet, ob negative Emotionen geteilt wurden (ja = 1, nein = 0). Jedes Transkript wurde doppelt kodiert, wobei bei unterschiedlichen Codes in einem Meeting der kodierenden Personen ein Konsens getroffen wurde. Die Interrater-Reliabilität der kodierenden Personen ist als gut zu bewerten ($ICC = .84$). Die abhängige Variable ist der momentane negative Affekt, der mit 11 negativen Emotionswörtern (z.B. verärgert) mit einer Likert-Skala von 1 (gar nicht) - 5 (äußerst), unmittelbar nach dem Gespräch erfasst wurde ($\alpha = .81$).

3. Ergebnisse

Zur Untersuchung der Hypothesen 1a und 1b wurden lineare Regressionsmodelle genutzt. Hierbei zeigte sich weder ein signifikanter Zusammenhang zwischen der Häufigkeit der geteilten negativen Emotionen und dem negativen Affekt der sprechenden Person ($b = 0.02$, $t(154) = 1.57$; $p = .118$) noch mit jenem der zuhörenden Person ($b < 0.01$, $t(154) = 0.31$, $p = .759$).

Für die Hypothesen 2a und 2b wurden Moderationsanalysen mittels multipler Regressionen gerechnet. Betrachtet wurde jeweils der Interaktionseffekt von Altersgruppe und Häufigkeit der negativen geteilten Emotionen. Dieser Effekt war für beide sowohl für den negativen Affekt der sprechenden Person ($b < 0.01$, $t(152) = -0.01$, $p = .989$) als auch für jenen der zuhörenden Person ($b < 0.04$, $t(152) = 1.56$, $p = .122$) nicht signifikant.

Weiterführende Analysen beinhalteten die Schätzung der Modelle ohne Interaktionsterm. Hierbei zeigt sich ein signifikanter Haupteffekt der Altersgruppe auf den berichteten negativen Affekt der sprechenden Person ($b = -0.20$, $t(153) = -1.98$, $p = .049$) – ältere Personen berichteten signifikant weniger negativen Affekt als jüngere Personen. Zudem zeigte sich in explorativen Analysen ein Effekt der Altersgruppe auf die Häufigkeit der geteilten negativen Emotionen – ältere Erwachsene teilten signifikant weniger als jüngere Erwachsene ($b = -2.59$, $t(154) = -4.15$, $p < .001$).

4. Diskussion

Konträr zu bisherigen Befunden, konnte keine Evidenz für einen positiven Zusammenhang zwischen dem Teilen negativer Emotionen und dem darauffolgenden Ausmaß des negativen Affekts gefunden werden. Genauso verhält es sich mit altersdifferentiellen Effekten. Die aufgestellten Hypothesen konnten somit nicht belegt werden.

Der in der explorativen Analyse gefundene Effekt der Altersgruppe auf sowohl den negativen Affekt als auch das tatsächliche Teilen von negativen Emotionen, stellt jedoch einen spannenden Befund dar, der mit bestehender Literatur zum Umgang älterer Erwachsener mit negativen Gefühlen vereinbar ist. So ist ein möglicher Schluss, dass ältere Erwachsene generell weniger negative Emotionen zum Ausdruck bringen, weder in Form eines Selbstberichts noch tatsächlich im beobachtbaren Verhalten. Ursache dafür könnte der sogenannte Positivitätseffekt (Carstensen & Mikels, 2005) sein – eine Fokussierung auf positive anstelle von negativen Inhalten mit dem Ziel, das eigene emotionale Wohlbefinden im Alter aufrechtzuerhalten.

Kritisch zu betrachten ist, dass die Stichprobe ausschließlich heterosexuelle Paare umfasst. Bezüglich der Generalisierbarkeit der Ergebnisse sollte die zukünftige Forschung diversere Stichproben berücksichtigen. Zudem wäre es sinnvoll, unterschiedliche Traits und Fähigkeiten der Person, wie Emotionsregulationsstrategien, als Moderationsvariablen zu berücksichtigen. Zuletzt ist anzumerken ist, dass die nicht-dyadische Analysestrategie zur Verzerrung der Schätzungen geführt haben könnte. Für weitere Studien ist also eine dyadischen Analysestrategie beachtenswert.

5. Literatur

- Carstensen, L. L., & Mikels, J. A. (2005). At the Intersection of Emotion and Cognition: Aging and the Positivity Effect. *Current Directions in Psychological Science*, 14(3), 117–121. <https://doi.org/10.1111/j.0963-7214.2005.00348.x>
- Charles, S. T. (2010). Strength and Vulnerability Integration: A Model of Emotional Well-Being Across Adulthood. *Psychological Bulletin*, 1068-1091.
- McRae, K., & Gross, J. J. (2020). *Emotion regulation*. *Emotion*, 20(1), 1–9. <https://doi.org/10.1037/emo0000703>.
- Rauers, A. (2020). *Coding social sharing of emotion in transcribed conversations*. *Coding manual*. <https://doi.org/https://osf.io/6pcvx/>
- Rauers, A., & Riediger, M. (2023). Ease of mind or ties that bind? Costs and benefits of disclosing daily hassles in partnerships. *Social Psychological and Personality Science*, 14(5), 551-561.
- Rimé, B., Mesquita, B., Boca, S., & Philippot, P. (1991). Beyond the emotional event: Six studies on the social sharing of emotion. *Cognition & Emotion*, 5(5-6), 435-465.
- Rimé, B., Bouchat, P., Paquot, L., & Giglio, L. (2020). Intrapersonal, interpersonal, and social outcomes of the social sharing of emotion. *Current opinion in psychology*, 31, 127-134.
- Rohr, M. K., Nestler, S., Kunzmann, U. (2019). A trouble shared is a trouble halved: Age differences in emotional experience and expression during couples' conversations. *Psychology and Aging*, 34(6), 848.

Schatz, erzähl mir mehr von deinen Emotionen!

Soziales Teilen von Emotionen in Partnerschaften und die Rolle von Gesprächszufriedenheit

Helene Merten, Felicitas Schneider-Rausch, Lisa Rabe, Kristin Stöhr

Leitung: Alissa von Großmann

1. Einleitung

Social Sharing of Emotions (SSE) beschreibt einen komplexen Prozess, bei dem ein Individuum selbstrelevante emotionale Informationen mit dem sozialen Umfeld teilt (Rimé et al., 2020). SSE spielt auch in engen Beziehungen wie Partnerschaften eine entscheidende Rolle. Der Fokus dieser Studie liegt deshalb auf der Frage, was Menschen zu sozialem Teilen von Emotionen in Partnerschaften motiviert und ob Gesprächszufriedenheit und vorangegangenes Teilen einen Einfluss darauf haben. Je mehr zwei Personen übereinander wissen, desto stärker wird das wahrgenommene Gefühl von Nähe und Verbundenheit zueinander (Sprecher et al., 2012). Wenn eine Person in einer Paarbeziehung emotionale Informationen teilt, sollte dies also zu mehr Wissen übereinander und dadurch zu mehr Beziehungsnähe führen. Dies wurde auch in einer Studie von Rauers & Riediger (2022) gezeigt. Sie fanden zudem, dass gleichzeitig das Aufrechterhalten oder sogar Vertiefen dieser Nähe zum zukünftigen Ziel der Beziehung wird. Dies könnte die andere Person in der Paarbeziehung dazu motivieren, ihr emotionales Befinden in einem folgenden Gespräch ebenfalls zu teilen, um so die Verbundenheit in der Beziehung zu vertiefen. Wir nehmen daher an, dass das SSE des einen Partners in einem Gespräch positiv mit dem SSE des anderen Partners im folgenden Gespräch zusammenhängt (Hypothese 1). Das *Interpersonal Process Model of Intimacy* von Reis und Shaver (1998) betont, dass zwischenmenschliche Kommunikation nicht allein von den gesendeten Informationen abhängt, sondern maßgeblich von den subjektiven Wahrnehmungen und Interpretationen der Beteiligten beeinflusst wird. Die eigene Wahrnehmung könnte auch das eigene SSE beeinflussen: wenn eine Person im vorherigen Gespräch zufrieden war, dann könnte sie ihr Gegenüber und das zweite Gespräch positiver wahrnehmen und deshalb motivierter sein, Verbundenheit aufzubauen oder aufrechtzuerhalten und mehr zu teilen. Vor diesem Hintergrund wird angenommen, dass die Zufriedenheit der zuhörenden Person mit dem Gespräch positiv mit ihrem SSE im folgenden Gespräch zusammenhängt (Hypothese 2).

2. Methode

Stichprobe. Die Stichprobe umfasste 78 heterosexuelle Paare aus zwei Altersgruppen: 74 jüngere Erwachsene (Alter: $M = 24.34$; $SD = 3.63$) mit einer durchschnittlichen Beziehungsdauer von 4 Jahren und 82 ältere Erwachsene (Alter: $M = 70.27$; $SD = 4.69$) mit einer durchschnittlichen Beziehungsdauer von 38 Jahren. 67,6% der jüngeren Paare lebten zusammen und 92,7% der älteren. *Datenerhebung und Design.* Die Datengrundlage wurde im Rahmen einer Paarstudie von Rohr et al. (2019) erhoben. Eine Person (Person A) begann die Studie mit der Emotionsaufgabe, sie sollte sich an ein aktuelles, negatives Erlebnis erinnern, über das im anschließenden ersten Paargespräch gesprochen werden sollte. Die andere Person (Person B) erhielt eine Alltagsaufgabe, in der sie ihren typischen Tagesablauf rekonstruieren sollte. Danach wechselten die Aufgaben und die Rollen im zweiten Paargespräch. In den Transkripten der Paargespräche wurde auf Basis eines bereits erprobten Kodiermanuals von trainierten Kodierer:innen das SSE pro Gespräch getrennt nach Person A und B kodiert. Pro Hauptsatz (die Analyseeinheit) wurde bewertet, ob SSE vorlag. Jedes Transkript wurde doppelt kodiert und bei unterschiedlicher Codevergabe fanden die beiden Kodierer:innen in Meetings einen Konsens. Die Inter-Rater-Reliabilität war gut ($ICC = .82$).

Maße. Als Maße wurden die absolute Häufigkeit des SSE der erlebnisteilenden Person im ersten Gespräch (Person A) und des SSE der erlebnisteilenden Person im zweiten Gespräch (Person B) verwendet. Die Zufriedenheit mit dem Gespräch wurde durch das Item „Wie zufrieden sind Sie mit dem Gespräch?“ auf einer Skala von 1 (gar nicht zufrieden) bis 5 (sehr zufrieden) erfasst.

Analyse. Der Datensatz enthielt keine fehlenden Werte. Ausreißer wurden zu Beginn der Analyse winsorisiert, wobei Extremwerte (z-Werte größer als 3.29 oder kleiner als -3.29) auf 3.29 bzw. -3.29 gesetzt wurden. Mittels linearen Regressionen wurde SSE der Person B im zweiten Paargespräch durch SSE von Person A im ersten Gespräch (Hypothese 1) und die Zufriedenheit von Person B mit dem ersten Gespräch (Hypothese 2) vorhergesagt. Zusätzlich wurden beide unabhängigen Variablen zusammen in einem multiplen Regressionsmodell als Prädiktoren aufgenommen. Aufgrund der Verletzung der Normalverteilungsannahme der Residuen wurde Bootstrapping angewandt.

3. Ergebnisse

Das Ausmaß des SSE im zweiten Gespräch ließ sich anhand des SSE im ersten Gespräch signifikant vorhersagen, $b = 0.27$, $\beta = 0.23$, $KI [0.03, 0.52]$, $p = .039$. Es konnten 5% der Varianz des SSE von Person B im zweiten Gespräch durch das SSE von Person A im ersten Gespräch aufgeklärt werden, $R^2 = .055$, $F(1,76) = 4.43$, $p = .039$. Die Zufriedenheit von Person B mit dem ersten Gespräch war kein signifikanter Prädiktor für das SSE von Person B im zweiten Gespräch, $b = 0.60$, $\beta = 0.09$, $KI [-0.91, 1.93]$, $p = .441$. Außerdem konnte mit dem Prädiktor Zufriedenheit 0.8% der Varianz im SSE im zweiten Gespräch aufgeklärt werden, $R^2 = .008$, $F(1,76) = 0.60$, $p = .441$. Durch die explorative Aufnahme beider Prädiktoren in ein multiples lineares Regressionsmodell änderten sich die Ergebnisse nicht, $R^2 = .071$, $F(2,75) = 2.86$, $p = .064$. Der Haupteffekt des SSE von Person A im ersten Gespräch blieb dabei signifikant, $b = 0.25$, $\beta = 0.23$, $KI [0.04, 0.55]$, $p = .027$ und der Haupteffekt der Zufriedenheit von Person B war weiterhin nicht signifikant, $b = 0.13$, $\beta = 0.23$, $KI [-0.70, 2.32]$, $p = .264$.

4. Diskussion

Die Ergebnisse zeigen, dass das SSE einer Person in einem Gespräch entsprechend der ersten Hypothese positiv mit dem SSE der anderen Person im folgenden Gespräch zusammenhing. Anders als in Hypothese 2 angenommen, gab es hingegen keinen linearen Zusammenhang zwischen der Zufriedenheit von Person B mit dem ersten Gespräch und ihrem SSE im zweiten Gespräch. Die Ergebnisse unterstützen die Annahme, dass das SSE unseres Gegenübers uns dazu motiviert, selbst mehr zu teilen. Allerdings könnte es auch sein, dass Personen sich tendenziell Partner:innen suchen, die ähnlich viel teilen, wie sie selbst, oder dass Personen in einer Beziehung sich diesbezüglich aneinander anpassen. Um den hier vermuteten Zusammenhang zu bestätigen, könnten zukünftige Studien kontrollierte Experimente einsetzen, bei denen Bedingungen systematisch variiert werden (z.B. Person A wird instruiert viel vs. wenig SSE zu betreiben). Ein weiterer Kritikpunkt ist die Erfassung von Zufriedenheit mit nur einem Item. Eine umfassendere Messung könnte genauere und zuverlässigere Ergebnisse liefern.

5. Literatur

- Rauers, A., & Riediger, M. (2022). Ease of Mind or Ties That Bind? Costs and Benefits of Disclosing Daily Hassles in Partnerships. *Social Psychological and Personality Science*, 14(5), 551-561.
- Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. Duck, D. F. Hay, S. E. Hobfoll, W. Ickes, & B. M. Montgomery (Hrsg.), *Handbook of personal relationships: Theory, research and interventions*, 367-389. *John Wiley & Sons*.
- Rimé, B., Bouchat, P., Paquot, L., & Giglio, L. (2020). Intrapersonal, interpersonal, and social outcomes of the social sharing of emotion. *Current opinion in psychology*, 31, 127-134. <https://doi.org/10.1016/j.copsyc.2019.08.024>
- Sprecher, S., Treger, S., & Wondra, J. (2012). Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interaction. *Journal of Social and Personal Relationships*, 30(4), 497-514.

#staypositive - Der Einfluss von Alter und Valenz berichteter Erfahrung auf das Soziale Teilen von Emotionen bei Frauen

Antonia Burzer, Pauline Götze, Wiebke Klages, Yolenn Mindt, Una Minten

Leitung: Alissa von Großmann

1. Einleitung

Social Sharing of Emotions (SSE) ist ein grundlegender Bestandteil sozialer Interaktion in verschiedenen Beziehungen, Situationen und Settings. Mit SSE ist eine emotionale Selbstoffenbarung, d. h. das Berichten von emotionalen Erlebnissen, gemeint (Rimé, 2009). Welche Rolle dabei die Valenz des Berichteten und das Alter der teilenden Person spielt, soll hier näher betrachtet werden.

Unmittelbar während des SSE findet der Prozess der *Capitalization* statt (Peters et al., 2018). Dieser beschreibt das Phänomen, dass das Teilen positiver Erfahrungen bei der sprechenden Person zu einem gesteigerten positiven Affekt führt. Wir nehmen an, dass dieser Prozess von Personen unbewusst zur Verstärkung des positiven Affekts genutzt wird. Aus bisherigen Studien (Dubois et al., 2016), bei denen Personen innerhalb ihres sozialen Umfelds Meinungen zu Konsumprodukten austauschen, ist bekannt, dass wahrgenommene interpersonelle Nähe zwischen Personen die Kommunikation negativer Informationen erhöht. Weniger Nähe geht hingegen mit dem Teilen positiver Informationen einher. Das kann mit dem Motiv der Selbstaufwertung begründet werden, das vor allem bei wenig Nähe vermehrt aktiviert wird (Dubois et al., 2016). Wir nehmen deshalb an, dass einander fremde Personen mehr SSE betreiben, wenn sie über positive Erfahrungen berichten als über negative Erfahrungen (H1).

Die Sozioemotionale Selektivitätstheorie nach Carstensen (2006) besagt, dass Personen mit einer subjektiv geringen verbleibenden Lebenszeit eher daran orientiert sind, ihr psychisches Wohlbefinden im Moment zu maximieren und weniger daran orientiert sind, Ressourcen zu verwenden, um neue Beziehungen zu knüpfen. Ältere Menschen könnten demnach dazu neigen, sich eher auf Positives zu konzentrieren und bei positiven Themen mehr auf ihre Emotionen einzugehen als bei negativen Themen. Daraus ergibt sich unsere zweite Hypothese, dass der Effekt aus H1 bei älteren Personen im Vergleich zu Jüngeren verstärkt auftritt.

2. Methode

Für die Untersuchung der Hypothesen wurde mit den Daten der Gesprächsstudie von Blanke et al. (2015) gearbeitet.

Die Stichprobe der Studie nach Blanke et al. (2015) besteht aus $N = 208$ Studienteilnehmerinnen. Diese umfasst 102 jüngere (20-31 Jahre) und 106 ältere Frauen (69-80 Jahre). Die Frauen, die sich vor der Studie alle nicht kannten, wurden in 52 altershomogene (jung-jung, alt-alt) oder in 52 altersheterogene (jung-alt) Dyaden aufgeteilt.

Der Großteil der Datenerhebung bestand aus einer strukturierten Unterhaltung, in der sich über positive und negative Erfahrungen ausgetauscht werden sollte. Jede Gesprächsteilnehmerin berichtete der anderen jeweils drei Minuten von einer negativen und einer positiven Erfahrung aus den letzten zwei Wochen.

Die Gespräche wurden transkribiert und anschließend anhand eines bereits erprobten Kodiermanuals hinsichtlich SSE kodiert. Pro Hauptsatz (die Analyseeinheit) wurde die An- oder Abwesenheit von SSE kodiert. Im Prozess der Kodierung wurde jedes Transkript zwei Mal von unterschiedlichen Studierenden kodiert. Bei Unterschieden in der Code-Vergabe wurde in Konsensmeetings eine Entscheidung getroffen. Die Inter-Rater-Reliabilität lag bei $ICC = .83$. Die Hypothesen wurden mit einer two way mixed ANOVA untersucht. Die beiden unabhängigen Variablen sind das Alter

(between-subject-factor) und die Valenz der vorgegebenen Gesprächssituation (within-subject-factor). Die abhängige Variable ist die Häufigkeit des SSE. Es gab keine fehlenden Werte und Ausreißer wurden winsorized. Da die Annahme der Normalverteilung der Residuen für die ANOVA verletzt war, wurde mit Bootstrap Verfahren gearbeitet und so eine robuste ANOVA durchgeführt.

3. Ergebnisse

Der Haupteffekt der Valenz der Gesprächssituation ist nicht signifikant, $F(1,206) = 3.66, p = .057, \eta^2 = .005$. Der Haupteffekt von Alter ist nicht signifikant, $F(1,206) = 0,74, p = .389, \eta^2 = .003$. Ebenso ist der Interaktionseffekt von Valenz und Alter nicht signifikant, $F(1,206) = 1.56, p = .213, \eta^2 = .002$.

4. Diskussion

Beide Hypothesen werden durch die Ergebnisse nicht unterstützt. In der vorliegenden Untersuchung zeigt sich kein Einfluss der Valenz des Berichteten auf die Häufigkeit von SSE. Es gab keinen Moderationseffekt des Alters der teilenden Person auf den Zusammenhang zwischen der Valenz des Berichteten und der Häufigkeit von SSE.

Das Ziehen von Schlüssen (insbesondere auch auf Alltagssituationen) ist durch den Aufbau der Studie limitiert. Folgende Punkte sind dabei besonders zu beachten: Die Valenz des Ereignisses, über das die Probandinnen berichten sollten, war in allen Situationen vorgegeben. Wir nehmen an, dass das Ausmaß an SSE mit der Motivation, die Aufgabe zu erfüllen, konfundiert sein könnte. Die Vorgabe, etwas Negatives zu berichten, könnte dann z. B. dazu geführt haben, dass mehr Negatives berichtet wurde, als es in Alltagssituationen der Fall gewesen wäre.

In der Auswertung und Hypothesenbildung wurde außerdem nicht beachtet, ob die Dyaden hinsichtlich der Alterszusammensetzung heterogen (alt – jung) oder homogen (jung – jung oder alt – alt) waren. In weiteren Studien könnten diese Variablen (Vorgabe der Valenz und Alterszusammensetzung) und ihr Zusammenhang mit SSE näher betrachtet werden.

5. Literatur

- Blanke, E. S., Rauers, A., & Riediger, M. (2015). Nice to meet you—adult age differences in empathic accuracy for strangers. *Psychology And Aging, 30*(1), 149–159. <https://doi.org/10.1037/a0038459>
- Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science, 312*(5782), 1913–1915. <https://doi.org/10.1126/science.1127488>
- Dubois, D., Bonezzi, A., & De Angelis, M. (2016). Sharing with Friends versus Strangers: How Interpersonal Closeness Influences Word-of-Mouth Valence. *Journal Of Marketing Research, 53*(5), 712–727. <https://doi.org/10.1509/jmr.13.0312>
- Peters, B. J., Reis, H. T., & Gable, S. L. (2018). Making the good even better: A review and theoretical model of interpersonal capitalization. *Social and Personality Psychology Compass, 12*(7). <https://doi.org/10.1111/spc3.12407>
- Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and Empirical review. *Emotion Review, 1*(1), 60–85. <https://doi.org/10.1177/1754073908097189>

Sonnenbrille oder Bart: Nettes Accessoire oder erfolgreiche Tarnung?

**Monique Becker, Katharina Bösendorfer, Lena Breidt, Emma Fieler,
Friederike Gehre, Lilly Hesse, Lena-Sophie Raspe, Elisa Ruehlmann,
Alexandra Zimmer, Tabea Neumann**

Leitung: Jürgen M. Kaufmann

1. Einleitung

Bei der Verbrechensaufklärung sind Polizei und Justiz von der Verlässlichkeit von Augenzeugenaussagen abhängig, um Justizirrtümer zu vermeiden. Dabei existieren individuelle Unterschiede in der Fähigkeit unbekannte Gesichter zu lernen und wiederzuerkennen (Bruce et al., 1999). Der negative Einfluss auf die Leistung der Gesichtserkennung durch die Veränderung externer Merkmale, wie zum Beispiel die Frisur, wurde bereits mehrfach nachgewiesen (Patterson, & Baddeley, 1977). Jedoch konnte noch nicht endgültig geklärt werden, welche Gesichtsregionen für die Identifikation entscheidend sind. Einige Studien verweisen auf die hohe Relevanz der Augenpartie für das Wiedererkennen gelernter Gesichter (Itier et al., 2007). Zwar liefern erste Studien durch Face-Matching und Cambridge Face Memory Tests (CFMT, Duchaine & Nakayama, 2006) erste Hinweise darauf, dass sich die Leistung signifikant verschlechterte (Kramer & Ritchie, 2016), wenn eines von zwei gezeigten Gesichtern eine Sonnenbrille trug, jedoch stellt sich die Frage, inwieweit sich diese Ergebnisse für den CFMT, indem es um das Erkennen von gelernten Gesichtern geht, replizieren lassen. Daraus resultiert das Forschungsinteresse, inwieweit eine Modifikation, z.B. das Tragen einer Sonnenbrille oder eines Bartes in der Lernphase die Wiedererkennungslleistung beeinflusst, wenn in der Testsituation diese Merkmale wegfallen.

2. Methode

Insgesamt wurden Daten von 62 Versuchspersonen (VPN, $M = 27$ Jahre, $SD = 10$), davon 41 Frauen, 20 Männer und eine Person ohne Angabe, mittels der PsyToolkit Version 3.4.4 (Stoet, 2010; 2017) online anonymisiert erhoben.

Das Experiment beruhte auf einem $2 \times 2 \times 2$ Mixed Design, welches sich aus einem between-Faktor bezüglich der Maskierungsart (Sonnenbrille oder Bart), dem within-Faktor Testblock (Block mit maskiert gelernten vs. unmaskiert gelernten Gesichtern) und dem within-Faktor Bekanntheit (gelernt vs. neu) zusammensetzte. Zu Beginn wurden die demographischen Daten der VPN erfragt. Anschließend wurden sie randomisiert einer der beiden Maskierungsarten zugeordnet. In der Lernphase wurden insgesamt 28 männliche Gesichter gelernt, davon 14 maskierte und 14 neutrale. Nach einer zweiminütigen Pause wurden in der Testphase alle gelernten Gesichter ohne Maskierung sowie 28 neue nicht-maskierte männliche Gesichter gezeigt. Die Präsentation erfolgte blockweise: ein Block („maskiert gelernt“) bestehend aus 14 zuvor maskiert gelernten Gesichtern, sowie 14 unbekanntem Gesichtern, und ein anderer Block („unmaskiert gelernt“) aus 14 zuvor unmaskiert gelernten Gesichtern und 14 unbekanntem Gesichtern. Per Tastendruck sollten die Gesichter als bekannt oder unbekannt eingeordnet werden. Zur Erfassung der individuellen Lernfähigkeit für Gesichter absolvierten die VPN anschließend den CFMT.

3. Ergebnisse

Als abhängige Variablen wurden für das Experiment die Signalentdeckungsparameter d' (Sensitivität) und das Kriterium C (Antwortbias) berechnet. Für den CFMT ermittelten wir den Gesamtscore. Die Datenanalyse erfolgte mittels R-Studio (2024). Grundlage für die Datenanalyse waren die Berechnungen von 2×2 mixed ANOVAs mit dem Gruppenfaktor Maskierungsart (Bart oder Brille) und dem Messwiederholungsfaktor Testblock. Da in die Berechnung von d' und C

sowohl gelernte, als auch neue Gesichter eingingen, entfällt in der ANOVA der Faktor Bekanntheit. Für d' zeigte sich kein Unterschied für die Art der Maskierung (Bart oder Brille), $F(1, 60) < 1$, $p = 0.785$, $\eta_p^2 < 0.011$. Der Haupteffekt für den Faktor Testblock legte bessere Leistungen für Gesichter nahe, die zuvor neutral gelernt wurden, als für Gesichter, die maskiert gelernt wurden, $F(1, 60) = 43.092$, $p < 0.001$, $\eta_p^2 = 0.418$. Die Faktoren Maskierungsart und Testblock interagierten nicht, $F(1, 60) = 2.204$, $p = 0.143$, $\eta_p^2 = .035$. Für den Antwortbias (Krit. C) zeigte sich ebenfalls lediglich ein Haupteffekt für den Faktor Testblock, $F(1, 60) = 12.83$, $p < 0.001$, $\eta_p^2 = 0.176$, mit einer konservativeren Antworttendenz für die Testblockbedingung „maskiert gelernt“. Die Korrelation zwischen der Akkuratheit im CFMT und dem Sensitivitätsmaß d' in der Testblockbedingung „unmaskiert gelernt“ war signifikant, $r(60) = 0.47$, $t = 4.147$, $p < 0.001$. Das heißt, je höher die Akkuratheit der VPN im CFMT war, desto besser war die Leistung in diesem Block. Im Gegensatz dazu war die Korrelation zwischen der Akkuratheit im CFMT und d' im Testblock „maskiert gelernt“ nicht signifikant, $r(60) = 0.20$, $t = 1.56$, $p = 0.124$.

4. Diskussion

Die vorliegende Studie untersuchte die Effekte verschiedener Maskierungsarten auf die Identifikationsleistung. Wie erwartet zeigte sich, dass die Maskierung von Gesichtern die Identifikationsleistung signifikant verschlechtert, wenn die Maskierung in der Testphase wegfällt. Jedoch hatte die Art der Maskierung keinen signifikanten Einfluss, d.h. eine Maskierung der oberen oder der unteren Gesichtsregionen die Identifikationsleistung ähnlich beeinträchtigten. Darüber hinaus zeigte eine Analyse der Antworttendenzen einen konservativeren Bias für zuvor manipuliert gelernte Gesichter. Das könnte darauf hinweisen, dass bei Unsicherheit eher Gesichter abgelehnt werden, was eine Strategie zur Verringerung von Fehlidentifikationen sein könnte, in der Praxis aber zu häufigeren Verpassern führen dürfte, also der Nichtidentifikation von Schuldigen.

Die individuelle Lernfähigkeit wurde mit dem CFMT erfasst. VPN mit höheren Werten waren zwar besser in der Identifikation von unmaskiert erlernten Gesichtern, aber nicht in der Identifikation von maskiert erlernten Gesichtern. Das könnte bedeuten, dass die Fähigkeit, vormals maskierte Gesichter zu erkennen, möglicherweise von anderen Faktoren beeinflusst wird, die nicht durch den CFMT erfasst werden, und Differenzen in der individuellen Lern- und Wiedererkennungslleistung bei maskierten Gesichtern geringer sein könnten als bei unmaskierten Gesichtern. Bei einer Verwendung des CFMT in der Praxis zur Einschätzung der Zuverlässigkeit von Augenzeugenaussagen sollten diese Unterschiede berücksichtigt werden.

5. Literatur

- Bruce, V., Henderson, Z., Greenwood, K., Hancock, P., Burton, A. & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology. Applied*, 5(4), 339–360. <https://doi.org/10.1037/1076-898x.5.4.339>
- 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. <https://doi.org/10.1016/j.neuropsychologia.2005.07.001>
- Itier, R. J., Alain, C., Sedore, K., & McIntosh, A. R. (2007). Early face processing specificity: it's in the eyes! *Journal of Cognitive Neuroscience*, 19(11), 1815–1826. <https://doi.org/10.1162/jocn.2007.19.11.1815>
- Kramer, R. S. S., and Ritchie, K. L. (2016) Disguising Superman: How Glasses Affect Unfamiliar Face Matching. *Applied. Cognitive Psychology*, 30, 841–845. doi: 10.1002/acp.3261.
- Patterson, K. E., & Baddeley, A. D. (1977). When face recognition fails. *Journal of Experimental Psychology. Human Learning and Memory*, 3(4), 406–417. <https://doi.org/10.1037/0278-7393.3.4.406>
- RStudio Team (2024). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL, Version 2024.04.2+764 [Software]. <http://www.rstudio.com/>.
- 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), 2431. <https://doi.org/10.1177/0098628316677643>

Wer war's? Forschungsansätze zur Verbesserung von Zeugenaussagen mittels der Stimme und virtueller Realität

Xenia Elenschleger, Johannes Fester, Camille Louise Fischer, Rebecca Fischer, Helin Hersek, Lena Konradi, Elina Lippmann, Celine Lindig, Filiz Scholz, Denis Pantschenko, Maximilia Schinner

Leitung: Jürgen M. Kaufmann

1. Einleitung

Augenzeugenaussagen sind in juristischen Verfahren oft unzuverlässig– zugleich jedoch unverzichtbar. Valentine et al. (2003) beschreiben eine korrekte Identifizierung von Tätern in nur 40% der Fälle. Hier ansetzende Forschung kann zu mehr Gerechtigkeit in unserer Gesellschaft beitragen. Das Ziel unseres Forschungsvorhabens ist es, die Fehlerquote in der Identifikation von Tätern zu minimieren. Unser Ansatz ist u. a. ein Vergleich von multimodaler Präsentation und Virtueller Realität (VR) gegenüber klassischen Präsentationsverfahren. Untersucht wurde jeweils die Wiedererkennungslleistung und die selbsteingeschätzte Sicherheit.

2. Studie des Wintersemesters

2.1 Methode

Es wurden 135 Datensätze (Alter: $M = 23.96$ Jahre) statistisch ausgewertet (R Core Team, 2022). Es lag ein 2×2 faktorielles Design mit den Stufen Stimme präsentiert und Stimme nicht präsentiert des between-Faktors Stimmbedingung sowie target-present und target-absent des within-Faktors Line-up-Typ vor. In der Lernphase sahen die Probanden zehn Videofrequenzen (mit Ton) verschiedener Personen, die den gleichen Satz sprachen. Nach einer Zwischenaufgabe sollten die Gesichter in einem Line-up wiedererkannt werden. Probanden der Experimentalgruppe hörten zusätzlich die Stimme der Personen. Es musste jedes Gesicht als „Täter“ oder als „unschuldig“ klassifiziert sowie die Entscheidungssicherheit angegeben werden. Probanden durchliefen dieselbe Anzahl von target-present und target-absent Line-ups.

2. 2 Ergebnisse und Diskussion

Eine 2×2 Mixed ANOVA mit Stimmbedingung und Line-up-Typ ergaben keinen Haupteffekt für den Faktor Stimmbedingung, $F(1, 133) = 0.33$, $p = .566$, $\eta^2 < .01$. Der Haupteffekt Line-up-Typ war signifikant, $F(1, 133) = 72.96$, $p < .001$, $\eta^2 = .35$: target-present Line-ups wurden damit häufiger korrekt beantwortet als in target-absent Line-ups, für die es häufig zu falschen Alarmen kam. Es lag keine signifikante Interaktion zwischen beiden Faktoren vor, $F(1, 133) = 0.32$, $p = .570$, $\eta^2 < .01$. Eine ANOVA für den Einfluss der beiden Faktoren auf die Sicherheit der Probanden ergab signifikante Haupteffekte für die Faktoren Stimmbedingung, $F(1, 133) = 4.35$, $p = .039$, $\eta^2 = .03$. und Line-up-Typ, $F(1, 133) = 39.46$, $p < .001$, $\eta^2 = .23$. Die Interaktion zwischen den Faktoren war nicht signifikant, $F(1, 133) = 2.80$, $p = .097$, $\eta^2 = .02$. Die Hinzunahme der Stimme im Line-Up konnte so zwar die gefühlte Sicherheit der Probanden signifikant verbessern, nicht aber deren Wiedererkennungslleistung. Wenn auch nicht signifikant, ergaben sich in der Stimme präsentiert-Bedingung allerdings numerisch etwas weniger falsche Alarme als bei Stimme nicht präsentiert. Falsche Gesichter wurden also etwas seltener als Target identifiziert.

3. Studie des Sommersemesters

3.1 Methode

Es wurden 70 Datensätze (Alter: $M = 21.8$ Jahre) statistisch ausgewertet. Es handelte sich um ein zweifaktorielles between-subjects Design mit dem Faktor Lernformat (2D-Video vs. VR) und dem Faktor Testformat (2D-Foto vs. animierte 3D Stimuli der Gesichter). Die Probanden sahen im Labor das Video einer gestellten Straftat, die entweder im 2D-Format oder mithilfe einer VR-Brille im VR-Format gezeigt wurde. Nach einer Woche durchliefen alle Probanden ein online PsyToolkit Experiment (Stoet, 2010; 2017). In zwei Line-ups (je einmal target-present und -absent) musste jeweils jedes der enthaltenen acht Gesichter entweder als Täter oder als unschuldig klassifiziert werden. Zusätzlich sollte die Entscheidungssicherheit auf einer Skala von 1 bis 6 angegeben werden.

3.2 Ergebnisse und Diskussion

VR-Videos in der Lernphase verbesserten die Identifikationsleistung im Vergleich zu 2D-Videos nicht, weder bei target-present ($\chi^2 = 0.552$, $df = 2$, $p = .759$) noch bei target-absent Line-ups ($\chi^2 = 0.040$, $df = 1$, $p = .841$). In der Testphase zeigte sich für das target-present Line-up kein Unterschied zwischen den Leistungen für 3D-Animationen im Vergleich zu 2D-Fotos ($\chi^2 = 5.419$, $df = 2$, $p = .067$). Für das target-absent Line-up ergab sich hierfür jedoch ein signifikanter Unterschied ($\chi^2 = 8.478$, $df = 1$, $p = .003$), der sich nach Aufschlüsselung in die vier Experimentalbedingungen (2D/2D, VR/2D, 2D/3D, VR/3D) in der 2D-Video x 2D-Foto-Bedingung als signifikant weniger falsche Alarmer widerspiegelte ($\chi^2 = 12.735$, $df = 3$, $p = .005$). Die Kombination von VR-Videos und 3D-Animation führte, entgegen unseren Erwartungen, zu keiner besseren Identifikationsleistung im Vergleich zu den anderen drei Bedingungen. Hinsichtlich der selbsteingeschätzten Sicherheit zeigte eine 2 x 2 between-subjects ANOVA keinen signifikanten Unterschied zwischen der VR- und 2D-Video Gruppe (2D: $M = 3.08$, $SD = 1.23$; VR: $M = 3.29$, $SD = 1.46$; $F(1, 66) = 0.61$, $p = .437$, $\eta^2 = .01$). Für das Testformat zeigte sich ein Vorteil für 2D-Fotos im Vergleich zu 3D-Animationen (2D: $M = 3.51$, $SD = 1.25$; 3D: $M = 2.86$, $SD = 1.37$; $F(1, 66) = 5.87$, $p = .018$, $\eta^2 = .08$). Insgesamt zeigten sich also sowohl für die Lern-, als auch für die Testbedingung eher Vorteile für die 2D-Stimuli. Da die 360°-Sicht der VR-Bedingung die Probanden mit mehr Informationen konfrontierte, könnte dies eine Ablenkung vom Gesicht verursacht haben. Die schlechteren Ergebnisse der 3D-Testbedingung könnten u.a. auf die geringere Auflösungsqualität der Textur zurückzuführen sein. Zudem sollten Probanden das Objekt selbst drehen können, um den 3D-Effekt klarer hervorzuheben. VR-Augenzeugenstudien sind nach wie vor mit Herausforderungen konfrontiert. Gezielte Nachbesserungen im Forschungsdesign könnten aber zu einer möglichst realitätsnahen Gegenüberstellung und so zu ökologisch valideren Ergebnissen betragen.

4. Zusammenfassung

Unsere Studien untersuchten potentielle positive Effekte der zusätzlichen Präsentation von auditiven Reizen in Form der menschlichen Stimme (WiSe) und des Einsatzes von 3D-Animationen und virtueller Realität (SoSe) auf die Identifikationsleistung von unbekanntem Gesichtern. In beiden Fällen zeigten sich die erwarteten Vorteile nicht. In der zweiten Studie zeichnete sich sogar ein Trend eines negativen Effekts des Nutzens von 3D-Animationen als Teststimuli ab. Das prinzipielle Problem der Unzuverlässigkeit von Augenzeugenaussagen für Line-ups bleibt bestehen.

5. Literatur

- Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096-1104.
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 2431.
- R Core Team. (2022). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology*, 17(8), 969-993. <https://doi.org/10.1002/acp.939>

Zwischen Bildschirmen und Begegnungen – die Rolle von parasozialen und sozialen Interaktionen bei der Wahrnehmung von Persönlichkeiten

Lara Hornischer, Linda Lehmann, Emma Lippold, Sophie Mintert, Laura Press, Annika Schmidt

Leitung: Gyula Zoltán Kovacs

1. Einleitung

Der Begriff der Parasozialen Interaktion, erstmals durch Horton & Wohl (1956) definiert, beschreibt eine einseitige Form der sozialen Interaktion zwischen Medienrezipient*innen und -figuren. Anders als soziale Interaktionen, charakterisiert durch eine bidirektionale Kommunikation – typischerweise face-to-face – basieren Parasoziale Interaktionen nicht auf Gegenseitigkeit (Schramm & Liebers, 2019). Parasoziale Interaktionen und Beziehungen sind heutzutage allgegenwärtig: Fast täglich konsumieren wir Inhalte von verschiedensten Personen online, sei es von unseren Freund*innen, oder von uns im echten Leben unbekanntem Personen. Die Interaktion mit Fremden oder Freund*innen geschieht auf Social Media auf ähnliche Weise, wodurch parasoziale und soziale Interaktion hier teilweise ineinander übergehen, was das Ziehen einer klaren Grenze zwischen den beiden erschwert (Baek et al., 2013). Wie bei Interaktionen im echten Leben ähnlich, bilden wir uns auch während Medienkonsum und parasozialer Interaktion Meinungen über das Gesehene und über die Medienpersönlichkeiten. Aber wie verändert sich die Wahrnehmung dieser (unbekanntem) Persönlichkeiten, wenn die parasoziale Interaktion zusätzlich durch eine soziale, bidirektionale Interaktion ergänzt wird? Diese Frage ist zentrales Thema unserer Forschung. Im Folgenden soll daher der Effekt parasozialer und sozialer Interaktion auf die Wahrnehmung von Persönlichkeiten, sowie auf den Grad der Parasozialen Interaktion zwischen Social Media Konsument*innen und Social Media Persönlichkeiten untersucht werden.

2. Methode

Die Stichprobe umfasste insgesamt 20 Personen (N=20, davon 18 weiblich, 2 männlich gelesen), die zufällig auf zwei Bedingungen verteilt wurden: Gruppe A (N=10, 9 weiblich, 1 männlich gelesen) hatte neben parasozialen Interaktionen auch real-life Interaktionen mit den Personae, während Gruppe B (N=10, 9 weiblich, 1 männlich gelesen) nur parasoziale Informationen erhielt. Alle Teilnehmenden mussten die für diese Studie kreierte Profile unserer drei Personae auf Instagram abonnieren. Als Stimulusmaterial verwendeten wir Instagram Posts in Form von Bildern mit kontextueller Valenz in Textform (Bildunterschriften), Videos in 10-Minuten-Länge als Instagram Stories (für 24 Stunden sichtbar) und persönliche Treffen. Die Trainingsphase dauerte 14 Tage, an denen täglich (Wochenenden ausgenommen) neun Bilder (drei pro Persona) gepostet wurden, welche die Personae möglichst umfangreich repräsentierten. An drei Tagen pro Woche erhielten Proband*innen der Gruppe A neben dem parasozialen Kontakt zusätzlich sozialen Kontakt in Form von 10-minütigen persönlichen Treffen, bei denen sie mit den Personae anhand von Gesellschaftsspielen oder Dialogen interagierten. Gruppe B bekam stattdessen an denselben drei Tagen zehnminütige Videos in Instagram „Close Friends“ Stories zu sehen, auf welche Gruppe A keinen Zugriff hatte. Die Inhalte der sechs Videos wurden so ähnlich wie möglich an den offline Treffen orientiert. Nach der Trainingsphase fand die Erhebungsphase statt. Dafür mussten die Versuchspersonen einen Fragebogen in PsychoPy ausfüllen, der sich aus einer modifizierten Version der Parasocial Interaction Process Scale (PSI-Process Scale, Schramm & Hartmann, 2008) und aus Fragen zu „perceived social traits“ (PST-Fragebogen) zusammensetzte. Die original PSI-Prozess-Skala wurde von 112 Items auf 26 Items gekürzt, welche in drei Teilprozesse der Parasozialen Interaktion (kognitiv, affektiv und konativ) unterteilt sind. Die 26 Items mussten neben einem globalen Valenzmaß für jede der drei Personae beantwortet werden. Das gleiche galt für die 10 Items

des PST-Fragebogens. Die Messung des PSI erfolgte über eine fünfstufige Likert-Skala („trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“). Während die PSI-Skala die Intensität der parasozialen Interaktionen mit den einzelnen Personae erfragte, erfasste der PST-Fragebogen die Wahrnehmung über die sozialen Persönlichkeitseigenschaften und Verhalten der drei Personae. Diese wurden anhand einer 7-stufigen Likert-Skala erhoben.

3. Ergebnisse

Parasoziale Interaktionen: Die Ergebnisse der mehrfaktoriellen ANOVA liefern einen signifikanten Unterschied auf Gruppenebene hinsichtlich der Affektivität ($F_{1,18} = 10.685$; $p = 0.004$) und der Kognition ($F_{1,18} = 4.335$; $p = 0.05$). Für die Valenz sowie die Konation ergab die ANOVA keinen signifikanten Unterschied (*Valenz:* $F_{1,18} = 3.200$; $p = 0.090$; *Konation:* $F_{1,18} = 2.494$; $p = 0.132$).

Perceived Social Traits: Die Analyse der Daten zeigt einen signifikanten Unterschied in der Eigenschaft Kompetenz auf Gruppenebene ($F_{1,18} = 5.198$; $p = 0.035$). In den Eigenschaften Wärme, Attraktivität, Charisma, Femininität und Jugendlichkeit zeigen sich signifikante Effekte auf Personaebene ($F_{2,36} > 5.490$; $p < 0.008$), jedoch nicht auf Gruppenebene (*Wärme:* $F_{1,18} = 0.895$; $p = 0.357$; *Attraktivität:* $F_{1,18} = 1.074$; $p = 0.314$; *Charisma:* $F_{1,18} = 1.742$; $p = 0.203$; *Femininität:* $F_{1,18} = 0.072$; $p = 0.791$; *Jugendlichkeit:* $F_{1,18} = 0.500$; $p = 0.488$). Für Verträglichkeit ergab sich ebenfalls kein signifikanter Effekt auf Gruppenebene, der aber potenziell als Tendenz zu interpretieren ist ($F_{1,18} = 2.425$; $p = 0.137$).

4. Diskussion

Einige Effekte der PSI-Prozess-Skala zeigen nur knapp keine Signifikanz, was die Vermutung, bei einer höheren Teilnehmendenzahl hier mehr signifikante Effekte zu erzielen, nahelegt. Unsere PST-Ergebnisse zeigen nur wenige signifikante Effekte auf Gruppenebene, jedoch lässt sich aufgrund der Tendenzen suggerieren, dass soziale Interaktionen höhere Bewertungen der Personae als parasoziale Interaktionen hervorrufen, allerdings nur in geringem Maße. Die Ergebnisse des PST-Tests deuten darauf hin, dass ein Großteil der Persönlichkeitseigenschaften ebenso durch Instagram (parasoziale Bedingung) interpretiert werden kann wie durch soziale Interaktionen. Der signifikante Effekt in der wahrgenommenen Kompetenz der Personae zeigt, dass Proband*innen, die direkten sozialen Kontakt mit den Personae hatten, diese deutlich kompetenter einschätzten, als jene, die keine bidirektionale Interaktion mit ihnen erlebten. Forschungen zu möglichen Gründen einer höher wahrgenommenen Kompetenz bei sozialen Interaktionen im Vergleich zu parasozialen Interaktionen stehen noch offen. Interessant wäre es auch Konsequenzen für beispielsweise die Lehre zu betrachten. Welche Folgen hat es für die Wahrnehmung über Lehrende an Universitäten, wenn Vorlesungsinhalte parasozial, also nur anhand von Aufzeichnungen vermittelt werden? Dies scheint eine sinnvolle Forschungsfrage für zukünftige Studien zu parasozialen Interaktionen zu sein.

5. Literatur

- Baek, Y. M., Bae, Y. & Jang, H. (2013). Social and Parasocial Relationships on Social Network Sites and Their Differential Relationships with Users' Psychological Well-Being. *Cyberpsychology, Behavior And Social Networking*, 16(7), 512–517. <https://doi.org/10.1089/cyber.2012.0510>
- Horton, D. & Wohl, R. R. (1956). Mass Communication and Para-Social Interaction. *Psychiatry*, 19(3), 215–229. <https://doi.org/10.1080/00332747.1956.11023049>
- Schramm, H. & Hartmann, T. (2008). Die Messung von parasozialen Interaktionen als mehrdimensionales Konstrukt: Entwicklung und Validierung von PSI-Prozess-Skalen auf Basis des Zwei-Ebenen-Modells parasozialer Interaktionen. In *Die Brücke zwischen Theorie und Empirie: Operationalisierung, Messung und Validierung in der Kommunikationswissenschaft* (Nummer 3, S. 48–69). <https://research.vu.nl/en/publications/die-messung-von-parasozialen-interaktionen-als-mehrdimensionales->
- Schramm, H. & Liebers, N. (2019). Parasocial Interactions and Relationships with Media Characters – An Inventory of 60 Years of Research. *Communication Research Trends*, 38(2), 4–31. <https://dialnet.unirioja.es/servlet/articulo?codigo=7895318>

Binge watching für kulturelle Sensibilität: Beeinflusst das Anschauen koreanischer Serien den Other-Race Effekt?

Pauline Fahle, Elham Kazemi, Samuel Melle, Katharina Riewe, Emily Schulz

Leitung: Gyula Zoltán Kovacs

1. Einleitung

Die Beliebtheit koreanischer Serien, insbesondere Netflix-Originale wie "Squid Game" mit über 260 Mio. Views in den ersten 91 Tagen, verzeichnet zurzeit einen regelrechten Boom (Netflix, 2024). Dieser Trend bietet eine interessante Perspektive auf den other-race-effect (ORE). Dabei handelt es sich um ein psychologisches Phänomen, das eine bessere Gesichtserkennung innerhalb der eigenen Ethnie beschreibt (Chance & Goldstein, 1996). Es gibt Hinweise darauf, dass die Exposition zu Gesichtern anderer Ethnien den ORE verringern könnte (Chiroro & Valentine, 1995). Wir wollen daher untersuchen, ob die regelmäßige Exposition zu koreanischen Medien die Gesichtserkennung für diese Gesichter verbessert und somit den ORE reduziert.

H_1 : Wir erwarten einen Unterschied in der *Performance d'* (Wiedererkennungseistung).

H_2 : Wir erwarten einen Unterschied in der *Reaktionszeit* zwischen den Gruppen

2. Methode

Stichprobe. Unsere initiale Stichprobe bestand aus 167 Personen. Mittels einer Umfrage auf www.sosicisurvey.de (Leiner, 2024) wurde demographischen Daten und eine Selbsteinschätzung der Gesichtserkennungsleistung erfasst. Auch sollten die Befragten angeben, ob sie regelmäßig koreanische Serien schauten. Traf dies zu, wurden sie der *Experimentalgruppe* (EG) zugeordnet, andernfalls gehörten sie zur *Kontrollgruppe* (KG). Insgesamt wurden neun Versuchspersonen (VPN) pro Gruppe für die Untersuchung ausgewählt.

Paradigma. Die VPN sahen sich für zwei Wochen täglich eine Folge einer bestimmten Serie an und verfassten einen kurzen Tagebucheintrag. Letzteres war Bestandteil unserer Cover Story, laut der wir den Unterschied in parasozialer Identifikation zwischen fiktiven koreanischen und europäischen Charakteren erforschten. Die VPN erhielten täglich Erinnerungsmails, um die gewissenhafte Durchführung zu sichern. Bei der Serie handelte es sich um eine koreanische (EG) - bzw. Hollywood-Serie (KG). Abschließend folgte ein 15-minütiges Experiment im Jentower.

Stimuli. Wir nutzten eine gekürzte Version des "old-new-paradigms" von Zhou et al. (2021). In der Trainingsphase wurden den VPN 12 asiatische und 12 europäische Gesichter präsentiert. Anschließend wurden in der Testphase 48 Gesichter gezeigt; 24 hatten die VPN bereits gesehen, 24 waren neu. Um die "within-person variability" (Jenkins et al., 2011) und den "picture-recognition effect" (Zhuo et al., 2021) zu berücksichtigen, wurden unterschiedliche Perspektiven derselben Gesichter verwendet. Insgesamt wurden 48 KI-generierte Stimuli mit ausgewogenem Geschlechterverhältnis genutzt. (Generated Photos, 2024).

Ablauf des Experiments. Das Experiment wurde in PsychoPy2 erstellt und durchgeführt (Peirce et al., 2019). In der Trainingsphase sahen die VPN in zwei Durchgängen die Stimuli für wenige Sekunden; in der Testphase sollten sie per Tastendruck angeben, ob das gezeigte Gesicht ihnen bekannt war.

3. Ergebnisse

Die Auswertung erfolgte in der Statistik-Software R (v4.4.0; R Core Team 2024). Keine der o.g. demographischen Daten zeigten eine signifikante Varianzaufklärung; daher wurden sie nicht im Modell berücksichtigt. Die *Performance d'* gab die Trefferrate der VPN unter Berücksichtigung der Ähnlichkeit der Stimuli und des Bias der einzelnen VPN an (Stanislaw & Todorov, 1999). Die *Reaktionszeit* in Sekunden wurde als zweites Maß für die Performance einbezogen. Für jede

Hypothese wurde eine 2x2 messwiederholte ANOVA mit Ethnie der Stimuli (*within*-Faktor) und Gruppenzugehörigkeit (*between*-Faktor) durchgeführt. Ad-hoc wurde ein Outlier entfernt.

H_1 : d' zeigte keinen signifikanten Gruppenunterschied, abhängig von der Art des gezeigten Stimulus; die Interaktion wurde nicht signifikant $F(1,15) = 0.70, p = 0.415; \eta_p^2 = 0.04, 90\%-KI [0.00, 0.29]$.

H_2 : Die Reaktionszeit zeigte einen signifikanten Gruppenunterschied, abhängig von der Art des Stimulus; die Interaktion wurde signifikant mit $F(1,15) = 11.53, p = 0.004; \eta_p^2 = 0.43 90\%-KI [0.12, 0.65]$.

Post-Hoc-Analyse. Durch Entfernung des Outliers hatten wir ungleiche n in EG/KG, daher mussten wir die Daten in ein Mehrebenenmodell fitten, um gepaarte t-Tests durchführen zu können. Anschließend wurden die geschätzten marginalen Mittel berechnet, anhand derer die Bedingungen verglichen wurden. Zur Adjustierung der p-Werte nutzten wir die Holm-Methode.

Post-Hoc-Ergebnisse. Die Vergleiche der Bedingungen waren nicht signifikant, aber: (1) Der Vergleich der Stimuli in der KG wurde fast signifikant ($t(15) = -2.893, p = 0.067$). Der t-Wert war negativ, da die Reaktionszeit für asiatische Gesichter größer war als für europäische. (2) Der Vergleich der Stimuli in der EG unterschied sich im p-Wert stark von den restlichen Bedingungen ($t(15) = 1.881, p = 0.398$). Der t-Wert war erwartungsgemäß positiv, da sich die Reaktionszeit für asiatische Gesichter verringern sollte.

4. Diskussion

Unsere Thesen konnten nicht bestätigt werden, es zeigten sich jedoch Effekte in die von uns postulierte Richtung. Für die Reaktionszeit ergab sich eine signifikante Interaktion; die anschließenden Post-Hoc Vergleiche wurden fast signifikant. Die Studie weist drei Schwachstellen auf: (1) die Stichprobengröße war zu gering, somit hatten wir nicht genug Power, um einen möglichen Effekt aufdecken zu können. (2) Die Stimuli waren sich untereinander zu ähnlich. Da wir aufgrund der Voraussetzungen auf KI-generierte Stimuli zurückgreifen mussten, war dies ein Trade-off, den wir eingehen mussten. Hierdurch haben wir möglicherweise in der Performance, besonders für d' , eine genauere Auflösung eingebüßt. (3) Eine Kontrolle des Ausmaßes des (koreanischen) Medienkonsums wäre zielfördernd gewesen. Wir schlagen daher vor, die Studie unter verbesserten Bedingungen erneut durchzuführen. Mehr VPN und Stimuli mit realen Personen könnten weitere Erkenntnisse liefern, ob der ORE mittels Medienkonsum reduziert werden kann.

5. Literatur

- Chance, J. E., & Goldstein, A. G. (1996). The other-race effect and eyewitness identification. In *Psychological issues in eyewitness identification*. (pp. 153-176). Lawrence Erlbaum Associates, Inc.
- Chiroro, P., & Valentine, T. (1995). An Investigation of the Contact Hypothesis of the Own-race Bias in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, 48(4), 879-894. <https://doi.org/10.1080/14640749508401421>
- Generated Photos (2024). *Unique, worry-free model photos*. <https://generated.photos/>
- Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition*, 121(3), 313-323. <https://doi.org/https://doi.org/10.1016/j.cognition.2011.08.001>
- Leiner, D. J. (2024). SoSci Survey (Version 3.5.02) [Computer software]. Available at <https://www.sosicurvey.de>
- Netflix. (2024). *Top 10 TV Shows On Netflix Right Now*. <https://www.netflix.com/tudum/top10/tv>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195-203. <https://doi.org/10.3758/s13428-018-01193-y>
- R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection measures. *Behavior research methods, instruments, & computers: a journal of the Psychonomic Society, Inc*, 31, 137-149.
- Zhou, X., Burton, A. M., & Jenkins, R. (2021). Two Factors in Face Recognition: Whether You Know the Person's Face and Whether You Share the Person's Race. *Perception*, 50(6), 524-539. <https://doi.org/10.1177/03010066211014016>

Listen to my voice – do you trust me?

Demiray, Canel; Huang, Ying; Kemmler, Lea; Märkl, Marlene; Nedziella, Angelina; Sahin, Beyza

Leitung: Romi Zäske

1. Einleitung

Die Fähigkeit, die Vertrauenswürdigkeit von Personen zu beurteilen, spielt eine entscheidende Rolle in verschiedenen sozialen und beruflichen Kontexten und kann menschliches Erleben und Verhalten beeinflussen. Dabei wird die Wahrnehmung von Vertrauenswürdigkeit fremder Personen u.a. durch den Klang der Stimme beeinflusst [1]. Weitgehend unerforscht ist, ob die Vertrauenswürdigkeit einer Stimme mit Vertrautheit, d.h. der Bekanntheit einer Stimme, variiert. Eine erste Studie lieferte hierzu gemischte Befunde [2]: So fand man weder nach einem intensiven Stimmlertraining für unbekannte Stimmen noch für bereits bekannte berühmte Stimmen einen eindeutigen Effekt von Bekanntheit auf wahrgenommene Vertrauenswürdigkeit. Im Vgl. zu Lavan et al. [2] soll die vorliegende Studie eine größere und repräsentativere Auswahl von Stimmen lernen lassen und durch einmalige Präsentation von unbekanntem Stimmen sicherstellen, dass diese während des gesamten Experiments gänzlich unbekannt bleiben.

Um den Effekt von Vertrautheit von Stimmen und auf deren wahrgenommene Vertrauenswürdigkeit zu erforschen, werden Stimmen zunächst mithilfe des Jena Voice Learning and Memory Tests [JVLMT, 3] gelernt und erfolgreiches Stimmenlernen überprüft. In einem darauffolgenden Vertrauenswürdigkeitsrating werden gelernte und neue Stimmen anhand von inhaltslosen Pseudosätzen bewertet (Lernbedingung). Manipuliert wird zusätzlich das Ausmaß des Lernens (Lernerfahrung), d.h. ob der komplette JVLMT durchlaufen wurde (Gruppe_{lang}) oder nur die erste Lernphase (Gruppe_{kurz}). Wir gingen davon aus, dass bekannten (gelernten) Stimmen mehr vertraut wird als neuen (unbekannten) Stimmen, insbesondere dann, wenn Stimmen intensiv gelernt wurden. Ein solcher Befund hätte zahlreiche Implikationen, z.B. für Mechanismen der Meinungsbildung.

2. Methode

Stichprobe Die Stichprobe besteht aus zwei Gruppen mit jeweils 19 Personen, die sich im Ausmaß der Lehrerfahrung (kurz/lang) unterscheiden.

Gruppe_{kurz}: 9 weiblich, 10 männlich; Alter: M = 33,4 Jahre; SD = 15; Spanne: 20-68

Gruppe_{lang}: 11 weiblich, 8 männlich; Alter: M = 33,3 Jahre; SD = 14,3; Spanne: 21-69

Alle Probanden gaben an, deutsch auf Muttersprachniveau zu sprechen und keine Hörschäden zu haben, sowie keine der verwendeten Sprecher zu kennen.

Materialien Für das Vertrauenswürdigkeitsrating wurden zusätzlich zum JVLMT 24 Sprachdateien von Pseudosätzen aus der JVLMT Datenbank benötigt [3]. Die Stimuli wurden mit der PRAAT-Software [4] geschnitten und bearbeitet. Es handelte sich um je einen Pseudosatz von 8 gelernten Sprechern und 16 neuen Sprechern.

Prozedur Das Experiment wurde online über PsyToolKit [5, 6] durchgeführt. Im ersten Teil wurden mithilfe des JVLMT [3] 8 Stimmen gelernt. Der JVLMT besteht aus einer Lern-, Wiederholungs- und Testphase, wobei die Wiedererkennungslleistung sowohl in der Lernphase als auch in der Testphase mittels einer Dreifachwahlreaktionsaufgabe abgeprüft wird (Lernscore/Testscore). Die Versuchspersonen durchliefen den JVLMT entweder komplett (Gruppe_{lang}) oder beendeten ihn bereits nach der Lernphase (Gruppe_{kurz}). Es folgte das Vertrauenswürdigkeitsrating, mit 32 Trials, bei denen 16 unbekannte Stimmen je einmal und die 8 gelernten Stimmen je zweimal den gleichen Pseudosatz sprachen, welcher nicht zuvor im JVLMT vorkam. Sprecher wurden anhand ihrer Vertrauenswürdigkeit auf einer Likert-Skala von 1 bis 6 bewertet. In einem Zusatztest (siehe Poster

1) wurde danach die Glaubwürdigkeit von Aussagen der Sprecher beurteilt. Die Erhebung betrug ca. 30 min (Gruppe_{kurz}) bzw. 45 min (Gruppe_{lang})

3. Ergebnisse

Die Lernscores in der Lernphase im JVLMT der Probanden in Gruppe_{kurz} ($t(18) = 17.91, p < .05$) und der Probanden in Gruppe_{lang} ($t(18) = 11.99, p < .05$) zeigen einen signifikanten Effekt, was darauf hindeutet, dass ein Lernen in dieser Phase des Experiments stattfand.

In der Testphase des JVLMT, welcher nur von der Gruppe_{lang} durchlaufen wurde, zeigte sich auch hier ein signifikanter Lerneffekt ($t(18) = 13.28, p < .05$). Dies deutet darauf hin, dass vorher gelernte, bereits bekannte Stimmen überzufällig wiedererkannt wurden (Zufallsniveau = .33).

Die 2 x 2 gemischte ANOVA mit dem Innersubjektfaktor Lernbedingung (gelernt/neu) und dem Zwischensubjektfaktor Lernerfahrung (kur/lang) ergab weder einen signifikanten Haupteffekt der Lernbedingung auf die Vertrauenswürdigkeit ($F(1,36) < 1, p = .633$), noch einen signifikanten Haupteffekt der Lernerfahrung auf die Vertrauenswürdigkeit ($F(1,36) < 1, p = .0871$). Auch gab es keinen signifikanten Interaktionseffekt zwischen der Lernbedingung und der Lernerfahrung ($F(1,36) < 1, p = .989$).

4. Diskussion

Die Ergebnisse des JVLMT zeigen, dass die Stimmen überzufällig häufig wiedererkannt wurden und damit das Lernen funktioniert hat. Dennoch wurden diese Sprecher im Vertrauenswürdigkeitsrating nicht höher bewertet als unbekannte, was unsere Hypothese nicht unterstützt. Auch für die Lernerfahrung konnten wir weder einen Haupteffekt noch eine Interaktion mit der Lernbedingung nachweisen. Diese Befunde könnten darauf hindeuten, dass der Bekanntheitsgrad von Stimmen keinen Einfluss auf deren eingeschätzte Vertrauenswürdigkeit hat. Dieses Ergebnis ist teilweise im Einklang mit Lavan et al., die einmal einen signifikanten Trainingseffekt auf Vertrauenswürdigkeit finden konnten und zweimal nicht [2]. Die Gründe dafür sind noch weitestgehend unklar. Wir könnten einen potenziellen Effekt nicht gefunden haben, da unsere Messung nicht fehlerfrei ablief. Die Probanden berichteten von langen Ladezeiten und Verständnisproblemen, wodurch deren Konzentration und Aufmerksamkeit gesunken sein könnte. Zudem könnten die Stimmen zu wenig gelernt worden sein, um Vertrauenswürdigkeit zu entwickeln. Möglicherweise müssen Stimmen persönlich bekannt sein, um als vertrauenswürdig wahrgenommen zu werden. In der Poweranalyse haben wir einen mittleren Effekt geschätzt, der sich nach der Erhebung als verschwindend gering herausstellte ($\eta^2=0.00$), sodass wir ein Powerproblem weitestgehend ausschließen. Wir empfehlen, in Zukunft Laborerhebungen durchzuführen, um die Umweltbedingungen der Probanden kontrollieren zu können. Darüber hinaus könnte die Sprecherauswahl verbessert werden, sodass die Stimmen repräsentativer sind.

5. Literatur

- [1] McAleer, P., Todorov, A., & Belin, P. (2014). How do you say 'hello'? Personality impressions from brief novel voices. *PLoS ONE*, 9(3)
- [2] Lavan, N., Mileva, M. & McGettigan, C. (2020). *How does familiarity with a voice affect trait judgement?* *British Journal of Psychology*, 112(1), 282-300
- [3] Humble, D., Schweinberger, S.R., Mayer, A., Jesgarzewsky, T.L., Dobel, C. & Zäske, R. (2023). The Jena Voice Learning and Memory Test (JVLMT): A standardized tool for assessing the ability to learn and recognize voices. *Behavior research methods*, 55(3), 1352-1371
- [4] Boersma, P. & Weenink, D. (2023). Praat: doing phonetics by computer (Version 6.3) [Computer software]
- [5] Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096-1104
- [6] Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, 44(1), 24-31.

Trust me, you know me! Familiarität und Glaubwürdigkeit von Stimmen

Daniel Bruins, Leonie Friedrich, Jordana Knappe, Elisa Krebs, Carina Marit Smit

Leitung: Romi Zäske

1. Einleitung

Die Wiederholung von Aussagen erhöht deren wahrgenommene Glaubwürdigkeit [1]. Ähnlich dazu steigert bereits die einmalige Präsentation eines unbekanntes Gesichts dessen wahrgenommene Glaubwürdigkeit [2]. Unsere Studie untersucht, ob dieser Befund auch auf Stimmen übertragbar ist. So sollte die Wiederholung bzw. das Lernen einer unbekanntes Stimme die Vertrautheit (Familiarität) mit dem Sprecher und damit die Glaubwürdigkeit von dessen Aussagen erhöhen. Dazu manipulierten wir die Familiarität der Sprecher, indem einige zuvor mittels des Jena Voice Learning and Memory Test [JVLMT, 3] gelernt wurden. Zudem sagten wir vorher, dass Aussagen von intensiv gelernten Stimmen glaubwürdiger bewertet werden als von oberflächlich gelernten Stimmen. Dies untersuchten wir mit zwei Gruppen, welche jeweils wahrheitsambiguen Aussagen (Trivia) von mehreren Sprechern als richtig oder falsch beurteilen sollten. Ein Teil der Sprecher war unbekannt, während ein anderer Teil aus einer vorherigen Lernphase bekannt war (Lernbedingung gelernt/neu). Während eine Gruppe nur eine kurze Lernphase des JVLMT durchlief, absolvierte die andere den gesamten JVLMT (Lernerfahrung kurz/lang). Zusätzlich erhoben wir die wahrgenommene Vertrauenswürdigkeit der gelernten und neuen Stimmen.

2. Methode

Stichprobe Die Proband*innen wurden zufällig einer von zwei experimentellen Bedingungen (Lernerfahrung kurz/lang) mit je $n = 19$ zugewiesen. Die Gruppe_{kurz} bestand aus 9 Frauen und 10 Männern, die Gruppe_{lang} bestand aus 11 Frauen und 8 Männern. Der Altersdurchschnitt der Teilnehmenden lag bei $M = 33.3$ Jahren ($SD = 14.9$, Altersspanne 20 bis 69 Jahre). Es gab keine signifikante Abweichung in der Altersverteilung der Gruppen. Alle Proband*innen gaben an, Deutsch auf Muttersprachniveau zu sprechen und keine Hörprobleme zu haben.

Materialien Für das Experiment wurden der JVLMT [3] sowie 64 Sprachdateien für die Glaubwürdigkeitsaufgabe genutzt (32 junge Erwachsene x 2 Aussagen). Die noch unbearbeiteten und zu schneidenden Stimuli wurden mit der PRAAT-Software [4] aufbereitet.

Prozedur Das Online-Experiment wurde mittels PsyToolkit durchgeführt [5, 6]. Im ersten Teil wurden 8 Stimmen gelernt, wobei Gruppe_{kurz} nur die Lernphase des JVLMT durchlief, während Gruppe_{lang} den kompletten JVLMT, bestehend aus einer Lern-, Wiederholungs- und Testphase durchlief. Sowohl während der Lern- als auch während der Testphase des JVLMT wird die Wiedererkennungslleistung mittels Dreifachwahlreaktionsaufgaben und einer Ratewahrscheinlichkeit von .33 erfasst (Lernscore/Testscore). Im zweiten Teil wurde die wahrgenommene Vertrauenswürdigkeit der gelernten und nicht gelernten Stimmen erfasst (siehe Ergebnisse Poster 2). Im dritten Teil erfolgte in beiden Gruppen die Abfrage der Glaubwürdigkeit von 32 wahrheitsambiguen Aussagen (richtig/falsch), die von 8 gelernten Stimmen (zweimal präsentiert) und 16 neuen Stimmen gesprochen wurden. Die Dauer der Erhebung betrug für die Gruppe_{kurz} ca. 30 min. und für die Gruppe_{lang} ca. 45 min.

Statistische Analysen Als Manipulationscheck untersuchten wir mit Hilfe eines Einstichproben-t-Test, ob es bei dem JVLMT eine überzufällige Erkennungsrate ($> 33\%$ korrekte Antworten) für gelernte Stimmen und somit ein Lerneffekt gibt. Um den vorhergesagten Interaktionseffekt von Lernbedingung und Lernerfahrung auf die mittlere Zustimmungsrates zu den Triviaaussagen zu untersuchen, führten wir eine gemischte 2×2 ANOVA mit dem Innersubjektfaktor Lernbedingung (gelernt/neu) und dem Zwischensubjektfaktor Lernerfahrung (kurz/lang) durch.

3. Ergebnisse

Stimmenlernen In beiden Gruppen lag der Lernscore signifikant über Zufallsniveau, d.h. bereits nach einer kurzen Lernphase gab es einen signifikanten Lerneffekt (Gruppe_{kurz} ($M = 0.76$): $t(18) = 17.91$, $p < .001$; Gruppe_{lang} ($M = 0.82$): $t(18) = 11.99$, $p < .001$). Der Testscore am Ende des gesamten JVLMT (Gruppe_{lang}) lag ebenfalls signifikant über Zufallsniveau ($M = 0.79$), $t(18) = 13.28$, $p < .001$. **Glaubwürdigkeitsaufgabe** Es gab weder Haupteffekte von Lernbedingung und Lernerfahrung, noch einen Interaktionseffekt beider Faktoren auf die mittlere Zustimmungsgate ($M = 0.47$) zu den wahrheitsambigen Aussagen (alle $F(36) < 1$).

4. Diskussion

Durch unsere Untersuchung wurde nachgewiesen, dass es eine überzufällige Erkennungsrate für gelernte Stimmen gibt. Dennoch konnte kein Effekt von Familiarität der Stimme auf die Glaubwürdigkeit wahrheitsambiger Aussagen gefunden werden. Dies widerspricht Befunden zu Gesichtern, welche bereits nach einmaliger Präsentation als glaubwürdiger geratet wurden als unbekannte Gesichter [2]. Diese Diskrepanz könnte auf eine unterschiedliche Gewichtung von Stimme und Gesicht bei der Beurteilung von Glaubwürdigkeit hindeuten, sodass einer Quelle eher getraut wird, wenn deren Gesicht bekannt ist, nicht jedoch, wenn deren Stimme bekannt ist. Auch könnten methodische Unterschiede für die widersprüchlichen Befunde verantwortlich sein, da Brown et al. [2] Glaubwürdigkeit nicht anhand von Aussagen, sondern indirekt anhand des Gesichts an sich beurteilen ließen. Unser Maß könnte Glaubwürdigkeit daher direkter und verhaltensrelevanter erfassen. Die Intensität der Lernerfahrung hat ebenfalls keinen Effekt auf die Glaubwürdigkeit von Aussagen. Dies steht im Einklang mit Ergebnissen, wonach selbst intensives Stimmenlernen die wahrgenommene Vertrauenswürdigkeit von Stimmen nicht zuverlässig erhöht [7]. Möglicherweise wurden Stimmen trotz signifikanten Lerneffekts nicht intensiv genug gelernt, um einen Glaubwürdigkeitseffekt zu finden. Alternativ könnten unsere Nulleffekte auch auf der Unkonzentriertheit des Online-Settings der Proband*innen beruhen. Es wird vermutet, dass solch eine Aufgabe in einem Labor gewissenhafter bearbeitet wird. Auch technische Schwierigkeiten, wie die mehrfach zurückgemeldete lange Ladezeit von bis zu zehn Minuten zwischen der Lernphase und dem Glaubwürdigkeitsrating, sollten zukünftig vermieden werden, da es hierbei zu Ablenkung und Unkonzentriertheit kommen kann, was wiederum Auswirkung auf das Stimmengedächtnis hat. Dies könnte zukünftig mithilfe von Catchtrials überprüft werden. Trotz einer vergleichsweise großen Sprecheranzahl in der vorliegenden Studie ist nicht auszuschließen, dass die Wahl der Stimuli nicht repräsentativ ausfiel (nur 24 Sprecher*innen in ähnlichem Alter) oder die Effektstärke in der a-priori Poweranalyse überschätzt und infolgedessen die Stichprobe zu klein gewählt wurde.

5. Literatur

- [1] Brashier, N.M. & Marsh, E.J. (2020). Judging Truth. *Annual Review of Psychology*, 71:499-515
- [2] Brown, A.S., Brown, L.A. & Zoccoli, S.L. (2002). Repetition-based credibility enhancement of unfamiliar faces. *American Journal of Psychology*, vol.115, no.2, 199-209
- [3] Humble, D., Schweinberger, S.R., Mayer, A., Jesgarzewsky, T.L., Dobel, C., Zäske, R. (2023). The Jena Voice Learning and Memory Test (JVLMT): A standardized tool for assessing the ability to learn and recognize voices. *Behavior research methods*, 55(3), 1352-1371
- [4] Lavan, N., Mileva, M. & McGettigan, C. (2021). How does familiarity with a voice affect trait judgement?. *British Journal of Psychology*, 112(1), 282-300
- [5] Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096-1104
- [6] Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, 44(1), 24-31
- [7] Boersma, Paul & Weenink, David (2024). Praat: doing phonetics by computer [Computer program]. Version 6.4.13, retrieved 10 June 2024 from <http://www.praat.org/>

Money, Money, Money! Does your political orientation determine whether you explore and invest in a local or foreign bank?

Manzke, Lisa; Möller, Mareike; Scheipers, Merle; Thiele, Victoria; Weihert, Marlene

Leitung: Michael Edem Fiagbenu

1. Introduction

Understanding the psychological differences between leftists and rightists can provide deeper insights into the origins of political beliefs. Research suggests that leftists are generally more risk-averse and less likely to invest in financial products and the stock market, which explains leftists' dislike of free-market capitalism (Fiagbenu & Kessler, 2022; Kaustia & Torstilla, 2011). Emerging evidence, however, suggests that the relationship between financial threat-sensitivity depends on the psychological features of the stimulus (Choma & Hodson, 2017). But is it unclear whether leftists' and rightists' financial exploration and attitudes differs as a function of the local or foreign origin of the financial stimulus. We hypothesised that leftists should be more likely to take financial risks in their local compared to foreign banks as the profits can be used to develop their local economies (i.e., home bias). But rightists should be more likely to take financial risks regardless of the origin of the bank. The current study therefore examined the relationship between political orientation and financial risk-taking as a function of stimulus origin.

All analysis codes, data and preregistration document can be found at the OSF website <https://osf.io/ext8y/>

2. Methods

Participants. A total of 118 students and community participants ($M_{\text{age}} = 22.8$, $SD = 7.40$, range = 18-64, female = 96, non-binary = 3) from Friedrich Schiller University of Jena participated in a 15 minute online "Sosci Survey" study about political attitudes and investment decisions.

Design and Procedure. We randomly assigned participants to two experimental conditions where they read a vignette describing a newly established local or foreign investment bank in Germany, offering the opportunity to invest in stocks. After reading the text, we measured risk taking (i.e., exploratorion intentions and attitudes towards the bans, and political orientation.

Measures. We measured participants exploratory intentions (5 statements, e.g., I will buy stocks or open an account in this bank) and attitudes towards the bank with (5 statements, e.g., I will recommend this bank to a friend). All statements were scored on a 5-point Likert scale: (1 = disagree, 5 = agree).

Attitudes towards the bank. Participants indicated their

Political orientation. Participants also indicated their political orientation on an 11-point scale: (1 = leftist, 11 = rightist)

Data analysis. We used the R statistical software to conducted t-test and regressions to examine differences between both experiment conditions and also how political orientation predicts exploration intention and attitudes towards the bank.

3. Results

Differences in Exploratory intentions and bank attitudes. A t-test revealed no significant difference in exploratory intentions between the local ($M = 2.57$, $SD = .75$) or foreign ($M = 2.62$, $SD = .68$) bank

conditions, $t(38) = 108, p = .7$. There was also no significant difference in attitudes towards the local ($M = 2.66, SD = .68$) or foreign ($M = 2.71, SD = .74$) bank conditions, $t(38) = 115.09, p = .7$.

Political Orientation, Exploration and Bank attitudes as a Function of Stimulus Origin. Regression analysis revealed no significant main effect of political orientation, $\beta = -.05, SE = .06, p = 0.38$, or stimulus origin (coded as local = 1 and foreign = 0), $\beta = -.41, SE = .31, p = 0.19$, on exploration intentions. There interaction effect of political orientation and stimulus origin on exploratory intentions was also not significant, $\beta = .09, SE = .07, p = .21$.

Also, there was no significant main effect of political orientation, $\beta = -.008, SE = .06, p = 0.89$, or stimulus origin (coded as local = 1 and foreign = 0), $\beta = -.084, SE = .32, p = .79$, on attitude formation. There interaction effect of political orientation and stimulus origin on attitude formation was also not significant, $\beta = .009, SE = .08, p = .91$.

4. Discussion

The results show that leftists and rightists did not differ significantly in their financial exploration and attitudes toward stocks as a function of local or foreign banks. The lack of support for our hypothesis may be primarily due to our small sample size. Future studies can rely on the current results to estimate the sample size required to produce a significant effect. Additionally, the study was conducted with mostly students who have below-average wealth and may not have much experience in investment behavior. Future studies should use more diverse participants with a broader age range and differentiated socio-economic status.

5. References

- Choma, B. L., & Hodson, G. (2017). Right-wing Ideology: Positive (and negative) relations to threat. *Social Cognition, 35*(4), 415-432. <https://doi.org/doi:10.1521/soco.2017.35.4.415>
- Fiagbenu, M. E., & Kessler, T. (2022). Fear and loathing of wall street: Political liberalism, uncertainty, and threat management in a dangerous economic world. *Political Psychology, 43*(6), 1101-1121. <https://doi.org/10.1111/pops.12805>
- Kaustia, M., & Torstila, S. (2011). Stock market aversion? Political preferences and stock market participation. *Journal of Financial Economics, 100*(1), 98-112. <https://doi.org/10.1016/j.jfineco.2010.10.017>

Would you eat this fruit? Effect of political orientation on exploration and attitudes towards foreign and local foods

Torben Dolgner, Sophia Hering, Isabelle Scheibler, Lina Bulmahn

Leitung: Michael Edem Fiagbenu

1. Introduction

Rightists are generally more sensitive to physical threats and exhibit less exploration in health/food foraging tasks than leftists (Shook & Fazio, 2009; Fiagbenu, Proch & Kessler, 2021). This difference is believed to explain why rightists are more sensitive to terrorism and immigration threats than leftists. However, recent studies show that ideological differences in food/health attitudes depend on the psychological features of the food stimuli (Purhonen & Heikkilä, 2017). But it remains an open question whether the geographic origin of the food determines exploration and attitudes among leftists and rightists. Because rightists are usually more nationalistic than leftists, we hypothesize that they will explore and report more positive attitudes towards local than foreign foods. In contrast, because leftists take more physical risks than rightists, they should be more likely to explore and report more positive attitudes towards both local and foreign foods. In the current study we investigated whether the origin of food stimuli influence exploration and attitude formation among leftists and rightists.

The pre-registration document, data and analysis scripts are available on the Open Science Framework repository at https://osf.io/p5myu/?view_only=774afa78e04244aa954281e275166216.

2. Method

Participants. We recruited 215 participants ($M_{age} = 22.31$, $SD = 5.39$, range = 18-64) from students from the Friedrich Schiller University Jena and community sample using the online SosciSurvey platform to participate in a study about political attitudes and food exploration.

Design. Participants were randomly assigned to one of two conditions foreign ($N_{foreign} = 113$) or local ($N_{local} = 102$) food. In the local food condition, they read about a newly discovered fruit from a remote region in Germany and in the foreign fruit condition they read about a newly discovered fruit from a foreign country. We used novel stimuli to avoid stimuli towards which participants already had prior attitudes. This will enable us to examine how the specific salient of stimuli origin alone influence attitudes and exploration. After reading the text, we measured risk taking (i.e., exploration intentions and attitudes towards the fruits, and political orientation).

Measures. We measured participants' exploratory intentions (e.g., I will pick up the fruit to smell it or I will buy the fruit) and attitudes towards the fruit (e.g., I will recommend this fruit to a friend). All statements were scored on a 5-point scale: (1 = disagree, 5 = agree). Participants also indicated their political orientation on an 11-point scale: (1 = leftist, 11 = rightist)

Data analysis. All data was analyzed using R statistical software. We conducted t-tests and regressions to examine differences between both experimental conditions and how political orientation predicts exploration intention and attitudes towards the fruits. We also controlled for purchase frequency, importance of fruits (self-invented) and demographic variables (e.g. age, gender).

3. Results

Differences in Exploratory Intentions and Fruit Attitudes. A t-test showed no significant differences, $t(-0.68) = 209.36$, $p = .5$ in intentions to explore foreign ($M = 4.12$, $SE = .096$) and local fruits ($M = 4.06$, $SE = 0.096$). Also, attitudes towards the fruit did not significantly differ in the foreign ($M = 4.46$, $SE = .076$) and local fruit ($M = 4.45$, $SE = .076$) conditions $t(-.03) = 207.49$, $p = .98$.

Relationship Between Political Orientation, Exploration and Attitudes as a Function of Stimulus Origin. A regression analysis revealed a significant main effect of political orientation, $\beta = -.09$, $SE = .044$, $p = .033$, but not stimulus origin, $\beta = -.15$, $SE = .234$, $p = .52$, on exploratory intentions.

There was no significant interaction effect of political orientation and stimulus origin on exploratory intentions, $\beta = .065$, $SE = 0.062$, $p = .30$). Also, political orientation, $\beta = -.096$, $SE = .034$, $p = .005$ but not stimulus origin, $\beta = -.17$, $SE = 0.182$, $p = .35$ significantly predicted attitudes towards the fruits. There was no significant interaction effect of political orientation and stimulus origin on attitudes towards, $\beta = .052$, $SE = .049$, $p = .29$.

4. Discussion

Contrary to our hypotheses, the stimulus origin did not moderate the relationship between political orientation, exploration intentions and attitudes towards the local or foreign fruits. Specifically, we found that leftists and rightists did not differ in their exploration intentions and attitudes towards as a function of stimulus origin. One limitation is that our sample was not representative of the target population as Psychology students are generally left-oriented. Secondly, political topics are controversial and thus social desirability could have distorted the findings. We hope future studies can address these limitations.

5. Literature

- Fiagbenu, M. E., Proch, J., & Kessler, T. (2021). Of deadly beans and risky stocks: Political ideology and attitude formation via exploration depend on the nature of the attitude stimuli. *British Journal of Psychology*, *112*(1), 342-357.
- Purhonen, S., & Heikkilä, R. (2017). Food, music and politics: The interweaving of culinary taste patterns, 'highbrow' musical taste and conservative attitudes in Finland. *Social Science Information*, *56*(1), 74-97.
- Shook, N. J., Fazio, R. H. (2009). Political ideology, exploration of novel stimuli, and attitude formation. *Journal of Experimental Social Psychology*, *45*(4), 995-998.

Programm

14.00 Uhr	Foyer	Eröffnung der Posterausstellung
14.00 – 16.00	Foyer	Präsentation der Poster
16.15 Uhr	HS 5	Festvortrag von Dr. Dwight Tse, Lecturer Psychology, University of Strathclyde Glasgow zum Thema: <i>„Let it Flow? The Challenges in Manipulating Flow (Optimal) Experience in Experiments“</i>
17.00 Uhr	HS 5	Auszeichnung der besten Poster
anschließend	Foyer	Ausklang der Veranstaltung

Thalia Buchhandlung

»Neue Mitte Jena«

Leutragraben 1

07743 Jena

Tel. 03641 4546-0

www.thalia.de

