

FULL PAPER

Priming von Stereotypen durch geschlechtergerechte Sprache in journalistischen Texten

Priming stereotypes through gender-sensitive language in journalistic texts

Frederieke Cirksena & Dominik J. Leiner

Frederieke Cirksena (B.A.), University of Groningen. Contact: frederieke.cirksena(at)gmx.net. ORCID: <https://orcid.org/0000-0001-7851-9118>

Dominik J. Leiner (Dr.), Ludwig Maximilian University of Munich, Department of Media and Communication, Oettingenstraße 67, 80538 München, Germany. Contact: dominik.leiner(at)ifkw.lmu.de. ORCID: <https://orcid.org/0000-0002-3862-3399>



© Frederieke Cirksena, Dominik J. Leiner

Priming von Stereotypen durch geschlechtergerechte Sprache in journalistischen Texten

Priming stereotypes through gender-sensitive language in journalistic texts

Frederieke Cirksena & Dominik J. Leiner

Abstract: In der vergangenen Dekade gewann geschlechtergerechte Sprache zunehmend an Akzeptanz und Verbreitung. Formulierungen, die Frauen bei der Nennung sozialer Gruppen sprachlich einbeziehen, sorgen bei Rezipientinnen und Rezipienten für mehr kognitive Repräsentation von Frauen. Allerdings wurde bisher kaum untersucht, ob geschlechtergerechte Formulierungen nicht-intendierte Nebenwirkungen entfalten: Die ständige Betonung des Geschlechts könnte die Geschlechter-Diskriminierung im Denken verstärken. Wir untersuchten anhand von zwei Befragungsexperimenten, ob solch ein Effekt in der Realität tatsächlich beobachtet werden kann. Die Stichprobe wählten wir ausreichend groß ($N = 2.066$), sodass selbst ein schwacher Effekt entdeckt werden würde. Dennoch fanden wir in einer akademisch geprägten und liberal orientierten Stichprobe keine Hinweise auf die skizzierten Nebenwirkungen – im Gegenteil deuten die Befunde darauf hin, dass geschlechtergerechte Sprache zumindest im befragten Milieu mittlerweile sprachliche Normalität geworden ist.

Schlagwörter: Geschlechtergerechte Sprache, Stereotype, Priming, Experimente.

Abstract: Over the past decade, gender-sensitive language has become increasingly accepted and widespread in Germany. Wording that includes women when mentioning social groups increases recipients' cognitive representation of women. However, there has been little research on whether gender-sensitive wording has unintended side effects: The constant emphasis on gender might reinforce gender discrimination in thinking. In two survey experiments, we investigated whether such an effect exists. We chose a sufficiently large sample ($N = 2,066$) to detect even a weak effect. Nevertheless, we found no evidence of the outlined side effects in a liberal and academically inclined sample. On the contrary, the findings suggest that gender-sensitive language has become linguistic normality, at least in the milieu surveyed.

Keywords: gender-sensitive language, stereotypes, priming, experiments, German language.

1. Einleitung

Das generische Maskulinum führt häufig dazu, dass soziale Gruppen in der Wahrnehmung und im Denken durch Männer repräsentiert werden (Hellinger, 1990; Pusch, 1979; Sczesny & Stahlberg, 2001). Dies kann vorhandene Diskri-

minierung stützen und andere Maßnahmen zur Gleichberechtigung von Frauen und Männern behindern (Demarmels & Schaffner, 2011). Werden weibliche und männliche Vertreter sozialer Gruppen explizit benannt („geschlechtergerechte Sprache“), so führt dies zu einem stärkeren gedanklichen Einbezug von Frauen (Blake et al., 2008; Blake & Klimmt, 2010; Sczesny & Stahlberg, 2001, 2005). Geschlechtergerechte Formulierungen finden zunehmend Verbreitung (Acke, 2019) und Akzeptanz (Steiger & Irmen, 2011); letztere hängt auf individueller Ebene aber stark mit der formalen Bildung zusammen (Lindqvist et al., 2019; Parks & Robertson, 2008).

Vergoossen et al. (2020) liefern eine Systematik möglicher Nachteile geschlechtergerechter Sprache. Die Sorge, geschlechtergerechte Formulierungen würden das Textverständnis einschränken, erwies sich als unbegründet (Friedrich & Heise, 2019; Pöschko & Prieler, 2018). Empirische Belege finden sich für eine reduzierte Lesegeschwindigkeit (Pöschko & Prieler, 2018) und eine schlechtere Beurteilung von geschlechtergerecht formulierten Texten (Blake et al., 2008; Rothmund & Scheele, 2004). Allerdings leidet die Beurteilung der Textästhetik (Blake & Klimmt, 2010) nur bei sperrigen Formulierungen, ein Effekt, welcher auch beim generischen Maskulinum zu beobachten ist (Trutkowski, 2018). Die Beurteilung der Textästhetik dürfte zudem damit zusammenhängen, ob die Formulierung zu den sprachlichen Gewohnheiten passt – schließlich genießen einheitliche Sprach- und Schreibnormen in Deutschland einen hohen Stellenwert. Diese Problematik verliert insofern an Relevanz, als geschlechtergerechte Sprache innerhalb bestimmter Kontexte zunehmend als normal erlebt wird (Acke, 2019).

Was bislang kaum untersucht wurde, sind unerwünschte Nebeneffekte geschlechtergerechter Sprache im Bereich der mentalen Repräsentation. Wir werfen die These auf, dass geschlechtergerechte Texte das Merkmal *Geschlecht* durch Priming-Effekte (Scheufele, 2016; Weingarten et al., 2016) besonders salient machen. Dadurch könnte das Geschlecht zum zentralen Erklärungsmuster avancieren. Der vorliegende Aufsatz stellt Befunde aus zwei Experimenten vor¹. Diese überprüften die Hypothese, geschlechtergerechte Formulierungen würden die Interpretation journalistischer Texte beeinflussen.

2. Theorie

Wenn wir „Ärztinnen und Ärzte“, „Professorinnen und Professoren“, „weibliche und männliche Führungskräfte“ stets explizit benennen, dann beinhaltet unser mentales Bild dieser Gruppen deutlich häufiger auch weibliche Vertreterinnen (Sczesny & Stahlberg, 2001). Zugleich macht der Text aber auch immer wieder bewusst: Unsere Sprache unterscheidet zwischen Frauen und Männern. Der Text untergliedert die soziale Gruppe immer wieder in zwei Teilgruppen, wiederholt stetig, dass es einen Unterschied gibt zwischen Frauen und Männern, dass wir „Richterinnen“ von „Richtern“ unterscheiden. Es gibt neben dem Geschlecht viele andere Merkmale, die zu Diskriminierung führten und führen, wie beispielswei-

1 Datensatz, Dokumentation und Auswertung sind öffentlich verfügbar unter <https://osf.io/ga4cp/>

se Alter, Hautfarbe, kulturelle Herkunft, Zugehörigkeit zu einer sozialen Schicht oder formale Bildung. Und abhängig vom Kontext gibt es auch Personenmerkmale, welche die Mitglieder einer sozialen Gruppe sinnvoll in Teilgruppen teilen: Wenn es um berufliche Leistung geht, könnten das beispielsweise individuelles Engagement sein oder familiäre Pflichten, welche mit dem Beruf konkurrieren. Dies mag ein Text explizit ansprechen, doch daneben bezeichnet die getrennte Nennung weiblicher Vertreterinnen und männlicher Vertreter stets das Geschlecht als Personenmerkmal. Geschlechtergerechte Texte heben das Merkmal *Geschlecht* zwangsläufig hervor.

2.1. Priming

Aus der Psychologie ist der Effekt bekannt, dass die Präsentation von Begriffen die Interpretation von nachfolgenden, ambivalenten Reizen beeinflusst (Weingarten et al., 2016). Dieser als *Priming* bezeichnete Effekt wurde von der Kommunikationswissenschaft aufgegriffen und aus dem individuellen in das gesellschaftliche Bewusstsein übertragen (Scheufele, 2016). Die ständige kognitive Aktivierung des Merkmals *Geschlecht* könnte auf Dauer dazu führen, eine mentale Diskriminierung (zunächst im Sinne von Unterscheidung) zu verfestigen (Degner et al., 2009). Dieser Effekt stünde natürlich in starker Konkurrenz zu anderen Ursachen von Geschlechterdiskriminierung (Diekman & Eagly, 2000; Sczesny et al., 2006; Steele, 1997).

Der Fokus des vorgestellten Experiments liegt deshalb auf kurzfristigen Effekten. Wir stellen die Frage, ob geschlechtergerechte Texte als *Prime* im psychologischen Sinne dienen können, welcher das Geschlecht als relevantes Merkmal salienter macht als andere Charakteristika von Personen. Solch ein Prime wäre in der Lage, die Interpretation ambivalenter Texte zu beeinflussen. Raum für Ambivalenz bietet die Komplexität und Unsicherheit der Realität bei nahezu jedem Thema der medialen Berichterstattung (Dunwoody, 2020; Friedman et al., 1999; Gill & Babrow, 2007).

2.2. Aktivierung von Geschlechterstereotypen

Wenn ein journalistischer Text das Geschlecht als Personenmerkmal betont, kann dies vorhandene Geschlechterstereotype aktivieren (Collins & Loftus, 1975; Eckes, 1997), wodurch diese als Schemata (Scheufele, 2004) an Salienz für die Interpretation des Textes gewinnen. Erlaubt ein Text unterschiedliche Erklärungsmuster, so folgen Leserinnen und Leser mit aktivierten Geschlechterstereotypen wahrscheinlicher einer Interpretation, bei welcher das Geschlecht die zentrale Ursache für Verhalten, Erfolg oder Misserfolg ist.

Studien von Sczesny und Stahlberg (2002) zeigen, dass bereits kleine Hinweisreize (z.B. die Anwesenheit des jeweiligen Geschlechts oder ein Männer- oder Frauenduft) genügen, um Geschlechterstereotype zu aktivieren. Darüber hinaus legen die Befunde von Eckes (2008) sowie von Steffens und Ebert (2016) nahe, dass Geschlechterstereotype trotz gewandelter Geschlechterrollen nach wie vor großen Einfluss auf unsere Wahrnehmung haben. Dies gilt nicht nur für explizite

Vorurteile, sondern ebenso für implizite Einstellungen (White & White, 2006). Das Modell der parallelen Informationsverarbeitung von Kunda und Thagard (1996) postuliert als konnektivistisches Netzwerkmodell, dass die Aktivierung einer Charakteristik auch zur Aktivierung von damit assoziierten Eigenschaften führt. Wenn also das Konzept *Frau* aktiviert wird, kann dies zur Aktivierung von Eigenschaften führen, die stereotyp mit Frauen verbunden werden, und Einfluss auf die Interpretation einer Handlung nehmen (Schmid Mast & Krings, 2008).

2.3. Geschlechterstereotype bei Berufen

Voraussetzung für den oben skizzierten Zusammenhang ist das Vorhandensein von geschlechtsbezogenen Stereotypen. Solche Stereotype werden zwar gesellschaftlich geteilt (konsensueller Stereotyp-Anteil), die Ausgestaltung und Intensität unterliegt aber individuellen Schwankungen (Eckes, 2008).

Ausführlich wurden gesellschaftliche Geschlechterstereotype bei Berufen untersucht (Lawrimore & Marcis, 1994; Parks & Robertson, 1998). Eine wesentliche Rolle, ob ein Beruf als männlich oder weiblich wahrgenommen wird, spielt demnach der Anteil von Männern und Frauen in dem Berufsfeld (Barth & Rice, 2016; Glick et al., 1995; Sáinz et al., 2016), zu einem geringeren Anteil auch die für die Tätigkeit präferierten Persönlichkeitsmerkmale (White & White, 2006). So gelten beispielsweise Pflegeberufe als weiblich, weil sie von Fachkräften Fürsorge erfordern – ein stereotypisch weibliches Merkmal (White & White, 2006).

2.4. Formen geschlechtergerechter Sprache

Als wäre die Verwendung geschlechtergerechter Formulierungen in journalistischen Publikationen nicht bereits mit ausreichend Risiken verbunden, besteht bislang auch kein Konsens über die genaue Form der Formulierung. Neben der *Paarform* („Leserinnen und Leser“) finden *Schrägstrichvarianten* („Leser/innen“, „Leser/-innen“ oder „Leser/Leserinnen“), *Klammern* („Leser(innen)“), das *Binnen-I* („LeserInnen“), seit 2003 die *Gender-Gap* („Leser_innen“; Herrmann, 2003) und seit 2015 das *Gendersternchen* („Leser*innen“; Diewald & Steinhauer, 2017) und der *Gender-Doppelpunkt* („Leser:innen“; kleinergast, 2016) Verwendung, um weibliche Mitglieder einer sozialen Gruppe besser zu repräsentieren als das generische Maskulinum („Leser“). Die drei letztgenannten Formen berücksichtigen zudem nicht nur Frauen und Männer, sondern sollen auch nicht-binäre und diversgeschlechtliche Menschen einbeziehen (Diewald & Steinhauer, 2017).

2.5. Kognitive Effekte geschlechtergerechter Sprache

Während das generische Maskulinum eher männliche Assoziationen weckt (Klann-Delius, 2005), kann das Binnen-I zu vorwiegend weiblichen Assoziationen führen aufgrund der starken schriftsprachlichen und phonologischen Ähnlichkeit zum Femininum („LeserInnen“ und „Leserinnen“; Rothmund & Scheele, 2004; Sczesny & Stahlberg, 2001, 2005). Für das Gendersternchen liegen noch keine empirischen Befunde vor, da es erst seit 2015 zunehmend Verwendung findet –

insbesondere im Hochschulkontext und in öffentlichen Verwaltungen (Rat für deutsche Rechtschreibung, 2018). Aufgrund der phonologischen Ähnlichkeit zum Femininum ist ein *female bias* jedoch zumindest plausibel. Wie beim Binnen-I wird zudem bei manchen Personenbezeichnungen im Plural nur die weibliche Variante optisch vollständig dargestellt („Professor*innen“).

2.6. Einflussfaktoren auf die Textinterpretation

Dass unterschiedliche Leserinnen und Leser einen Text unterschiedlich interpretieren, wurde hinlänglich gezeigt (Artelt et al., 2001; Kintsch, 1974). Deren Geschlecht ist hier nur ein Faktor unter vielen (Howard & Allen, 1990; Iser, 1980; Kolodny, 1980; Rubin, 1993). Ebenso ist bekannt, dass eine Person denselben Text in unterschiedlichen Situationen und Kontexten unterschiedlich interpretiert (Groeben, 1982). Die Rezipientin oder der Rezipient konstruiert (Schmitz, 2016) auf Basis des Texts unterschiedliche mentale Repräsentationen (*Situationsmodelle*; Dutke, 1998). Dabei werden verfügbare Hinweisreize (*cues*; Dascal & Weizman, 1987; Roskos-Ewoldsen et al., 2008) ausgedeutet und fehlende Informationen durch Kontext (Parks & Roberton, 1998; Rothmund & Scheele, 2004), Vorwissen und andere kognitiv verfügbare Inhalte (Baden & Lecheler, 2012; Brewer et al., 2003; Domke et al., 1998; Matthes, 2007) ergänzt.

2.7. Unerwünschte Effekte geschlechtergerechter Sprache

Geschlechtergerechte Sprache kann neben der mentalen Repräsentation von Frauen – einem erwünschten Effekt – auch zu unerwünschten Effekten führen. So wurden beispielsweise Bewerberinnen, die man mit einer weiblichen Berufsbezeichnung vorstellte, negativer bewertet als Bewerberinnen, denen eine männliche Berufsbezeichnung zugewiesen wurde (Formanowicz et al., 2013). Zudem kann die Verwendung der Paarform für Berufsbezeichnungen dazu führen, dass der Beruf mit einem geringeren Einkommen assoziiert wird (Horvath et al., 2016).

Wie eingangs dargelegt, soll auch in diesem Aufsatz ein nicht-intendierter kognitiver Effekt geschlechtergerechter Sprache betrachtet werden: Fungiert die Verwendung geschlechtergerechter Formulierungen in journalistischen Texten als Prime, welcher die Dichotomie Männer/Frauen im Bewusstsein der Rezipientinnen und Rezipienten aktualisiert? Gewinnen Geschlechterstereotype dadurch an Salienz, sodass sie wahrscheinlicher als Interpretationsmuster für die rezipierten Texte angewendet werden?

3. Forschungsdesign

Zur Überprüfung, ob geschlechtergerechte Sprache die Interpretation journalistischer Texte beeinflusst, greifen wir zwei traditionelle Stereotype auf: Zum einen, dass Männer kompetenter mit Technik umgehen würden als Frauen. Hierbei handelt es sich um eine Konkretisierung des generellen Konzepts der *Instrumentalität* (Eckes, 2008). Zum anderen, dass Frauen kompetenter mit Kindern umgehen würden als Männer. Eigenschaften wie *fürsorglich* und *zärtlich* wurden von Tro-

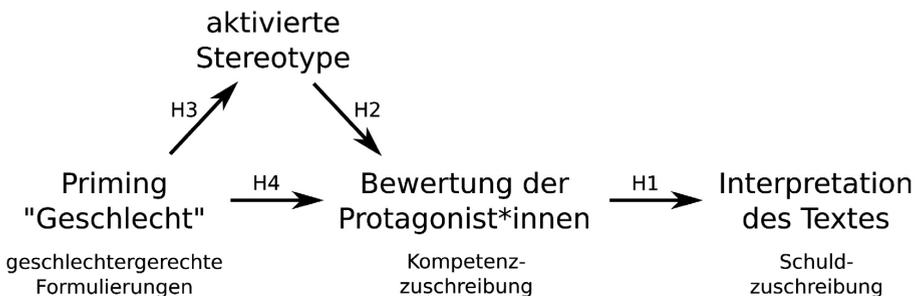
che und Rammsayer (2011) in einer Revision des Bem-Sex-Role-Inventory (BSRI) als sozial erwünschte Eigenschaften – und somit präskriptive Stereotype – für Frauen identifiziert. Prentice und Carranza (2002) führen auch *interest in children* als erwünschte Eigenschaft für Frauen auf. Ein Stereotyp, dass Männer generell kompetenter wären (Eckes, 2008) ist empirisch nur eingeschränkt belegt (Alfermann, 1992) und sollte deutlich weniger Einfluss nehmen als das spezifischere Stereotyp, wonach der fürsorgliche Umgang mit Kindern Frauensache sei. Die starke gesellschaftliche Assoziation zwischen Technik und Männlichkeit thematisieren etwa Axell und Boström (2019).

Zu jedem der beiden Stereotype wurde in Anlehnung an tatsächliche Nachrichtenbeiträge ein Nachrichtenartikel mit einer Länge von ca. 2.000 Zeichen formuliert (Anhang 1). Der Text schildert einen Schadensfall und erlaubt zwei Interpretationen zu der Frage, wer den Schaden zu verantworten hat: Die Rahmenbedingungen oder die Fachkräfte.

Die beiden Texte dienten als Stimulus in zwei getrennten einfaktoriellen Experimenten: Der jeweilige Text wurde Leserinnen und Lesern entweder im generischen Maskulinum, in der Paarform oder mit Gendersternchen vorgelegt.

Die Fachkräfte, die den Schaden potenziell verantworten, sind jeweils Gruppen bestehend aus Männern und Frauen. Die Beteiligung von Frauen geht aber nur dann explizit aus dem Text hervor, wenn dieser nicht im generischen Maskulinum formuliert ist. Bei Verwendung des generischen Maskulinums können die Leserinnen und Leser plausibel annehmen, es handle sich bei den „App-Entwicklern“ nur um Männer, denn es sind weit überwiegend Männer, die Software programmieren (Bitkom, 2019; Nier, 2018). Bei den „Erziehern“ einer Kinderkrippe hingegen konkurriert im generischen Maskulinum der *male bias* der grammatischen Form mit berufsspezifischen Stereotypen (Blake & Klimmt, 2010; Irmen & Linner, 2005) aufgrund der weiblichen Dominanz in Pflege- und Erziehungsberufen (Cremers et al., 2015).

Abbildung 1. Empirisches Modell



Zu überprüfen ist, ob die geschlechtergerechte Formulierung des Textes Stereotype aktiviert. Diesen Stereotypen zufolge wäre zumindest ein Teil der Fachkräfte (weibliche Programmiererinnen und männliche Erzieher) weniger kompetent, als

es Fachkräfte mit dem berufstypischen Geschlecht wären. Wenn das mentale Modell der Leserinnen und Leser von inkompetenten Fachkräften ausgeht, äußert sich das in der Kompetenzzuschreibung und darin, dass die Schuld für den Schaden eher bei den Fachkräften als bei den Rahmenbedingungen gesehen wird. Dieser Zusammenhang ist natürlich vorab zu prüfen:

H1. Leserinnen und Leser, welche die Fachkräfte als weniger kompetent beurteilen, sehen die Schuld eher bei den Fachkräften als bei Situation und Rahmenbedingungen.

Ebenfalls vorab zu prüfen ist, ob aktivierte Stereotype einen Einfluss auf die Bewertung der Fachkräfte haben:

H2. Leserinnen und Leser, bei welchen geschlechtsspezifische Stereotype in hohem Maße aktiv sind, bewerten die Kompetenz von (auch weiblichen) Fachkräften im Bereich der Softwareentwicklung und (auch männlichen) Fachkräften im Bereich der Kindererziehung schlechter als Leserinnen und Leser, bei denen diese Stereotype in geringerem Maße aktiv sind.

Um den Einfluss geschlechtergerechter Formulierungen auf die Interpretation des Textes zu untersuchen, sind folgende Kernhypothesen zu prüfen:

H3. Wenn der Text eine geschlechtergerechte Formulierung verwendet, so werden geschlechtsspezifische Stereotype in höherem Maße aktiviert, als wenn der Text im generischen Maskulinum formuliert ist.

Weil die Messung von individuell vorhandenen Stereotypen mit starken Verzerrungen einhergehen kann, etwa mit sozialer Erwünschtheit, unterscheidet die empirische Umsetzung zwischen der individuellen Ausprägung von Stereotypen und gesellschaftlich verbreiteten Stereotypen. Selbst wenn die Leserinnen und Leser sich individuell nicht oder nur geringfügig in der Ausprägung der Stereotype unterscheiden, wäre aufgrund der gesellschaftlich verbreiteten Stereotype folgender Unterschied zu erwarten:

H4. Wenn der Text eine geschlechtergerechte Formulierung verwendet, so wird den Fachkräften eine geringere Kompetenz zugesprochen, als wenn der Text im generischen Maskulin formuliert ist.

Gemeinsam mit dem Zusammenhang zwischen Kompetenz- und Schuldzuschreibung (H1) ergibt sich folgende Hypothese zur Wirkung geschlechtergerechter Sprache auf die Textinterpretation:

H5. Wenn der Text eine geschlechtergerechte Formulierung verwendet, so wird die Schuld eher den Fachkräften zugeschrieben, als wenn der Text im generischen Maskulinum formuliert ist.

Ergänzend zu diesen Hypothesen überprüfen wir, ob die Effekte mit dem Geschlecht der Leserinnen und Leser variieren, ob Paarform und Gendersternchen zu unterschiedlichen Effekten führen und ob die geschlechtergerechte Formulierung sich auf die Lesbarkeit des Textes auswirkt.

4. Stichprobe

Das Experiment wurde mithilfe eines Online-Fragebogens realisiert. Mithilfe des *SoSci Panels* konnten im Juni 2020 insgesamt 2.888 Personen rekrutiert werden, um an der Online-Befragung teilzunehmen („Studie zur journalistischen Aufbereitung von Unfällen“). Beim *SoSci Panel* handelt es sich um einen nicht-repräsentativen Pool für wissenschaftliche Befragungen, der sich insbesondere für experimentelle Erhebungen eignet (Leiner, 2016a). Der anvisierte Rücklauf von jeweils 930 Männern und Frauen mit dem Ziel einer statistischen Power von .80 (geschätztes $d = 0.20$, $\alpha = .95$) wurde übertroffen, wahrscheinlich eine Folge der seit März 2020 wegen der Corona-Pandemie geltenden Einschränkungen für Arbeit und Freizeitgestaltung.

Die Teilnehmerinnen und Teilnehmer wurden unter der Randbedingung gleicher Gruppengrößen zufällig einer der Experimentalgruppen in einem der beiden Experimente zugewiesen. Aus der Analyse ausgeschlossen wurden 615 Abbrecherinnen und Abbrecher sowie Personen, die entweder jünger als 15 Jahre waren, nicht in die Verwendung ihrer Angaben einwilligten oder nicht in Deutschland wohnhaft waren. Weiterhin ausgeschlossen wurden 11 Datensätze, in denen mehr als 10 % der Items (gewichtet, Leiner, 2019) nicht beantwortet wurden und 196 Datensätze, bei welchen die Befragten den Stimulustext nur sehr kurz (weniger als die Hälfte des Medians für diesen Text) oder so lange betrachteten, dass eine Nebenbeschäftigung anzunehmen war (Verweildauer länger als der dreifache Median). Durch die parallele Auswertung des unbereinigten Datensatzes wurde sichergestellt, dass die strenge Datenbereinigung keine Artefakte (abweichende Befunde) verursacht. Mithin flossen nach Bereinigung Datensätze von 954 Frauen, 1.105 Männern und 7 Personen, die sich als *divers* einordneten, in die Auswertung ein ($N = 2.066$), entsprechend 144 bis 198 Personen pro Experimentalgruppe und Geschlecht (ohne diverse Geschlechter, Anhang 5.3). Die Befragten verteilen sich gleichmäßig über die Altersgruppen ab 25 Jahre (erhoben in Kategorien von jeweils 5 Jahren, 15 % waren 65 Jahre oder älter) und können überwiegend eine hohe formale Bildung vorweisen (25 % Fach-/Hochschulreife, 64 % Fach-/Hochschulabschluss).

5. Methode

Der Artikel (Stimulustext, Anhang 1) im ersten Experiment („App-Desaster bei der Demokratischen Partei“) legte zwei Interpretationen nahe bezüglich der Frage, weshalb es bei der Übermittlung der Wahlergebnisse massive Verzögerungen gab: Entweder hat (1) die mangelhafte Sorgfalt der App-Entwicklerinnen und -Entwickler zu einer fehlerhaften App geführt oder (2) die Wahlen waren schlecht organisiert und fanden unter ungünstigen Bedingungen statt. Der Artikel im zweiten Experiment („Körperverletzung bei Krippenkind“) legte zwei Interpretationen nahe bezüglich der Frage, wie es zu dem Unfall eines Krippenkindes kam: Entweder waren (1) die Erzieherinnen und Erzieher der Kindergruppe unaufmerksam und fahrlässig oder (2) die Bedingungen, unter denen sie die Kindergruppe betreuen mussten, waren unangemessen und gefährlich.

5.1. Fragebogen

Für die Experimente wurde ein Fragebogen so gestaltet, dass die Thematik Vorurteile/Geschlecht/geschlechtergerechte Sprache möglichst spät angesprochen wurde, um die empfindlichen Priming-Effekte nicht zu beeinflussen. Nach Begrüßung, Einwilligung und Eisbrecherfrage wurde bereits der Stimulus präsentiert, direkt gefolgt von der Schuldfrage. Anschließend erfolgte die Beurteilung der Kompetenz der Fachkräfte, die generalisierte geschlechtsspezifische Kompetenzzuschreibung und die Ausprägung der Stereotype. Erst danach wurde geprüft, ob die im Stimulus verwendete Formulierung korrekt erinnert wurde (*treatment recognition*), zentrale demografische Charakteristika erhoben und zuletzt gefragt, ob der oder die Befragte selbst geschlechtergerechte Formulierungen verwendet. Im Anschluss daran erfolgte die Aufklärung über das Untersuchungsziel.

5.2. Operationalisierung

Die *Schuldzuschreibung* wurde anhand einer visuellen Analogskala mit drei bipolaren Items bestimmt (z.B. „Die Kindergruppe war zu groß für die Anzahl der Fachkräfte“ ↔ „Die Fachkräfte waren fahrlässig bei der Aufsicht“). In den beiden Experimenten wurden unterschiedliche, für den jeweiligen Stimulustext entwickelte Items verwendet (Anhang 2). Die *Kompetenzbeurteilung* der Fachkräfte erfolgte anhand von acht bipolaren Items (Anhang 3).

Die *Ausprägung des Stereotyps* wurde mithilfe von fünfstufigen Skalen erhoben, welche für die Studie frei konstruiert worden waren (Anhang 4) und jeweils positiv und negativ formulierte Items enthielten. Im ersten Experiment *Softwareentwicklung* wurde anhand von vier Items abgefragt, ob Frauen sich mit Technik schwertun. Im zweiten Experiment *Kindererziehung* wurde anhand von sechs Items erfragt, ob sich Männer mit Kindern schwertun. Gegenläufig abgefragt – aber nicht für die Auswertung verwendet – wurde auch, ob Männer gut mit Technik und Frauen gut mit Kindern umgehen können. Durch das vorhergegangene Priming sollten die genannten Stereotype aktiviert werden und handlungsrelevant sein. Jedoch besteht hier das Risiko, dass bereits die Messung der Stereotype dieselben aktivieren könnte.

Zur Anknüpfung an die Literatur wurden die Befragten um eine Bewertung von *Lesefluss* (8 Items) und *Textästhetik* (6 Items) gebeten. Die Items dieser Skalen wurden in Anlehnung an Rothmund und Christmann (2002) formuliert.

Ob die Manipulation des Textes überhaupt wahrgenommen wurde, wurde mit einer geschlossenen Erinnerungsfrage kontrolliert („Erinnern Sie sich noch, welche dieser Formen im Artikel verwendet wurde?“). Zur Auswahl stand die im Text verwendete Bezeichnung in den drei möglichen Formen, zum Beispiel *App-Entwickler*, *App-Entwicklerinnen* und *App-Entwickler* und *App-Entwickler*innen*.

Das Geschlecht der Befragten wurde in den Kategorien *weiblich*, *männlich* und *divers* abgefragt. Ob die Befragten selbst geschlechtergerechte Sprache verwenden, wurde ganz am Ende des Fragebogens 4-stufig mit den Kategorien *nie*, *gelegentlich*, *meistens* und *immer* erfragt.

Außerhalb des Stimulustextes vermied der Fragebogen die Nennung sozialer Gruppen und wo das nicht möglich war, verwendete er geschlechtsneutrale Formulierungen (*Fachkräfte*).

6. Ergebnisse

Die Erinnerungsfrage nach der Textformulierung unterstreicht die Verknüpfung der Berufe mit dem Geschlecht. Während sich 97 % der Befragten bei den „App-Entwicklern“ korrekt an das generische Maskulinum erinnern konnten, erinnern nur 48 % die Paarform und 59 % das Gendersternchen. Bei den „Erziehern“ hingegen wurde das generische Maskulinum von nur 65 % der Befragten korrekt erinnert, aber 73 % erinnerten oder errieten die Paarform korrekt und sogar 84 % das Gendersternchen. Diese Unterschiede kommen mutmaßlich durch jene Befragte zustande, die sich nicht sicher erinnerten: In der von Frauen dominierten Kindererziehung ist eine geschlechtergerechte Formulierung naheliegender als im Bereich der männlich assoziierten Softwareentwicklung.

Generell müssen wir auch davon ausgehen, dass die Stichprobe geschlechtergerechten Formulierungen gegenüber aufgeschlossen ist: 79 % der Befragten gaben an, selbst zumindest gelegentlich (39 %) geschlechtergerechte Formulierungen zu verwenden (32 % *meistens*, 8 % *immer*).

Alle Hypothesen wurden separat für die beiden Experimente mit den Stimulustexten zu App-Entwicklerinnen und Entwicklern (Software, S) und Erzieherinnen und Erziehern (Kindererziehung, K) überprüft. Entsprechend sind die statistischen Kennwerte jeweils für beide Experimente ausgewiesen (r_S , r_K , F_S , F_K , η_S , η_K). Als statistische Verfahren nutzen wir die Produkt-Moment-Korrelation nach Pearson (Hypothesen 1 und 2) und die einfaktorielle Varianzanalyse (ANOVA, Hypothesen 3, 4 und 5) mit den drei experimentellen Bedingungen generisches Maskulinum, Paarform und Gendersternchen als Faktorstufen. Zur besseren Vergleichbarkeit zwischen Korrelationen und Varianzanalysen, und angesichts der im vorliegenden Fall geringen statistischen Diskriminierung (jeweils $\eta^2 < 0.01$), weisen wir die nicht-quadrierte Effektstärke ($\eta = \sqrt{\eta^2}$) aus.

Mit Hypothese 1 wurde zunächst die Voraussetzung bestätigt, dass die antizipierte Kompetenz der Fachkräfte die Zuschreibung vorhersagt, inwieweit die Schuld bei den Fachkräften liegt ($r_S = -.45$, $r_K = -.55$, jeweils $p < .001$, H1). Eine weitere Voraussetzung für den Gesamtzusammenhang zwischen geschlechtergerechter Sprache, Aktivierung von Stereotypen und Textinterpretation (Hypothese 2) ließ sich hingegen nicht belegen: In der Stichprobe ist unabhängig vom Experiment nur ein marginaler Zusammenhang ($r^2 < 0.01$) zwischen der Ausprägung fachspezifischer Stereotype und der Kompetenzzuschreibung zu beobachten ($r_S = -.02$, $p = .25$, $n = 1014$, $r_K = -.10$, $p < .001$, $n = 1053$, H2). Auf Basis dieser geringen Effektstärke ist nicht von einem indirekten, durch die Aktivierung von Stereotypen medierten Zusammenhang zwischen geschlechtergerechten Formulierungen und Kompetenzzuschreibung auszugehen.

Die direkte Wirkung geschlechtergerechter Formulierungen überprüfen die Hypothesen 3 und 4. Eine Aktivierung von Stereotypen durch geschlechtergerechte Formulierungen (H3) ist nicht zu beobachten – zumindest nicht anhand der

explizit abgefragten Zustimmung zum jeweiligen fachspezifischen Stereotyp. Diesbezüglich finden wir keine systematischen Unterschiede zwischen den Experimentalgruppen ($\eta_S = 0.02$, $F_S(2,1011) = 0.21$, $p = .81$, $\eta_K = 0.03$, $F_K(2,1050) = 0.55$, $p = .58$, H3).

Ebenso beobachten wir in den Stichproben keinen systematischen Einfluss der geschlechtergerechten Formulierungen auf die Kompetenzzuschreibung der Fachkräfte ($\eta_S = 0.01$, $F_S(2,1011) = 0.04$, $p = .97$, $\eta_K = 0.02$, $F_K(2,1050) = 0.14$, $p = .87$, H4). Dies gilt im Übrigen auch dann, wenn man Befragte aus der Analyse ausschließt, die sich nicht korrekt an die verwendete Formulierung erinnern konnten (Anhang 6). Können wir also davon ausgehen, dass geschlechtergerechte Sprache keinen Einfluss auf die Kompetenzzuschreibung hat? Ein paarweiser Äquivalenztest (Lakens, 2017; Lakens et al., 2018) zeigt, dass der Effekt in beiden Experimenten zumindest deutlich geringer ist, als dass er eine praktische Bedeutung hätte ($d_{EQ} = [-0.2|+0.2]$, TOST, jeweils $p < .05$).

Angesichts des ausgeprägten Fehlens eines Effekts auf die Kompetenzzuschreibung der Fachkräfte betrachteten wir post-hoc eventuelle Effekte auf die Schuldzuschreibung. Es wäre schließlich denkbar, dass eine geschlechtergerechte Formulierung die Interpretation des Textes durch andere als die hier betrachteten Mechanismen beeinflusst. Doch auch hinsichtlich der Schuldzuschreibung beobachten wir in der Stichprobe keine relevanten Unterschiede zwischen den Experimentalgruppen ($\eta_S = 0.04$, $F_S(2,1011) = 0.90$, $p = .41$, $\eta_K = 0.06$, $F_K(2,1050) = 1.97$, $p = .14$, H5). Ebenso wenig hatte die geschlechtergerechte Formulierung Einfluss auf die Bewertung des Leseflusses ($\eta_S = 0.05$, $F_S(2,1011) = 1.16$, $p = .32$, $\eta_K = 0.05$, $F_K(2,1050) = 1.48$, $p = .23$) oder der Textästhetik ($\eta_S = 0.04$, $F_S(2,1011) = 0.74$, $p = .48$, $\eta_K = 0.03$, $F_K(2,1050) = 0.56$, $p = .57$). Auch beobachteten wir keine Interaktionen zwischen dem Geschlecht der Befragten und den getesteten Zusammenhängen (Anhang 6).

7. Diskussion

Frühere Studien zu den Effekten geschlechtergerechter Sprache konzentrierten sich im Wesentlichen auf erwünschte Effekte im Bereich der inhaltlichen Wahrnehmung (z. B. mentale Repräsentation sozialer Gruppen) und auf unerwünschte Effekte im Bereich der Rezeption (z. B. Störung des Leseflusses). Die vorliegende Studie untersucht anhand zweier Experimente, ob geschlechtergerechte Formulierungen auch unerwünschte Effekte auf die inhaltliche Wahrnehmung haben könnten – schließlich betonen geschlechtergerechte Formulierungen, *dass* es unterschiedliche Geschlechter gibt. Geschlechtergerechte Formulierungen vermitteln möglicherweise, *dass* das Geschlecht ein relevantes Merkmal von Akteuren sei, zum Beispiel relevant für Verhalten und Leistung. Daraus wird die These abgeleitet, dass eine geschlechtergerechte Formulierung eine stereotypgeleitete Interpretation journalistischer Texte verstärkt.

Die Experimente liefern keine Belege, welche die These stützen würden. Nun ist es nicht ungewöhnlich, dass die in einem Experiment beobachteten Unterschiede zu gering ausfallen, als dass sie statistische Signifikanz erreichen. Hinzu kommt, dass Stereotype ohnehin schwer zu beobachten sind: Explizite Abfragen

sind anfällig für Verzerrungen und können die Stereotype alleine durch die Messung aktivieren. Es könnte also dem Einsatz einer ungeeigneten Messung für Stereotype geschuldet sein, dass wir keinen Zusammenhang zwischen der Aktivierung eines Stereotyps und der Kompetenzbeurteilungen beobachten konnten.

Diese Einschränkung kann allerdings nicht die umfassenden Nicht-Effekte erklären, welche wir in der Stichprobe beobachten. Ebenso wenig greifen andere gängige Erklärungsmuster, wie etwa eine unzureichende Stichprobengröße oder eine zurückhaltende Manipulation. Das Experiment wurde anhand von zwei unterschiedlichen beruflichen Kontexten durchgeführt und in Erwartung schwacher Effekte wurde eine große Stichprobe gewählt (in jedem Experiment $N > 1.000$). Die Ergebnisse zeigen deshalb im Äquivalenztest sogar statistisch belastbar, dass geschlechtergerechte Formulierungen entgegen der ursprünglichen Vermutung *keinen* Effekt auf die Textinterpretation haben. Die Befunde sprechen vielmehr dafür, dass die Befragten die geschlechtergerechte Formulierung geflissentlich ignorieren – nicht nur im Hinblick auf die vermuteten negativen Effekte, sondern auch im Hinblick auf die Textbeurteilung. Dies deckt sich mit früheren Befunden von Blake und Klimmt (2010).

Womöglich verantwortet die Wahl der Stichprobe das Ausbleiben einer Reaktion. Befragte aus dem *SoSci Panel* sind überwiegend akademisch geprägt und zeigten schon in früheren Studien (Leiner, 2016b) eine ausgeprägt liberale politische Ausrichtung. Laut Selbstauskunft verwendet die überwiegende Mehrheit der Befragten selbst geschlechtergerechte Formulierungen. Mehrere Studien zeigen, dass die Einstellung gegenüber inklusiver Sprache wesentlich mit der Bildung zusammenhängt (Harrigan & Lucic, 1988; Parks & Robertson, 2008; Prentice, 1994). Zudem fördert ein höherer Bildungsstand die Fähigkeit zur Erkennung sexistischer Sprache (McMinn et al., 1991; McMinn & Foster, 1991). Ob geschlechtergerechte Sprache in Gruppen mit geringerer Bildung Wahrnehmungs- und Primingeffekte auslöst, können wir anhand der vorliegenden Stichprobe nicht beurteilen. Wir überprüften aber, ob der Befund marginaler Effekte auch für einzelne Teilgruppen innerhalb der Stichprobe haltbar ist, von denen man annehmen könnte, dass sie geschlechtergerechter Sprache kritischer gegenüberstehen (Anhang 6). Ein Äquivalenztest ist aufgrund der kleineren Fallzahlen in den Teilgruppen nicht sinnvoll anwendbar – die in den Teilgruppen beobachteten Effektstärken deuten aber auch nicht darauf hin, dass relevante Effekte in der Gesamtstichprobe untergegangen wären. Zwar tauchen in den Teilgruppen vereinzelt Zusammenhänge mit relevanter Effektstärke ($d > 0.1$) auf, aber dabei handelt es sich mutmaßlich um zufällige Ausreißer, denn weder ist eine offensichtliche Systematik erkennbar, noch tauchen die Zusammenhänge im jeweils anderen Experiment auf.

Weiterhin ist es plausibel, dass wir in der Studie langfristige Effekte der Sprache übersehen haben: Die starke Bedeutung des grammatischen Geschlechts in der deutschen Sprache könnte über die Zeit kumulativ dazu führen, dass das Geschlecht als Personenmerkmal mehr Bedeutung erhält oder erhalten hat. Die Beobachtung solch eines generationenübergreifenden und tief in der Sprache verwurzelten Effekts liegt aber weit außerhalb der Möglichkeiten eines Befragungsexperiments.

Die in dieser Studie erhobenen Daten konnten die Befürchtung zerstreuen, dass geschlechtergerechte Formulierungen entgegen ihrer Zielsetzung indirekt Sexismus verstärken könnten. Im Gegenteil beobachteten wir, dass eine akademisch und liberal geprägte Stichprobe nahezu überhaupt nicht darauf reagierte, ob ein journalistischer Text im generischen Maskulinum, in Paarform oder mit Gendersternchen vorgelegt wurde. Dies passt zu den Befunden von Acke (2019), wonach geschlechtergerechte Formulierungen zumindest in Teilen der Gesellschaft zunehmend zur Norm werden. Insgesamt liefern die vorliegenden Daten keine Hinweise, welche gegen die Verwendung der Paarform oder des Gendersternchens in journalistischen Texten sprächen.

Literatur

- Acke, H. (2019). Sprachwandel durch feministische Sprachkritik: Geschlechtergerechter Sprachgebrauch an den Berliner Universitäten [Language change through feminist language criticism: Gender-equitable language use at Berlin universities]. *Zeitschrift für Literaturwissenschaft und Linguistik*, 49, 303–320. <https://doi.org/10.1007/s41244-019-00135-1>
- Alfermann, D. (1992). Frauen in der Attributionsforschung: Die fleißige Liese und der kluge Hans [Women in attribution research: Hardworking Liese and clever Hans]. *Zeitschrift für Personalforschung / German Journal of Research in Human Resource Management*, 6, 301–317.
- Artelt, C., Stanat, P., Schneider, W., & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. [Reading literacy: Test design and results]. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann, & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_4
- Axell, C., & Boström, J. (2021). Technology in children's picture books as an agent for reinforcing or challenging traditional gender stereotypes. *International Journal of Technology and Design Education*, 31(1), 27–39. <https://doi.org/10.1007/s10798-019-09537-1>
- Baden, C., & Lecheler, S. (2012). Fleeting, fading, or far-reaching? A knowledge-based model of the persistence of framing effects: Persistence of framing effects. *Communication Theory*, 22(4), 359–382. <https://doi.org/10.1111/j.1468-2885.2012.01413.x>
- Barth, J. M., & Rice, L. (2016). A tale of two gender roles: The effects of implicit and explicit gender role traditionalism and occupational stereotype on hiring decisions. *Gender Issues*, 34, 86–102. <https://doi.org/10.1007/s12147-016-9175-4>
- Bitkom. (2019, März 6). *IT-Fachkräfte: Nur jeder siebte Bewerber ist weiblich* [IT professionals: Only one in seven applicants is female]. <https://www.bitkom.org/Presse/Presseinformation/IT-Fachkraefte-Nur-jeder-siebte-Bewerber-ist-weiblich>
- Blake, C., & Klimmt, C. (2010). Geschlechtergerechte Formulierungen in Nachrichtentexten [Gender-appropriate wording in news texts]. *Publizistik*, 55(3), 289–304. <https://doi.org/10.1007/s11616-010-0093-2>

- Blake, C., Klimmt, C., & Pompetzki, V. (2008). Geschlechterrepräsentation in Nachrichtentexten: Der Einfluss von geschlechterbezogenen Sprachformen und Fallbeispielen auf den gedanklichen Einbezug von Frauen und die Bewertung der Beitragsqualität. [Gender representation in news texts: The influence of gendered language forms and case studies on the mental inclusion of women and ratings of news story quality.]. *Medien & Kommunikationswissenschaft*, 56(1), 3–20.
- Brewer, P. R., Graf, J., & Willnat, L. (2003). Priming or framing: Media influence on attitudes toward foreign countries. *Gazette (Leiden, Netherlands)*, 65(6), 493–508. <https://doi.org/10.1177/0016549203065006005>
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Cremers, M., Krabel, J., & Calmbach, M. (2015). *Männliche Fachkräfte in Kindertagesstätten. Eine Studie zur Situation von Männern in Kindertagesstätten und in der Ausbildung zum Erzieher* [Male professionals in day care centers. A study on the situation of men in day care centers and in the training to become educators]. Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ). <https://www.bmfsfj.de/bmfsfj/service/publikationen/maennliche-fachkraefte-in-kindertagesstaetten-80936>
- Dascal, M., & Weizman, E. (1987). Contextual exploitation of interpretation clues in text understanding: An integrated model. In J. Verschueren & M. Bertuccelli Papi (Hrsg.), *The pragmatic perspective: Selected papers from the 1985 International Pragmatics Conference* (Bd. 5, S. 31–46). John Benjamins Publishing Company. <https://doi.org/10.1075/pbcs.5.08das>
- Degner, J., Meiser, T., & Rothermund, K. (2009). Kognitive und sozial-kognitive Determinanten: Stereotype und Vorurteile [Cognitive and social-cognitive determinants: Stereotypes and prejudices]. In A. Beelmann & K. J. Jonas (Hrsg.), *Diskriminierung und Toleranz: Psychologische Grundlagen und Anwendungsperspektiven* (S. 75–93). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91621-7_4
- Demarmels, S., & Schaffner, D. (2011). Gendersensitive Sprache in Unternehmenstexten [Gender-sensitive language in corporate texts]. In *Textsorten in der Wirtschaft: Zwischen textlinguistischem Wissen und wirtschaftlichem Handeln* (S. 98–120). VS Verlag für Sozialwissenschaften.
- Diekmann, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin*, 26(10), 1171–1188. <https://doi.org/10.1177/0146167200262001>
- Diewald, G., & Steinhauer, A. (2017). *Richtig gendern. Wie Sie angemessen und verständlich schreiben* [Correct gendering. How to write appropriately and understandably]. Dudenverlag.
- Domke, D., Shah, D. V., & Wackman, D. B. (1998). Media priming effects: Accessibility, association, and activation. *International Journal of Public Opinion Research*, 10(1), 51–74. <https://doi.org/10.1093/ijpor/10.1.51>
- Dunwoody, S. (2020). Science journalism and pandemic uncertainty. *Media and Communication*, 8(2), 471–474. <https://doi.org/10.17645/mac.v8i2.3224>
- Dutke, S. (1998). Zur Konstruktion von Sachverhaltsrepräsentationen beim Verstehen von Texten: 15 Jahre nach Johnson-Lairds Mental Models. [On the construction of factual representations in text comprehension: 15 years after Johnson-Laird's mental models]. *Zeitschrift für Experimentelle Psychologie*, 45(1), 42–59.

- Eckes, T. (1997). *Geschlechterstereotype: Frau und Mann in sozialpsychologischer Sicht* [Gender stereotypes: Woman and man in a social psychological perspective]. Centaurus.
- Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen [Gender stereotypes: Of roles, identities and prejudices]. In R. Becker & B. Kortendiek (Hrsg.), *Handbuch Frauen- und Geschlechterforschung* (2. Aufl.). VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-91972-0>
- Formanowicz, M., Bedynska, S., & Cislak, A. (2013). Side effects of gender-fair language: How feminine job titles influence the evaluation of female applicants. *European Journal of Social Psychology*, 43(1), 62–71. <https://doi.org/10.1002/ejsp.1924>
- Friedman, S. M., Dunwoody, S., & Rogers, C. L. (Hrsg.). (1999). *Communicating uncertainty: Media coverage of new and controversial science*. L. Erlbaum.
- Friedrich, M. C. G., & Heise, E. (2019). Does the use of gender-fair language influence the comprehensibility of texts? An experiment using an authentic contract manipulating single role nouns and pronouns. *Swiss Journal of Psychology*, 78(1–2), 51–60. <https://doi.org/10.1024/1421-0185/a000223>
- Gill, E. A., & Babrow, A. S. (2007). To hope or to know: Coping with uncertainty and ambivalence in women's magazine breast cancer articles. *Journal of Applied Communication Research*, 35(2), 133–155. <https://doi.org/10.1080/00909880701263029>
- Glick, P., Wilk, K., & Perreault, M. (1995). Images of occupations: Components of gender and status in occupational stereotypes. *Sex Roles*, 32, 565–582. <https://doi.org/10.1007/BF01544212>
- Groeben, N. (1982). *Leserpsychologie: Textverständnis Textverständlichkeit* [Reader psychology: Text comprehension text comprehensibility]. Aschendorff.
- Harrigan, J. A., & Lucic, K. S. (1988). Attitudes about gender bias in language: A reevaluation. *Sex Roles*, 19, 129–140. <https://doi.org/10.1007/BF00290150>
- Hellinger, M. (1990). *Kontrastive Feministische Linguistik. Mechanismen sprachlicher Diskriminierung im Englischen und Deutschen* [Contrastive feminist linguistics. Mechanisms of linguistic discrimination in English and German]. Hueber.
- Herrmann, S. K. (2003). Performing the Gap. Queere Gestalten und geschlechtliche Aneignung [Performing the gap. Queer gestalten and gendered appropriation]. *arranca!*, 28. <https://arranca.org/archive?path=%2Fausgabe%2F28%2Fperforming-the-gap>
- Horvath, L. K., Merkel, E. F., Maass, A., & Sczesny, S. (2016). Does gender-fair language pay off? The social perception of professions from a cross-linguistic perspective. *Frontiers in Psychology*, 6(2018). <https://doi.org/10.3389/fpsyg.2015.02018>
- Howard, J. A., & Allen, C. (1990). The gendered context of reading. *Gender & Society*, 4(4), 534–552. <https://doi.org/10.1177/089124390004004006>
- Irmen, L., & Linner, U. (2005). Die Repräsentation generisch maskuliner Personenbezeichnungen. Eine theoretische Integration bisheriger Befunde [The representation of generically masculine personal names. A theoretical integration of previous findings]. *Zeitschrift für Psychologie*, 213(3), 167–175. <https://doi.org/10.1026/0044-3409.213.3.167>
- Iser, W. (1980). Interaction between text and reader. In S. R. Suleiman & I. Crosman (Hrsg.), *The reader in the text: Essays on audience and interpretation* (S. 106–119). Princeton University Press. <https://doi.org/10.1515/9781400857111>
- Kintsch, W. (1974). *The representation of meaning in memory*. Lawrence Erlbaum. <https://doi.org/10.2307/1421786>

- Klann-Delius, G. (2005). *Sprache und Geschlecht: Eine Einführung* [Language and gender: An introduction]. Metzler.
- kleinergast. (2016, September 26). „Alles Handeln ist politisch“ – ein Plädoyer für das Gendern mit Doppelpunkt und ein Gespräch mit Lann Hornscheidt | Kleinerdreieck [“All action is political”—A plea for gendering with colon and a conversation with Lann Hornscheidt | Kleinerdreieck] [Gastbeitrag]. *kleinerdreieck*. <http://kleinerdreieck.org/2016/09/alles-handeln-ist-politisch-ein-plaedoyer-fuer-das-gendern-mit-doppelpunkt-und-ein-gespraech-mit-lann-hornscheidt/>
- Kolodny, A. (1980). A map for rereading: Or, gender and the interpretation of literary texts. *New Literary History*, 11(3), 451. <https://doi.org/10.2307/468938>
- Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, 103(2), 284–308. <https://doi.org/10.1037/0033-295X.103.2.284>
- Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269. <https://doi.org/10.1177/2515245918770963>
- Lawrimore, K. W., & Marcis, J. G. (1994). Current occupational stereotypes. *Atlantic Economic Journal*, 22(3), 87. <https://doi.org/10.1007/BF02301808>
- Leiner, D. J. (2016a). Our research’s breadth lives on convenience samples: A case study of the online respondent pool “SoSci Panel”. *Studies in Communication | Media*, 5(4), 367–396. <https://doi.org/10.5771/2192-4007-2016-4-367>
- Leiner, D. J. (2016b). *Stabilität öffentlicher Meinung* [Public opinion stability]. Springer VS. <https://doi.org/10.1007/978-3-658-12285-0>
- Leiner, D. J. (2019). Too fast, too straight, too weird: Non-reactive indicators for meaningless data in internet surveys. *Survey Research Methods*, 229–248. <https://doi.org/10.18148/SRM/2019.V13I3.7403>
- Lindqvist, A., Renström, E. A., & Gustafsson Sendén, M. (2019). Reducing a male bias in language? Establishing the efficiency of three different gender-fair language strategies. *Sex Roles*, 81(1–2), 109–117. <https://doi.org/10.1007/s11199-018-0974-9>
- Matthes, J. (2007). Beyond accessibility? Toward an on-line and memory-based model of framing effects. *Communications*, 32(1). <https://doi.org/10.1515/COMMUN.2007.003>
- McMinn, M. R., & Foster, J. D. (1991). A computer program to teach nonsexist language. *Teaching of Psychology*, 18(2), 115–117. https://doi.org/10.1207/s15328023top1802_16
- McMinn, M. R., Troyer, P. K., Hannum, L. E., & Foster, J. D. (1991). Teaching nonsexist language to college students. *Journal of Experimental Education*, 59(2), 153–161. <https://doi.org/10.1080/00220973.1991.10806558>
- Nier, H. (2018, März 20). *Wie weiblich ist die IT?* [How female is IT?]. Statista. <https://de.statista.com/infografik/13283/frauen-in-der-tech-branche/>
- Parks, J. B., & Robertson, M. A. (1998). Influence of age, gender and context on attitudes toward sexist/nonsexist language: Is sport a special case? *Sex Roles*, 38(5–6), 477–493. <https://doi.org/10.1023/A:1018766023667>

- Parks, J. B., & Robertson, M. A. (2008). Generation gaps in attitudes toward sexist/nonsexist language. *Journal of Language and Social Psychology, 27*(3), 276–283. <https://doi.org/10.1177/0261927X08317956>
- Pöschko, H., & Prieler, V. (2018). Zur Verständlichkeit und Lesbarkeit von geschlechtergerecht formulierten Schulbuchtexten [On the comprehensibility and readability of gender-appropriate textbook texts]. *Zeitschrift für Bildungsforschung, 8*, 5–18. <https://doi.org/10.1007/s35834-017-0195-2>
- Prentice, D. A. (1994). Do language reforms change our way of thinking? *Journal of Language and Social Psychology, 13*(1), 3–19. <https://doi.org/10.1177/0261927X94131001>
- Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly, 26*(4), 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>
- Pusch, L. F. (1979). Der Mensch ist ein Gewohnheitstier, doch weiter kommt man ohne ihr. Eine Antwort auf Kalverkämpfers Kritik an Trömel-Plötz' Artikel über Linguistik und Frauensprache [Man is a creature of habit, but one gets further without her. A response to Kalverkämpfer's criticism of Trömel-Plötz's article on linguistics and women's language]. *Linguistische Berichte, 63*, 84–102.
- Rat für deutsche Rechtschreibung. (2018). *Empfehlungen zur „geschlechtergerechten Schreibung“*. Beschluss des Rats für deutsche Rechtschreibung vom 16. November 2018 [Recommendations on „gender-appropriate spelling“. Decision of the Council for German Orthography of November 16, 2018]. https://www.rechtschreibrat.com/DOX/rfdr_PM_2018-11-16_Geschlechtergerechte_Schreibung.pdf
- Roskos-Ewoldsen, D. R., Roskos-Ewoldsen, B., & Dillman Carpentier, F. (2008). Media priming. An updated synthesis. In J. Bryant & M. B. Oliver (Hrsg.), *Media Effects Advances in Theory and Research* (S. 74–93). Routledge.
- Rothmund, J., & Christmann, U. (2002). Auf der Suche nach einem geschlechtergerechten Sprachgebrauch: Führt die Ersetzung des generischen Maskulinums zu einer Beeinträchtigung von Textqualitäten? [In search of gender-equitable language use: Does the substitution of the generic masculine lead to an impairment of textual qualities?]. *Muttersprache, 112*(2), 115–136.
- Rothmund, J., & Scheele, B. (2004). Personenbezeichnungsmodelle auf dem Prüfstand [Person designation models put to the test]. *Journal of Psychology, 212*, 40–54. <https://doi.org/10.1026/0044-3409.212.1.40>
- Rubin, D. (1993). *Gender influences: Reading student texts*. Southern Illinois University Press.
- Sáinz, M., Meneses, J., López, B.-S., & Fàbregues, S. (2016). Gender stereotypes and attitudes towards information and communication technology professionals in a sample of Spanish secondary students. *Sex Roles, 74*, 154–168. <https://doi.org/10.1007/s11199-014-0424-2>
- Scheufele, B. (2004). Framing-Effekte auf dem Prüfstand. Eine theoretische, methodische und empirische Auseinandersetzung mit der Wirkungsperspektive des Framing-Ansatzes [Framing effects under scrutiny. A theoretical, methodological, and empirical examination of the effects perspective of the framing approach]. *Medien & Kommunikationswissenschaft, 52*(1), 30–55. <https://doi.org/10.5771/1615-634x-2004-1-30>
- Scheufele, B. (2016). *Priming*. Nomos. <https://doi.org/10.5771/9783845263120>

- Schmid Mast, M., & Krings, F. (2008). Stereotype und Informationsverarbeitung [Stereotypes and information processing]. In L.-E. Petersen & B. Six (Hrsg.), *Stereotype, Vorurteile und soziale Diskriminierung. Theorien, Befunde und Interventionen*. Beltz. <https://archive-ouverte.unige.ch/unige:102101>
- Schmitz, A. (2016). *Verständlichkeit von Sachtexten: Wirkung der globalen Textkohäsion auf das Textverständnis von Schülern* [Comprehensibility of nonfiction texts: The effect of global text cohesion on students' text comprehension]. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-658-12016-0_3
- Sczesny, S., Spreemann, S., & Stahlberg, D. (2006). Masculine = competent? Physical appearance and sex as sources of gender-stereotypic attributions. *Swiss Journal of Psychology*, 65(1), 15–23. <https://doi.org/10.1024/1421-0185.65.1.15>
- Sczesny, S., & Stahlberg, D. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen [Effects of the generic masculine and alternative forms of language on the mental inclusion of women]. *Psychologische Rundschau*, 52(3), 131–140. <https://doi.org/10.1026/0033-3042.52.3.131>
- Sczesny, S., & Stahlberg, D. (2002). The influence of gender-stereotyped perfumes on leadership attribution. *European Journal of Social Psychology*, 32, 815–828. <https://doi.org/10.1002/ejsp.123>
- Sczesny, S., & Stahlberg, D. (2005). Cognitive effects of masculine generics in German: An overview of empirical findings. *Communications*, 30(1), 1–21. <https://doi.org/10.1515/comm.2005.30.1.1>
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613–629. <https://doi.org/10.1037//0003-066x.52.6.613>
- Steffens, M. C., & Ebert, I. D. (2016). *Frauen – Männer – Karrieren: Eine sozialpsychologische Perspektive auf Frauen in männlich geprägten Arbeitskontexten* [Women – men – careers: A social psychological perspective on women in male-dominated work contexts.]. Springer. <https://doi.org/10.1007/978-3-658-10750-5>
- Steiger, V., & Irmen, L. (2011). Recht verständlich und „gender-fair“: Wie sollen Personen in amtlichen Texten bezeichnet werden? Ein Vergleich verschiedener Rezipientengruppen zur Akzeptanz geschlechtergerechter Rechtssprache [Legally understandable and „gender-fair“: How should persons be referred to in official texts? A comparison of different recipient groups on the acceptance of gender-fair legal language]. *Linguistische Berichte*, 227, 297–322.
- Troche, S. J., & Rammsayer, T. H. (2011). Eine Revision des deutschsprachigen Bem Sex-Role Inventory [A revision of the German-Language Bem Sex-Role Inventory]. *Klinische Diagnostik und Evaluation*, 4, 262–283.
- Trutkowski, E. (2018). Wie generisch ist das generische Maskulinum? Über Genus und Sexus im Deutschen [How generic is the generic masculine? About genus and sexus in German]. *ZAS Papers in Linguistics*, 59, 83–96. <https://doi.org/10.21248/zaspil.59.2018.437>
- Vergoossen, H. P., Renström, E. A., Lindqvist, A., & Gustafsson Sendén, M. (2020). Four dimensions of criticism against gender-fair language. *Sex Roles*, 83(4–5), 328–337. <https://doi.org/10.1007/s11199-019-01108-x>
- Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, 142(5), 472–497. <https://doi.org/10.1037/bul0000030>
- White, M. J., & White, G. B. (2006). Implicit and explicit occupational gender stereotypes. *Sex Roles*, 55(3–4), 259–266. <https://doi.org/10.1007/s11199-006-9078-z>

Anhang 1: Stimulustexte

Die Stimulustexte sind nachfolgend in der Variante mit Gendersternenchen dargestellt. Alle Verweise auf soziale Gruppen, wie z.B. //App-Entwickler*innen//, //Expert*innen// oder //Erzieher*innen//, unterlagen der experimentellen Manipulation.

Stimulustext „Softwareentwicklung“ in der Variante Gendersternenchen

5. Februar 2020, 14:00 Uhr

App-Desaster bei der Demokratischen Partei

Eigentlich sollte bei den Vorwahlen der Demokratischen Partei in Iowa die neue App mit dem Namen IowaReporterApp die Ergebnisse übermitteln. Doch am Montag mussten die Ergebnisse stattdessen auf völlig überlasteten Telefonleitungen durchgegeben werden oder sogar persönlich in die Parteibüros gebracht werden. Dadurch hat sich die Auszählung erheblich verzögert. Wie konnte es dazu kommen?

Der New York Times zufolge bastelten die App-Entwickler*innen der Firma Shadow noch zwei Tage vor der Wahl an der Software herum. Wahlkreisleiter Jonathan Green räumte gegenüber der Webseite Vice Motherboard ein, dass er erst zwei Wochen vor der Wahl zum ersten Mal Gelegenheit bekommen habe, die App zu testen. Erst am Tag der Wahl habe die endgültige Gebrauchsanweisung vorgelegen. Eine durchdachte Teststrategie war damit unmöglich. Auch konnten die Programmierer*innen nicht berücksichtigen, dass sich viele Wahllokale häufig in Schulturnhallen oder Kellern von Kirchen mit schlechtem Mobilfunkempfang befinden. Für die Übermittlung der Wahlergebnisse per App stellt dies ein Problem dar.

Die App-Entwickler*innen haben Fehler ihrer Firma mit „katastrophalen Auswirkungen“ zugegeben. Sie erklärten zwar, dass ihr Unternehmen zuvor noch nie mit einer Wahl zu tun gehabt habe, wiesen jedoch Vorwürfe bezüglich der Sorgfalt bei der App-Entwicklung entschieden zurück. Tatsächlich hielt die Firma Shadow die App geheim, sodass unabhängige Expert*innen den Programmcode nicht auf seine Sicherheit überprüfen konnten. Man verließ sich auf die Kompetenz der eigenen Programmierer*innen.

IT-Expert*innen warnen seit Jahren immer wieder davor, dass die Digitalisierung von Wahlen mehr Probleme schaffe als löse. Hackerangriffe und technische Fehler wie in Iowa könnten das Vertrauen der Bevölkerung in den Wahlprozess untergraben. Dennoch schreitet die Digitalisierung der Abstimmungen weiter voran, ohne zentral beaufsichtigt oder geprüft zu werden.

Stimulustext „Kindererziehung“ in der Variante Gendersternenchen

4. Februar 2020, 17:53 Uhr

Körperverletzung bei Krippenkind

Drei Erzieher*innen einer Krippe wurden wegen fahrlässiger Körperverletzung jeweils verwarnt und zu Geldstrafen von 1500 Euro auf Bewährung verurteilt.

Ein 20 Monate alter Junge stürzte im April 2019 in ein Regenrückhaltebecken. Dort saß er mehrere Minuten im kalten Wasser fest. Die Erzieher*innen der fünf-

zehnköpfigen Krippengruppe befanden sich etwa 30 Meter entfernt, als sie sein Fehlen bemerkten und ihn retteten. Der Junge wurde in eine Klinik gebracht, in der man eine Unterkühlung feststellte. Am nächsten Tag konnte er die Klinik wieder verlassen.

Mit persönlichen Worten hat sich Richter Henning Lüdtkke in seiner Urteilsbegründung an die Eltern des verletzten Krippenkindes und die angeklagten Betreuer*innen im Wolfsburger Gerichtssaal gewandt. Aus Sicht des Gerichts bemerkten die Beschuldigten beim Ausflug zu einem nahegelegenen Spielplatz nicht rechtzeitig, dass das Kind weggelaufen war.

Das Urteil ist vergleichsweise milde. Das begründete der Richter damit, dass die individuelle Schuld trotz der Folgen eher gering zu bewerten sei. Lüdtkke betonte jedoch eine gesteigerte Aufsichtspflicht mit Blick auf das Alter des Jungen.

Die Angeklagten drückten ihr Bedauern aus. Das Geschehene tue ihnen sehr leid, jedoch sei es nicht einfach, zu jedem Zeitpunkt alle Kinder im Blick zu behalten. Jede*r Erzieher*in habe bei ihnen fünf Kleinkinder zu beaufsichtigen, was in bestimmten Situationen wie z. B. im Straßenverkehr überfordernd sei.

Die Eltern forderten ein Berufsverbot für die angeklagten Betreuer*innen. Ihr Sohn sei durch den Vorfall sehr verängstigt worden. Wäre das Wasser nur etwas tiefer gewesen, hätte der Unfall tödlich verlaufen können, betonten sie. Für eine*n Erzieher*in sei es die wichtigste Pflicht, ihrer*seiner Aufsichtspflicht ordentlich nachzukommen. Richter Lüdtkke wies diese Forderung mit dem Hinweis zurück, dass die Krippe aufgrund der geringen Personalstärke eine Mitschuld trage.

Die Bertelsmann Stiftung empfiehlt für Krippengruppen ein Maximum von drei Kindern pro Fachkraft. Die Realität zeigt jedoch, dass in Deutschland der Durchschnitt bei ca. 4,2 Kleinkindern pro Erzieher*in liegt – in Sachsen sind es sogar 6,2 Kinder.

Die Verteidiger*innen hatten auf Freispruch plädiert und können noch Rechtsmittel gegen das Urteil einlegen.

Anhang 2: Schuldzuschreibung

Erhebung der Schuldzuschreibung im Experiment „Softwareentwicklung“

Bitte lesen Sie die Aussagen unten und setzen Sie den Schieberegler an die Stelle, die Ihre persönliche Einschätzung am besten beschreibt.

Pol mit Wert=0	Pol mit Wert=1	M	SD
Die Fähigkeiten der Fachkräfte bei der App-Entwicklung führten zum Scheitern der App	Die Organisation der Wahlen führte zum Scheitern der App	0.57	0.28
Die Fachkräfte haben im Entwicklungsprozess fahrlässig gehandelt	Der Vorfall hätte nicht durch die Fachkräfte des App-Entwicklungsunternehmens verhindert werden können	0.63	0.26
Die Fachkräfte des App-Entwicklungsunternehmens hätten vorausschauender denken müssen	Im Vornherein waren die aufgetretenen Probleme nicht absehbar	0.77	0.22

Erhebung der Schuldzuschreibung im Experiment „Kindererziehung“

Bitte lesen Sie die Aussagen unten und setzen Sie den Schieberegler an die Stelle, die Ihre persönliche Einschätzung am besten beschreibt.

Pol mit Wert=0	Pol mit Wert=1	M	SD
Die Fachkräfte waren fahrlässig bei der Aufsicht	Die Kindergruppe war zu groß für die Anzahl der Fachkräfte	0.35	0.23
Die Fachkräfte haben sich falsch verhalten	Die mangelhaften Sicherheitsvorkehrungen machten es den Fachkräften unmöglich, den Unfall zu verhindern	0.49	0.22
Die Fachkräfte waren den Anforderungen ihres Berufs nicht gewachsen	Die Bedingungen bei der Betreuung der Krippenkinder waren nicht zumutbar	0.35	0.19

Anhang 3: Kompetenzbeurteilung

Wie würden Sie die zuständigen Fachkräfte allein auf Grundlage des Textes einschätzen?

[Experiment Softwareentwicklung] Hierbei geht es um die für die App-Entwicklung zuständigen Fachkräfte, die im Text thematisiert wurden.

[Experiment Kindererziehung] Hierbei geht es um die pädagogischen Fachkräfte, die im Text thematisiert wurden.

Pol mit Wert=1	Pol mit Wert=6	Softwareentw.		Kindererzieh.	
		M	SD	M	SD
schlecht ausgebildet	gut ausgebildet	4.0	1.1	4.2	1.0
unzuverlässig	zuverlässig	3.5	1.2	3.8	1.1
inkompetent	kompetent	3.5	1.2	4.0	1.0
unaufrichtig	aufrichtig	3.9	1.3	4.8	1.0
verantwortungslos	verantwortungsvoll	3.2	1.2	4.1	1.0
nicht vertrauenswürdig	vertrauenswürdig	3.3	1.	4.1	1.1
gewissenlos	gewissenhaft	3.6	1.0	4.2	0.9
unprofessionell	professionell	2.8	1.3	3.9	1.2

Anhang 4: Zustimmung zu Stereotypen

4.1 Stereotyp „Frauen tun sich schwer mit Technik“

Natürlich kann man das nicht so pauschal sagen, aber wenn Sie ein spontanes Urteil zu den folgenden Aussagen abgeben müssten, wie sehr würden Sie ihnen zustimmen? [Zustimmung 5-stufig mit $W = [1;5]$]

Tabelle 4.1. Items zur Erhebung des Stereotyps Frauen/Technik

	<i>Softwareentw.</i>		<i>Kindererzieh.</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Viele Frauen ...				
... können gut mit Technik umgehen (umgepolt)	2.9	0.93	2.9	0.92
... können technische Geräte zwar bedienen, verstehen aber ihre Funktionsweise nicht	3.1	1.02	3.1	1.01
... sind mathematisch begabt (umgepolt)	2.7	0.81	2.7	0.82
... können logische Aufgaben gut lösen (umgepolt)	2.5	0.84	2.5	0.83

Die Mittelwerte sind in beiden Experimenten für alle 4 Items auf 2 gültige Stellen gleich. Es handelt sich hierbei nicht um einen Darstellungsfehler.

Die Zustimmung zum Stereotyp „Frauen tun sich schwer mit Technik“ wurde in beiden Experimenten erhoben, floss aber nur in die Auswertung des Experiments mit dem Stimulustext „Softwareentwicklung“ ein.

4.2. Stereotyp „Männern tun sich schwer mit Erziehung“

Natürlich kann man das nicht so pauschal sagen, aber wenn Sie ein spontanes Urteil zu den folgenden Aussagen abgeben müssten, wie sehr würden Sie ihnen zustimmen? [Zustimmung 5-stufig mit $\mathbb{W} = [1;5]$]

Tabelle 4.2. Items zur Erhebung des Stereotyps Männer/Kinder

	<i>Softwareentw.</i>		<i>Kindererzieh.</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Viele Männer ...				
... empfinden es als lästig, auf Kinder aufzupassen	3.1	0.84	3.2	0.88
... gehen verantwortungsvoll mit Kindern um (umgepolt)	2.3	0.64	2.3	0.64
... sind fürsorglich (umgepolt)	2.5	0.68	2.4	0.66
... konzentrieren sich lieber auf andere Dinge (z.B. Hobbys) als Kinder	3.1	0.85	2.1	0.86
... wissen gut, wie sie auf ein Kind richtig aufpassen (umgepolt)	2.8	0.74	2.7	0.74
... können gut mit Kindern umgehen (umgepolt)	2.5	0.68	2.5	0.65

Die Zustimmung zum Stereotyp „Männern tun sich schwer mit Erziehung“ wurde in beiden Experimenten erhoben, floss aber nur in die Auswertung des Experiments mit dem Stimulustext „Kindererziehung“ ein.

Anhang 5: Deskriptive Informationen

Die folgenden Tabellen weisen die Kennwerte der eingesetzten Skalenindizes getrennt nach Experiment aus. Neben der Bezeichnung des Konstrukts ist in eckigen Klammern jeweils der mögliche Wertebereich der Operationalisierung (Mittelwertindex) angegeben. Folgende Abkürzungen werden verwendet:

<i>MW</i>	Arithmetisches Mittel
<i>CI</i> _{95%}	Konfidenzintervall mit einer Wahrscheinlichkeit von 95 %
<i>SD</i>	Standardabweichung
α	Skalenkonsistenz, Cronbach's Alpha (standardisiert)
<i>N</i>	Anzahl gültiger Fälle für den Skalenindex

Sofern in den Antworten einer befragten Person einzelne Items nicht beantwortet wurden (item non-response < 0.006), ersetzen wir fehlende Werte zur Bildung des Mittelwertindex durch Schätzungen auf Basis einer linearen Regression. Der resultierende Skalenmittelwert korreliert in der vorliegenden Stichprobe mit einem Skalenmittelwert, welcher fehlende Werte entfernt, mit $r > 0.998$.

Weiterhin sind Mittelwerte (*MW*) und Standardabweichungen (*SD*) für alle Experimentalbedingungen ausgewiesen.

Ebenfalls ausgewiesen sind die Korrelationen (Pearson) zwischen den Indizes sowie das partielle Eta (η) zwischen Experimentalbedingung und Index. Für den Zusammenhang zwischen Altersgruppe bzw. der eigenen Verwendung geschlechtergerechter Sprache (*ge.ger. Sprache*) und Indizes ist die Rangkorrelation (Spearman) ausgewiesen. Die Zusammenhänge sind für die Personen, die sich einem Geschlecht zugeordnet haben, zusätzlich auch noch einmal separat nach Teilnehmerinnen (♀) und Teilnehmern (♂) ausgewiesen.

5.1 Experiment mit Stimulustext „Softwareentwicklung“

Tabelle 5.1.1. Mittelwerte und Skalenkennwerte nach Geschlecht

Konstrukt	<i>MW</i>	<i>CI</i> _{95%}	<i>SD</i>	α	<i>N</i>
Schuldzuschreibung [0; 1]	0.656	[0.64; 0.67]	0.19	.64	1014
♀	0.649	[0.63; 0.67]	0.20	.69	480
♂	0.663	[0.65; 0.68]	0.19	.60	530
Kompetenzbeurteilung [1; 6]	3.50	[3.44; 3.55]	0.89	.89	1014
♀	3.51	[3.43; 3.59]	0.88	.89	480
♂	3.49	[3.41; 3.57]	0.90	.89	530
Zustimmung Stereotyp ¹ [1; 5]	2.78	[2.73; 2.82]	0.68	.76	1014
♀	2.67	[2.60; 2.73]	0.68	.76	480
♂	2.88	[2.82; 2.93]	0.67	.76	530
Lesefluss [1; 5]	3.89	[3.84; 3.93]	0.74	.88	1014
♀	3.88	[3.81; 3.95]	0.78	.90	480
♂	3.90	[3.84; 3.96]	0.70	.86	530
Textästhetik [1; 5]	3.11	[3.06; 3.15]	0.73	.83	1014
♀	3.03	[2.97; 3.10]	0.84	.84	480
♂	3.18	[3.12; 3.24]	0.81	.81	530

*Anmerkung.*¹Zustimmung zu dem Stereotyp „Frauen können schlecht mit Technik umgehen“ (gemessen durch 4 Items, Anhang 4)

Tabelle 5.1.2. Mittelwerte nach Experimentalbedingung

Experimentelle Bedingung →	Generisches Maskulinum (<i>n</i> = 329)	Paarform (<i>n</i> = 329)	Gendersternchen (<i>n</i> = 356)
Schuld-zuschreibung	0.647 (0.19) [0.63; 0.67]	0.667 (0.20) [0.64; 0.69]	0.654 (0.19) [0.63; 0.67]
Kompetenz-beurteilung	3.49 (0.91) [3.39; 3.59]	3.51 (0.85) [3.41; 3.60]	3.50 (0.91) [3.41; 3.60]
Zustimmung Stereotyp1	2.77 (0.68) [2.69; 2.84]	2.80 (0.69) [2.72; 2.87]	2.77 (0.68) [2.70; 2.84]
Leseffluss	3.93 (0.75) [3.85; 4.01]	3.84 (0.74) [3.76; 3.92]	3.90 (0.74) [3.82; 3.97]
Testästhetik	3.11 (0.75) [3.03; 3.19]	3.07 (0.72) [3.00; 3.15]	3.14 (0.71) [3.07; 3.22]

Anmerkung. Angegeben sind Mittelwerte, Standardabweichungen (in runden Klammern) und das Konfidenzintervall des Mittelwerts (95 %) in eckigen Klammern.

5.2 Experiment mit Stimulustext „Kindererziehung“

Tabelle 5.2.1. Mittelwerte und Skalenskennwerte nach Geschlecht

Konstrukt	MW	CI _{95%}	SD	α	N
Schuldzuschreibung [0; 1]	0.398	[0.39; 0.41]	0.18	.75	1053
♀	0.389	[0.37; 0.41]	0.18	.77	474
♂	0.404	[0.39; 0.42]	0.17	.73	575
Kompetenzbeurteilung [1; 6]	4.16	[4.10; 4.21]	0.84	.93	1053
♀	4.14	[4.06; 4.21]	0.83	.93	474
♂	4.17	[4.10; 4.24]	0.86	.93	575
Zustimmung Stereotyp ² [1; 5]	2.69	[2.66; 2.72]	0.50	.77	1053
♀	2.64	[2.60; 2.69]	0.51	.79	474
♂	2.73	[2.69; 2.77]	0.50	.76	575
Leseffluss [1; 5]	4.17	[4.13; 4.21]	0.68	.87	1053
♀	4.16	[4.09; 4.22]	0.70	.88	474
♂	4.17	[4.12; 4.23]	0.66	.87	575
Textästhetik [1; 5]	3.35	[3.31; 3.39]	0.67	.78	1053
♀	3.31	[3.25; 3.37]	0.69	.80	474
♂	3.38	[3.32; 3.43]	0.66	.77	575

Anmerkung. ² Zustimmung zu dem Stereotyp „Männern tun sich schwer mit Erziehung“ (gemessen mit 6 Items, Anhang 4).

Tabelle 5.2.2. Mittelwerte nach Experimentalbedingung

Experimentelle Bedingung →	Generisches Maskulinum (<i>n</i> = 352)	Paarform (<i>n</i> = 359)	Gendersternchen (<i>n</i> = 343)
Schuld-zuschreibung	0.404 (0.17) [0.39; 0.42]	0.383 (0.18) [0.36; 0.40]	0.408 (0.18) [0.39; 0.43]
Kompetenz-beurteilung	4.17 (0.78) [4.09; 4.25]	4.16 (0.85) [4.07; 4.25]	4.14 (0.90) [4.04; 4.23]
Zustimmung Stereotyp ²	2.69 (0.48) [2.64; 2.74]	2.68 (0.52) [2.62; 2.73]	2.71 (0.51) [2.66; 2.77]
Lesefluss	4.17 (0.65) [4.10; 4.24]	4.21 (0.70) [4.14; 4.28]	4.12 (0.69) [4.05; 4.19]
Testästhetik	3.37 (0.68) [3.30; 3.44]	3.35 (0.66) [3.28; 3.42]	3.32 (0.68) [3.24; 3.39]

Anmerkung. Angegeben sind Mittelwerte, Standardabweichungen (in runden Klammern) und das Konfidenzintervall des Mittelwerts (95 %) in eckigen Klammern.

5.3 Experimentalgruppen

Tabelle 5.3. Fallzahlen (bereinigt) nach Experimentalgruppe und Geschlecht

Experiment	Stimulus	N	♀	♂	divers	k.A.
Software	Generisches Maskulinum	329	159	170	0	0
	Paarform	329	146	182	1	0
	Gendersternchen	356	175	178	2	1
	Gesamt (Software)	1014	480	530	3	1
Kindererziehung	Generisches Maskulinum	351	172	179	0	0
	Paarform	359	158	198	3	0
	Gendersternchen	343	144	198	1	0
	Gesamt (Kindererziehung)	1053	474	575	4	0

Anhang 6: Zusammenhänge

Weiterhin wurden die Zusammenhänge der Hypothesen 1 bis 5 und die Zusammenhänge zwischen Experimentalgruppe und beobachteter Lesedauer sowie empfundenem Lesefluss und Textästhetik für einzelne Teilgruppen separat berechnet. Dies dient der Prüfung, ob die Annahme marginaler Effekte auch für Teilgruppen korrekt ist, welche geschlechtergerechter Sprache mutmaßlich kritisch gegenüberstehen. Als Teilgruppen wurden betrachtet: Personen ab einem Alter von 60 Jahren (60+), Männer (*Männer*), Personen, die laut eigener Aussage selbst „nie“ geschlechtergerechte Formulierungen verwenden (*selbst kgS*) und Personen, welche sich im Manipulation Check korrekt an die im Text verwendete Formulierung erinnern konnten (*Mn.Check*).

Die Effektstärken weisen wir als Korrelationskoeffizienten r (mit Vorzeichen) bzw. als nicht-quadrirtes Eta ($\eta = \sqrt{\eta^2}$) aus.

6.1. Experiment mit Stimulustext „Softwareentwicklung“

Zur Überprüfung einer möglichen Interaktion der Zusammenhänge mit dem Geschlecht wurden die Zusammenhänge zwischen den zentralen Konstrukten getrennt nach dem Geschlecht der Befragten berechnet.

Tabelle 6.1.1. Zusammenhänge insgesamt und nach Geschlecht

	Experimentelle Bedingung (η)	Schuld-zuschreibung	Kompetenz-beurteilung	Zustimmung zu Stereotypen
Experimentelle Bedingung (η)	–	.04° ♀ .04 .10 ♂	.01° ♀ .02 .01 ♂	.02° ♀ .06 .03 ♂
Schuld-zuschreibung	.04° ♀ .04 .10 ♂	–	-.45 ♀ -.44 -.46 ♂	+.01 ♀ +.13 -.11 ♂
Kompetenzbeurteilung	.01° ♀ .02 .01 ♂	-.45 ♀ -.44 -.46 ♂	–	-.02 ♀ -.02 -.01 ♂
Zustimmung Stereotyp	.02° ♀ .06 .03 ♂	+.01 ♀ +.13 +.11 ♂	-.02 ♀ -.02 -.01 ♂	–
Lesefluss	.05° ♀ .02 .10 ♂	+.12 ♀ +.18 +.05 ♂	+.03 ♀ +.03 +.03 ♂	-.05 ♀ -.05 -.05 ♂
Testästhetik	.04° ♀ .03 .09 ♂	+.08 ♀ +.13 +.01 ♂	+.10 ♀ +.13 +.07 ♂	-.01 ♀ +.02 -.07 ♂
Altersgruppe	.02° ♀ .03 .03 ♂	+.02 ♀ +.01 +.02 ♂	-.03 ♀ +.07 -.11 ♂	+.06 ♀ -.02 +.08 ♂
Verwendung ge. ger. Sprache	.03° ♀ .07 .02 ♂	.00 ♀ .00 .00 ♂	-.01 ♀ -.04 .00 ♂	-.17 ♀ -.07 -.22 ♂

Anmerkung. °Eine zweifaktorielle Varianzanalyse zeigte *keinen* signifikanten Interaktionseffekt zwischen experimenteller Bedingung und Geschlecht der Befragten ($p \geq 0.05$). Dies gilt für alle aufgeführten abhängigen Variablen. Nur ausgewiesen für Zusammenhänge mit der experimentellen Bedingung.

Die Altersgruppe wurde zwischen 15 und 64 Jahren in 10 Kategorien zu jeweils 5 Jahren abgefragt sowie durch eine elfte Kategorie „65 Jahre oder älter“.

Für die Zusammenhänge zwischen der experimentelle Bedingung und den anderen Konstrukten ist das Eta (η) auf Basis einer einfaktoriiellen Varianzanalyse ausgewiesen (die Altersgruppe wurde hierfür als näherungsweise intervallskaliert angenommen). Zudem wurde anhand einer zweifaktoriellen Varianzanalyse überprüft, ob ein Interaktionseffekt zwischen experimenteller Bedingung und Geschlecht vorliegt, was jeweils nicht der Fall war. Zwischen den weiteren Variablen sind Korrelationskoeffizienten nach Pearson bzw. Spearman (Korrelationen mit der Altersgruppe) ausgewiesen.

Tabelle 6.1.2. Zusammenhänge in Teilgruppen

	60+	Männer	selbst kgS	Mn.Check	gesamt
Hypothese 1 (r)	-0.50	-0.46	-0.43	-0.47	-0.45
Hypothese 2 (r)	-0.15 [†]	-0.01	-0.08	+0.00	-0.02
Hypothese 3 (η)	0.05	0.03	0.10	0.02	0.02
Hypothese 4 (η)	0.06	0.01	0.11 [†]	0.03	0.01
Hypothese 5 (η)	0.03	0.10	0.04	0.05	0.04
Lesedauer (η)	0.01	0.07	0.24 [†]	0.12 [†]	0.04
Lesefluss (η)	0.04	0.10	0.10	0.06	0.05
Textästhetik (η)	0.06	0.09	0.03	0.04	0.04
Teilstrichprobe n	235	530	226	690	1014

Anmerkung. Für Lesedauer, Lesefluss und Textästhetik ist der Zusammenhang mit der experimentellen Manipulation angegeben. [†] Zusammenhänge, welche die Effektstärke von 0,10 überschreiten und bei denen in der Gesamtstichprobe kein relevanter Effekt beobachtet wurde, sind mit einem Kreuz markiert.

6.2. Experiment mit Stimulustext „Kindererziehung“

Tabelle 6.2.1. Zusammenhänge insgesamt und nach Geschlecht

	Experimentelle Bedingung (η)	Schuld-zuschreibung	Kompetenz-beurteilung	Zustimmung zu Stereotypen
Experimentelle Bedingung (η)	–	.06° ♀ .05 .09 ♂	.02° ♀ .03 .04 ♂	.03° ♀ .07 .07 ♂
Schuld-zuschreibung	.06° ♀ .05 .09 ♂	–	-.55 ♀ -.51 -.57 ♂	-.00 ♀ +.01 -.02 ♂
Kompetenzbeurteilung	.02° ♀ .03 .04 ♂	-.55 ♀ -.51 -.57 ♂	–	-.10 ♀ -.10 -.10 ♂
Zustimmung Stereotyp	.03° ♀ .07 .07 ♂	-.00 ♀ +.01 -.02 ♂	-.10 ♀ -.10 -.10 ♂	–
Lesefluss	.05° ♀ .05 .12 ♂	-.09 ♀ -.11 -.08 ♂	+1.2 ♀ +.13 +.11 ♂	-.06 ♀ -.01 -.12 ♂
Textästhetik	.03° ♀ .01 .06 ♂	-.07 ♀ -.10 -.06 ♂	+1.5 ♀ +.14 +.16 ♂	-.03 ♀ +.02 -.08 ♂
Altersgruppe	.07° ♀ .13 ¹ .02 ♂	-.01 ♀ -.07 +.02 ♂	-.01 ♀ -.03 -.02 ♂	+0.3 ♀ +.06 -.02 ♂
Verwendung ge.ger. Sprache	.01° ♀ .06 .03 ♂	-.05 ♀ -.10 -.01 ♂	-.02 ♀ +.02 -.04 ♂	-.06 ♀ -.04 -.06 ♂

Anmerkung. °Eine zweifaktorielle Varianzanalyse zeigte *keinen* signifikanten Interaktionseffekt zwischen experimenteller Bedingung und Geschlecht der Befragten ($p \geq 0.05$). Dies gilt für alle aufgeführten abhängigen Variablen. ¹ Im Experiment „Kindererziehung“ führte die zufällige Zuteilung bei den Teilnehmerinnen zu geringfügigen Altersunterschieden zwischen den experimentellen Bedingungen.

Tabelle 6.2.2. Zusammenhänge in Teilgruppen

	60+	Männer	selbst kgS	Mn.Check	gesamt
Hypothese 1 (r)	-0.67	-0.57	-0.60	-0.55	-0.55
Hypothese 2 (r)	-0.02	-0.10	-0.02	-0.11	-0.10
Hypothese 3 (η)	0.06	0.01	0.04	0.03	0.03
Hypothese 4 (η)	0.11 [†]	0.04	0.09	0.02	0.02
Hypothese 5 (η)	0.10	0.09	0.11 [†]	0.06	0.06
Lesedauer (η)	0.04	0.04	0.02	0.06	0.04
Lesefluss (η)	0.08	0.12 [†]	0.05	0.10	0.05
Textästhetik (η)	0.05	0.06	0.03	0.08	0.03
Teilstrichprobe n	268	575	201	778	1053

Anmerkung. Für Lesedauer, Lesefluss und Textästhetik ist der Zusammenhang mit der experimentellen Manipulation angegeben. [†] Zusammenhänge, welche die Effektstärke von 0,10 überschreiten und bei denen in der Gesamtstichprobe kein relevanter Effekt beobachtet wurde, sind mit einem Kreuz markiert.

EXTENDED ABSTRACT

Priming stereotypes through gender-sensitive language in journalistic text

Frederieke Cirksena & Dominik J. Leiner

Frederieke Cirksena (B.A.), University of Groningen. Contact: frederieke.cirksena(at)gmx.net. ORCID: <https://orcid.org/0000-0001-7851-9118>

Dominik J. Leiner (Dr.), Ludwig Maximilian University of Munich, Department of Media and Communication, Oettingenstraße 67, 80538 München, Germany. Contact: dominik.leiner(at)ifkw.lmu.de. ORCID: <https://orcid.org/0000-0002-3862-3399>



© Frederieke Cirksena, Dominik J. Leiner

EXTENDED ABSTRACT

1. Introduction

The grammatical gender takes a prominent role in the German language. Different word forms in German refer to individuals or social groups, depending on whether they are women or men. “Professorinnen” refers to female professors, while “Professoren” refers to male professors. According to language standards, the male form may also refer to a mixed group of male and female professors or even a group of only female professors. This usage of the male form is known as the *generic masculine*. Literature found the *generic masculine* to cause mental representations that include primarily male members of a group (Sczesny & Stahlberg, 2001). *Gendered language* addresses this issue and results in a broader mental inclusion of women (Blake et al., 2008; Sczesny & Stahlberg, 2005). Examples of such language are the *pair form* (“Professorinnen und Professoren”) and newer ones such as the *gender asterisk* (“Professor*innen”).

While gendered language aims to solve issues in cognitive representation, critics (Vergoossen et al., 2020) have mainly focused on readability (Friedrich & Heise, 2019; Pöschko & Prieler, 2018) and text aesthetics (Blake & Klimmt, 2010; Trutkowski, 2018). Our study looked for detrimental effects of gendered language regarding cognitive representation.¹ We expected gendered language to increase the availability and salience of gender as a trait and its relevance in attributing the causes of failure.

2. Theory

The *pair form* explicitly names both female (“Professorinnen”) and male (“Professoren”) members of a social group. The *pair form* not only reminds readers that a social group includes female members (Sczesny & Stahlberg, 2001). It also reminds them of the existence of two different subgroups within the social group, namely female and male members. Moreover, there is a subtext to naming both subgroups: The distinction between the subgroups implies a relevant difference between them. Since the *pair form* highlights gender, it has the potential to make this trait more salient.

Priming theory (Scheufele, 2016; Weingarten et al., 2016) tells us that the world is far too complex for us to assess all of its aspects simultaneously. Priming specific characteristics will guide thinking about observations, especially in ambiguous situations. The topics covered by the news typically involve a lot of ambiguity (Friedman et al., 1999), so priming can effectively affect how readers interpret the news.

Gender stereotypes (Eckes, 1997) are powerful interpretation schemes (Steffens & Ebert, 2016), well-researched in the field of professions (Lawrimore & Marcis, 1994; Parks & Robertson, 1998). Gender stereotypes regarding professions are

1 Dataset, documentation and analyses are publicly available at <https://osf.io/ga4cp/>

reinforced by the ratio between male and female employees in a field (Barth & Rice, 2016; Sáinz et al., 2016) and by personality traits that are preferable for a job (White & White, 2006). One such pair of stereotypes is that women are better with care and children (Prentice & Carranza, 2002; Troche & Rammsayer, 2011), and men are better with technology and computers (Axell & Boström, 2019; Eckes, 2008). Assume that a news report allows for different interpretations, for example, regarding a possible explanation of why things went wrong. People will use cues (Roskos-Ewoldsen et al., 2008) and context (Parks & Robertson, 1998) to address ambiguity and create a coherent cognitive model of the situation (Dutke, 1998). Consequently, if language primes people that gender is an important characteristic of people, gender could advance to become the central explanatory pattern.

3. Research design and hypotheses

We created a 2,000-character news article based on actual news stories for each stereotype mentioned above. The first article (*child care*) deals with an accident during a nursery excursion. The second article (*software*) describes the chaos that occurred during the candidate elections of the Democratic Party in Iowa due to technical issues with their voting app. Both texts describe a case of damage and allow for two interpretations regarding the question of who is responsible for the damage: The circumstances or the professionals (*educators* or *programmers*).

In two experiments – one for each article/stereotype – we manipulated the word form on three levels. Subjects read the text either in a condition in which the professionals were referred to by (1) the generic masculine, (2) the pair form, or (3) by the gender asterisk.

We hypothesized that priming the gender by a gendered article would activate stereotypes (H3), negatively affecting how much competence readers attribute to the professionals. There would either be an indirect effect by activating the stereotypes (H2) or a direct effect because the word form made clear that some professionals had a gender that did not “match” the profession (H4), as far as social stereotypes are concerned. The lower competence perception would then (H1) cause readers to attribute guilt to the professionals rather than the conditions. Our overarching hypothesis was that readers who read one of the gendered articles would be more likely to attribute guilt to the professionals (H5). Based on our theory, the *pair form* would exert a stronger effect than the *gender asterisk* form.

4. Sample

We expected weak effects and assumed that these effects would vary between male and female readers. Subjects were addressed through the non-commercial *SoSci Panel* for an alleged “study on the journalistic depiction of accidents”. A power analysis suggested 930 men and women each. We slightly exceeded this goal with 2,066 valid responses (954 women, 1,105 men, 7 diverse) after removing dropouts, respondents not residing in Germany, incomplete records, and respondents who spent too little time on the article. Parallel analysis of the non-cleaned data set en-

sured that data cleaning did not cause artifacts. The sample represented all age groups of 25 years and above. The majority of the respondents indicated a high level of formal education (89% matriculation standard or above).

5. Method

Both experiments (*child care* and *software*) were designed as online experiments and followed the same scheme. The online questionnaire addressed the issues of priming and gendered language as late as possible to avoid interfering with sensitive priming effects. The news article was presented after a short introduction, informed consent, and an icebreaker question. It was immediately followed by a visual analog scale, asking to what degree the professionals were to blame compared to the prevailing circumstances. Subsequently, the questionnaire asked for the professionals' competence (3 bipolar items, designed for each experiment separately), text aesthetics and readability (Rothmund & Christmann, 2002), general gender competence perception (8 bipolar items), and agreement with stereotypes (4 items for women and technology, 6 items for men and child care, both asked in both experiments). A closed choice question measured treatment recognition, and lastly, we asked for demographics and how often respondents used gendered language (single-item with 4 response options).

Outside of the stimulus text, the questionnaire avoided naming social groups; where that was not possible, we used gender-neutral wording ("professionals"). Where possible, scales used positive and negative items to avoid faux-correlations caused by general acquiescence.

6. Results

As the experiments used slightly different measures (especially for the attribution of guilt and the professionals' competence), we tested all hypotheses separately for both experiments using Pearson's correlation and analyses of variance (ANOVA).

The main finding is that we did not find any relevant differences between the three article forms. For better comparability between correlations and analyses of variance, and given the low discrimination in the present case ($\eta^2 < 0.01$ in each case), we report the non-squared effect size ($\eta = \sqrt{\eta^2}$).

We observed a stable correlation between the anticipated competence of the professionals and to what extent professionals are to blame ($r_{\text{Software}} = -.45$, $r_{\text{ChildCare}} = -.55$, each $p < .001$, H1). Even small cues could activate stereotypes, so we attribute limited meaning to the fact that we did not observe any effect of the wording on stereotype agreement ($\eta_{\text{Software}} = 0.02$, $\eta_{\text{ChildCare}} = 0.03$, each n.s., H3) and that stereotype agreement did no more than marginally correlate with attribution of guilt ($r_{\text{Software}} = -.02$, $n = 1014$, n.s., $r_{\text{ChildCare}} = -.10$, $n = 1053$, $p < .001$, H2).

More revealing is the observation that there was no effect of the use of gendered language on how competent the professionals were rated ($\eta_{\text{Software}} = 0.01$, $\eta_{\text{ChildCare}} = 0.02$, each n.s., H4). Observing no significant effects does not mean that there are none. Consequently, we performed a pairwise equivalence test (Lakens et al., 2018) on this non-effect. This test also showed that in both experi-

ments, the effect of gendered language is too small to have any practical meaning (d within $[-0.2|+0.2]$, TOST, $p < .05$).

As one might expect from the previous findings, gendered language also had no effect on overall guilt attribution ($\eta_{\text{Software}} = 0.04$, $\eta_{\text{ChildCare}} = 0.06$, each n.s., H5). Furthermore, our experiments did not find any relevant effects ($\eta^2 < 0.01$) of gendered language on text aesthetics and readability assessment. We also tested all (non-)effects for interaction with the respondents' gender but found none.

7. Discussion

Previous research focused on desired effects of gendered language on cognition and undesired effects on reading. In two experiments, we looked for undesirable effects of gendered language on cognition, namely if gendered language might emphasize the interpretation of journalistic texts driven by gender stereotypes.

Based on 2,066 valid questionnaires in two experiments, we find no evidence to support the hypothesis of such undesired effects. Instead, an equivalence test showed that gendered language had no relevant effect on text interpretation. Our findings show that respondents more or less ignored forms of gendered language, also in their text ratings. These results are in line with earlier findings from Blake and Klimmt (2010).

Choosing a highly educated sample that has shown liberal political attitudes in earlier studies (Leiner, 2016) may have obscured the effects of gender stereotypes. However, subsamples that we assumed to be more susceptible to gender stereotypes (elder male respondents, for example) showed no systematic effects in additional analyses.

We cannot exclude the possibility of long-term effects of language. Yet, the results of our experiments do not provide any evidence against using the *pair form* or the *gender asterisk* in journalistic texts.

References

- Axell, C., & Boström, J. (2019). Technology in children's picture books as an agent for reinforcing or challenging traditional gender stereotypes. *International Journal of Technology and Design Education*, 31(2021), 27–39. <https://doi.org/10.1007/s10798-019-09537-1>
- Barth, J. M., & Rice, L. (2016). A tale of two gender roles: The effects of implicit and explicit gender role traditionalism and occupational stereotypes on hiring decisions. *Gender Issues*, 34, 86–102. <https://doi.org/10.1007/s12147-016-9175-4>
- Blake, C., & Klimmt, C. (2010). Geschlechtergerechte Formulierungen in Nachrichtentexten [Gender-appropriate wording in news texts]. *Publizistik*, 55(3), 289–304. <https://doi.org/10.1007/s11616-010-0093-2>
- Blake, C., Klimmt, C., & Pompetzki, V. (2008). Geschlechterrepräsentation in Nachrichtentexten: Der Einfluss von geschlechterbezogenen Sprachformen und Fallbeispielen auf den gedanklichen Einbezug von Frauen und die Bewertung der Beitragsqualität. [Gender representation in news texts: The influence of gendered language forms and

- case studies on the mental inclusion of women and ratings of news story quality]. *Medien & Kommunikationswissenschaft*, 56(1), 3–20.
- Dutke, S. (1998). Zur Konstruktion von Sachverhaltsrepräsentationen beim Verstehen von Texten: 15 Jahre nach Johnson-Lairds Mental Models. [On the construction of factual representations in text comprehension: 15 years after Johnson-Laird's mental models.]. *Zeitschrift für Experimentelle Psychologie*, 45(1), 42–59.
- Eckes, T. (1997). *Geschlechterstereotype: Frau und Mann in sozialpsychologischer Sicht* [Gender stereotypes: Woman and man in a social psychological perspective]. Centaurus.
- Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen [Gender stereotypes: Of roles, identities and prejudices]. In R. Becker & B. Kortendiek (Eds.), *Handbuch Frauen- und Geschlechterforschung* (2nd ed.). VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-91972-0>
- Friedman, S. M., Dunwoody, S., & Rogers, C. L. (Eds.). (1999). *Communicating uncertainty: Media coverage of new and controversial science*. L. Erlbaum.
- Friedrich, M. C. G., & Heise, E. (2019). Does the use of gender-fair language influence the comprehensibility of texts? An experiment using an authentic contract manipulating single role nouns and pronouns. *Swiss Journal of Psychology*, 78(1–2), 51–60. <https://doi.org/10.1024/1421-0185/a000223>
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269. <https://doi.org/10.1177/2515245918770963>
- Lawrimore, K. W., & Marcis, J. G. (1994). Current occupational stereotypes. *Atlantic Economic Journal*, 22(3), 87. <https://doi.org/10.1007/BF02301808>
- Leiner, D. J. (2016). Our research's breadth lives on convenience samples: A case study of the online respondent pool "SoSci Panel". *Studies in Communication & Media*, 5(4), 367–396. <https://doi.org/10.5771/2192-4007-2016-4-367>
- Parks, J. B., & Robertson, M. A. (1998). Influence of age, gender and context on attitudes toward sexist/nonsexist language: Is sport a special case? *Sex Roles*, 38(5–6), 477–493. <https://doi.org/10.1023/A:1018766023667>
- Pöschko, H., & Prieler, V. (2018). Zur Verständlichkeit und Lesbarkeit von geschlechtergerecht formulierten Schulbuchtexten [On the comprehensibility and readability of gender-appropriate textbook texts]. *Zeitschrift für Bildungsforschung*, 8, 5–18. <https://doi.org/10.1007/s35834-017-0195-2>
- Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, 26(4), 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>
- Roskos-Ewoldsen, D. R., Roskos-Ewoldsen, B., & Dillman Carpentier, F. (2008). Media priming. An updated synthesis. In J. Bryant & M. B. Oliver (Eds.), *Media Effects Advances in Theory and Research* (pp. 74–93). Routledge.
- Rothmund, J., & Christmann, U. (2002). Auf der Suche nach einem geschlechtergerechten Sprachgebrauch: Führt die Ersetzung des generischen Maskulinums zu einer Beeinträchtigung von Textqualitäten? [In search of gender-equitable language use: Does the substitution of the generic masculine lead to an impairment of textual qualities?]. *Muttersprache*, 112(2), 115–136.
- Sáinz, M., Meneses, J., López, B.-S., & Fàbregues, S. (2016). Gender stereotypes and attitudes towards information and communication technology professionals in a sample of

- Spanish secondary students. *Sex Roles*, 74, 154–168. <https://doi.org/10.1007/s11199-014-0424-2>
- Scheufele, B. (2016). *Priming*. Nomos. <https://doi.org/10.5771/9783845263120>
- Sczesny, S., & Stahlberg, D. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen [Effects of the generic masculine and alternative forms of language on the mental inclusion of women]. *Psychologische Rundschau*, 52(3), 131–140. <https://doi.org/10.1026//0033-3042.52.3.131>
- Sczesny, S., & Stahlberg, D. (2005). Cognitive effects of masculine generics in German: An overview of empirical findings. *Communications*, 30(1), 1–21. <https://doi.org/10.1515/comm.2005.30.1.1>
- Steffens, M. C., & Ebert, I. D. (2016). *Frauen – Männer – Karrieren: Eine sozialpsychologische Perspektive auf Frauen in männlich geprägten Arbeitskontexten* [Women – men – careers: A social psychological perspective on women in male-dominated work contexts]. VS Verlag für Sozialwissenschaften.
- Troche, S. J., & Rammsayer, T. H. (2011). Eine Revision des deutschsprachigen Bem Sex-Role Inventory [A revision of the German-Language Bem Sex-Role Inventory]. *Klinische Diagnostik und Evaluation*, 4, 262–283.
- Trutkowski, E. (2018). Wie generisch ist das generische Maskulinum? Über Genus und Sexus im Deutschen [How generic is the generic masculine? About genus and sexus in German]. *ZAS Papers in Linguistics*, 59, 83–96. <https://doi.org/10.21248/zaspil.59.2018.437>
- Vergoossen, H. P., Renström, E. A., Lindqvist, A., & Gustafsson Sendén, M. (2020). Four dimensions of criticism against gender-fair language. *Sex Roles*, 83(4–5), 328–337. <https://doi.org/10.1007/s11199-019-01108-x>
- Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, 142(5), 472–497. <https://doi.org/10.1037/bul0000030>
- White, M. J., & White, G. B. (2006). Implicit and explicit occupational gender stereotypes. *Sex Roles*, 55(3–4), 259–266. <https://doi.org/10.1007/s11199-006-9078-z>