



Stimmzettel nutzerfreundlicher gestalten: Empfehlungen auf Grundlage eines Survey-Experimentes

Dominic Nyhuis · Michael Jankowski · Felix Münchow ·
Morten Harmening · Jannik Bruns · Max Holly · Josip Karacic ·
Christoph Hönnige

Online publiziert: 10. September 2020
© Der/die Autor(en) 2020

Zusammenfassung Die vorliegende Arbeit nimmt die große empirische Vielfalt in der Gestaltung von Stimmzetteln zum Anlass, um sich mit Frage zu beschäftigen, welche Stimmzettelmerkmale von Wählerinnen und Wählern gewünscht werden. Dazu greift die Studie auf ein Discrete-Choice-Experiment im Rahmen eines Online-Surveys mit rund 1200 deutschen Teilnehmenden zurück. Im Experiment entscheiden sich die Teilnehmenden zwischen fiktiven Stimmzetteln, die sich im Hinblick auf die Informationen zum Wahlsystem, die Informationen zu den Kandidierenden und das Design unterscheiden. Im Ergebnis zeigt sich, dass die Teilnehmenden einer zu ausführlichen, aber auch einer zu knappen Erläuterung des Wahlsystems kritisch gegenüberstehen. Weiterhin schätzen die Teilnehmenden die Angabe von Beruf und Alter der Kandidierenden, während die Angabe von Kandidatenadressen skeptischer bewertet wird. Zudem kann festgestellt werden, dass die Verständlichkeit das wichtigste Kriterium für die Bewertung von Stimmzetteln ist. Auf Grundlage der Befunde werden Merkmale geeigneter Stimmzettel herausgearbeitet und ein gelungener Stimmzettel vorgestellt.

Das Projekt wurde im Rahmen des Seminars „Experimente in der Politikwissenschaft“ entwickelt, das im Sommersemester 2019 unter der Leitung von Dominic Nyhuis an der Leibniz Universität Hannover durchgeführt wurde.

Zusatzmaterial online Zusätzliche Informationen sind in der Online-Version dieses Artikels (<https://doi.org/10.1007/s11615-020-00267-x>) enthalten.

D. Nyhuis (✉) · F. Münchow · M. Harmening · J. Bruns · M. Holly · J. Karacic · C. Hönnige
Institut für Politikwissenschaft, Leibniz Universität Hannover, Schneiderberg 50, 30167 Hannover,
Deutschland
E-Mail: d.nyhuis@ipw.uni-hannover.de

M. Jankowski
Institut für Sozialwissenschaften, Carl von Ossietzky Universität Oldenburg, Oldenburg, Deutschland

Schlüsselwörter Stimmzettel · Experimentalforschung · Survey-Experiment · Discrete-Choice-Experiment · Kommunalwahlen

Making Ballots More User Friendly: Recommendations Based on a Survey Experiment

Abstract This paper takes the considerable variation in ballot design as a point of departure to ask which ballot features are preferred by voters. Using a discrete choice experiment as part of an online survey of 1200 German participants, we asked respondents to choose among fictitious ballots that differed in terms of the information about the electoral system, information about the candidates, and the ballot design. The results show that respondents are opposed to lengthy explanations of the electoral system, but they are also critical when the explanation is too brief. Providing the candidates' occupations and ages is welcomed, whereas candidate addresses are viewed with more skepticism. Overall, comprehension is the most important criterion for ballot evaluation. Based on the results, we discuss features of good ballots and present an example of a well-designed ballot.

Keywords Ballots · Experiments · Survey experiments · Discrete choice experiments · Local elections

1 Einleitung

Seit mehreren Jahren wird die Gestaltung von Stimmzetteln in der politikwissenschaftlichen Forschung problematisiert. Angefangen mit den viel beachteten Studien über den Effekt des Stimmzettel-Layouts bei den umstrittenen US-Präsidentenwahlen im Jahr 2000 (Sinclair et al. 2000; Wand et al. 2001), gibt es mittlerweile eine Reihe von Arbeiten, die sich mit Wahltechnologien im Allgemeinen und mit Stimmzetteln im Besonderen beschäftigen (Ansolabehere und Stewart 2005; Calvo et al. 2009; Hanmer et al. 2010). In diesem Forschungszweig ist die Einsicht gewachsen, dass Stimmzettel unterschiedlich gut gestaltet sind. Vor diesem Hintergrund ist die große Vielfalt von Stimmzetteln in der Verwaltungspraxis kritisch zu hinterfragen. Die Vermutung liegt nahe, dass Entscheidungen über die Gestaltung von Stimmzetteln in Politik und Verwaltung häufig auf Gewohnheiten zurückzuführen sind und eine systematische Prüfung der Eignung von Stimmzetteln im Regelfall unterbleibt. Die vorliegende Arbeit soll dazu beitragen, diesen Missstand zu beheben und die Gestaltung von Stimmzetteln auf ein empirisch abgesichertes Fundament zu stellen. Zu diesem Zweck werden verschiedene Gestaltungsmerkmale von Stimmzetteln aus der gegenwärtigen Verwaltungspraxis im Rahmen eines Survey-Experimentes untersucht, um auf dieser Basis praktische Hinweise über die Gestaltung von Stimmzetteln geben zu können.

Obwohl die Aufmerksamkeit für Wahltechnologien in den vergangenen Jahren zugenommen hat, so gibt es doch deutliche Lücken in der Literatur. Aufgrund der Vielfalt von Wahltechnologien in den Vereinigten Staaten liegt der Schwerpunkt der US-zentrierten Forschung auf dem Vergleich und dem Effekt der verschiedenen

Technologien (Bullock und Hood 2002; Herrnson et al. 2012). Im Gegensatz dazu spielt die Gestaltung von Stimmzetteln in der Forschung bisher nur eine untergeordnete Rolle (Carman et al. 2008; Kimball und Kropf 2005; Pachón et al. 2017). Zudem werten die bestehenden Arbeiten zur Gestaltung von Stimmzetteln zumeist Beobachtungsdaten sowie in seltenen Fällen Daten natürlicher Experimente aus. Im Gegensatz dazu ist die vorliegende Studie eine von wenigen Arbeiten, welche die Gestaltungsmerkmale von Stimmzetteln im Rahmen eines Survey-Experimentes untersucht.

Um Empfehlungen über die Gestaltung von Stimmzetteln aussprechen zu können, muss ein Bewertungsmaßstab angelegt werden, den es offenzulegen gilt. In dieser Arbeit handelt es sich dabei um die Bewertungen der Wählerinnen¹. Während die Einstellungen der Wählerinnen in unseren Augen ein entscheidender Maßstab für die Gestaltung von Stimmzetteln sein sollten, gibt es durchaus andere denkbare Kriterien, zwischen denen bisweilen Zielkonflikte bestehen. Ein offensichtlicher Zielkonflikt liegt beispielsweise zwischen dem Informationsgehalt und der Nutzerfreundlichkeit von Stimmzetteln vor, da ein informationsreicher Stimmzettel in der Tendenz ein wenig unübersichtlicher ist. Ein weiterer denkbare Maßstab ist die Neutralität von Stimmzetteln. Ein gelungener Stimmzettel sollte die Wahlentscheidung möglichst nicht beeinflussen und entsprechend problematisch sind Effekte des Stimmzettel-Layouts auf das Wahlverhalten (Carman et al. 2008; Sinclair et al. 2000; Wand et al. 2001). Dabei gibt es jedoch Grenzen des Neutralitätsgebots. So bieten Stimmzettel verschiedene Kandidateninformationen, beispielsweise über den Beruf von Kandidierenden, die sich durchaus wahlentscheidend auswirken und bestimmte Kandidierende benachteiligen mögen. Welche solcher Informationen jedoch zulässig oder unzulässig sind, lässt sich nur schwer herleiten. Um derartige Wertentscheidungen zu treffen, ist der Wille der Wählerinnen ein naheliegendes Kriterium zur Ausgestaltung von Stimmzetteln.

Während wir in dieser Arbeit deshalb den Willen der Wählerinnen als Maßstab für einen gelungenen Stimmzettel anlegen, bleibt festzuhalten, dass auch andere Maßstäbe denkbar sind, die womöglich einen anderen Stimmzettel erforderlich machen. Dass Zielkonflikte zwischen den Kriterien bestehen, bedeutet allerdings nicht, dass ein gelungener Stimmzettel nicht auch mehrere Kriterien maximieren kann. So liegt es auf der Hand, dass ein gut gestalteter Stimmzettel den Anteil ungültiger Stimmen verringert. Unabhängig von den anzulegenden Kriterien für die Stimmzettelgestaltung ist eine zentrale Botschaft dieser Arbeit, dass ein Stimmzettel im Hinblick auf ein oder mehrere Kriterien optimiert werden kann. Dagegen kommen in der gegenwärtigen Verwaltungspraxis häufig Stimmzettel zur Anwendung, die im Hinblick auf kein sichtbares Kriterium optimiert wurden.

Für das Experiment wurden fiktive Stimmzettel entworfen und einem quotierten Sample von Deutschen zur Bewertung vorlegt, um Merkmale gelungener Stimmzettel zu identifizieren. Dabei sind die Eigenschaften der fiktiven Stimmzettel an die Stimmzettel angelehnt, die gegenwärtig bei deutschen Kommunalwahlen verwendet werden. Untersucht wurden im Experiment die Erläuterungen des Wahlsystems, die verfügbaren Kandidateninformationen und das Design des Stimmzettels. Die Studie

¹ Mit der Verwendung des generischen Femininums sollen alle Geschlechter eingeschlossen werden.

betritt in der Stimmzettelforschung Neuland, da sie auf Grundlage eines Discrete-Choice-Experimentes den simultanen Einfluss verschiedener Faktoren messen kann. Im Gegensatz dazu hat die bestehende Experimentalforschung lediglich den Einfluss weniger Gestaltungselemente in den Blick genommen (bspw. Reidy und Buckley 2015).

Im Ergebnis kann festgestellt werden, dass die Teilnehmerinnen differenzierte Erwartungen an die Erläuterung des Stimmverfahrens haben. Während zu knappe Informationen abgelehnt werden, ist auch ein Übermaß an Informationen unerwünscht, da es bei den Wählerinnen zu Überforderung und Unverständnis führt. Darüber hinaus bevorzugen die Teilnehmerinnen Stimmzettel mit der Angabe des Berufs und des Alters der Kandidierenden. Der durchaus üblichen Angabe von Adressdaten stehen die Wählerinnen dagegen sichtlich reservierter gegenüber. Insgesamt ist die Verständlichkeit der wichtigste Faktor bei der Bewertung der Stimmzettel.

Im folgenden Abschnitt diskutieren wir drei grundlegende Dimensionen der Gestaltung von Stimmzetteln. Auf Basis dieses Analyserasters werden im darauffolgenden Abschnitt die Merkmale von Stimmzetteln klassifiziert, die gegenwärtig bei deutschen Kommunalwahlen zur Anwendung kommen. Die so identifizierten Merkmale werden bei der Gestaltung der fiktiven Stimmzettel aufgegriffen, die den Teilnehmerinnen im Rahmen des Discrete-Choice-Experimentes vorlegt werden, um die Präferenzen für die verschiedenen Stimmzettelmerkmale zu erheben. Der Beitrag schließt mit einem Appell an die Länderverwaltungen, ihre Stimmzettel zu optimieren und deren Eignung systematisch zu überprüfen.

2 Dimensionen der Stimmzettelgestaltung

Um relevante Faktoren der Stimmzettelgestaltung zu identifizieren, bedarf es eines Analyserasters. Zur Entwicklung eines solchen Rasters wählen wir in dieser Arbeit ein zweistufiges Verfahren. In einem ersten Schritt unterscheiden wir zwischen zwei grundlegenden Dimensionen der Stimmzettelgestaltung: dem Informationsgehalt und dem Layout von Stimmzetteln, die in diesem Abschnitt knapp in die bestehende Literatur eingebettet werden. Im folgenden Abschnitt werden dann die kommunalen Stimmzettel in den deutschen Bundesländern entlang dieser beiden Dimensionen ausgewertet, um einen systematischen Eindruck der entscheidenden Merkmale zu gewinnen. Die auf diese Weise identifizierten Faktoren werden anschließend im Discrete-Choice-Experiment aufgegriffen, um die Eignung der empirischen Varianten zu prüfen.

Das beschriebene Verfahren ermöglicht es uns, die große Vielfalt der Stimmzettel in der deutschen Wahlpraxis systematisch abzubilden. Insbesondere kann mit dem gewählten Vorgehen ein Experiment entworfen werden, das sich nah an der Verwaltungspraxis bewegt, sodass sich aus den Ergebnissen leicht praktische Handlungsempfehlungen ableiten lassen. Auf eine naheliegende Einschränkung dieses Vorgehens sei jedoch explizit hingewiesen. Wenn besonders geeignete Merkmalsausprägungen in der Praxis nicht auftreten, dann werden sie durch das gewählte Verfahren nicht erfasst. Zwei Punkte seien genannt, um diesem Einwand zu begegnen. Erstens ist die empirische Varianz der Stimmzettel derart groß, dass entscheidende

Varianten durch das beschriebene Verfahren eher nicht unentdeckt bleiben. Zweitens konnten weder bei der Entwicklung des Experimentaldesigns, noch bei der Sichtung der einschlägigen Forschungsliteratur Merkmale ausgemacht werden, die zugleich geeignet und für den gewählten Anwendungsfall zweckmäßig erscheinen. Dennoch liegt bei dem gewählten Verfahren ein gewisser Status Quo Bias auf der Hand, da radikalere Neuentwürfe zur Gestaltung von Stimmzetteln praktisch ausgeschlossen werden. Zu nennen wären hier beispielsweise Parteilogos oder Kandidatenbilder, die auf Stimmzetteln anderer Länder gelegentlich abgedruckt werden (Banducci et al. 2008; Reynolds und Steenbergen 2006; Moehler und Conroy-Krutz 2016).

2.1 Der Informationsgehalt von Stimmzetteln

Die Informationen auf einem Stimmzettel lassen sich grob in zwei Kategorien unterteilen. Zum einen enthalten Stimmzettel üblicherweise Informationen über das Wahlsystem und die korrekte Stimmabgabe. Entsprechende Hinweise sollen sicherstellen, dass Wählerinnen den Stimmzettel korrekt ausfüllen können. Gerade vor dem Hintergrund der mangelhaften Kenntnisse über Wahlsysteme (Behnke 2015; Jankowski et al. 2020; Karp 2006; Schmitt-Beck 1993) erscheinen solche Hinweise gar zwingend, um einen hohen Anteil fehlerhafter oder zumindest fehlintendierter Stimmabgaben zu vermeiden. An diesem Punkt stellt sich vor allem die Frage, wie die relevanten Informationen auf dem Stimmzettel knapp präsentiert werden können.

Zum anderen stellen Stimmzettel Informationen über die Wahlvorschläge bereit. Durch zahlreiche Studien ist bekannt, dass solche Kandidateninformationen, sogenannte „ballot cues“, einen Einfluss auf das Verhalten der Wählerinnen haben. Dabei ist es bei der Gestaltung von Stimmzetteln unvermeidlich, dass die Namen der Kandidierenden genannt werden.² Aus dem Namen lassen sich in der Regel jedoch bereits das biologische Geschlecht und ein vermuteter Migrationshintergrund der Kandidierenden ableiten. Tatsächlich leiten Wählerinnen diese Informationen aus den Kandidatennamen auch ab und lassen sie in ihre Wahlentscheidungen einfließen (Kelley und McAllister 1984; Matson und Fine 2006; Portmann und Stojanović 2019).

Häufig werden auf Stimmzetteln zudem das Alter, der Beruf, oder der Wohnort der Kandidierenden genannt. Daneben sind noch weitere Faktoren denkbar, in der Praxis aber selten. Auch für diese Faktoren ist bekannt, dass sie einen Einfluss auf die Kandidatenpräferenzen der Wählerinnen haben. So zeigen verschiedene Studien, dass der Wohnort einen erheblichen Einfluss auf den Wahlerfolg von Kandidierenden hat (Arzheimer und Evans 2012; Campbell and Cowley 2014; Campbell et al. 2019; Jankowski 2016; Tavits 2010). Ähnliche Ergebnisse können für den Kandidatenberuf (Campbell und Cowley 2014; Carnes und Lupu 2016; McDermott 2005; Mechtel 2014) und das Kandidatenalter festgestellt werden (Piliavin 1987; Sigelman und Sigelman 1982; Webster und Pierce 2019). Dass solche Informationen einen wahlentscheidenden Effekt haben, ist jedoch unabhängig von der Frage, ob ein Stimmzettel diese Informationen auch bieten sollte. Der vorliegende Beitrag

² Ein Stimmzettel ohne Nennung von Kandidatennamen wäre in Verhältniswahlsystemen mit geschlossenen Listen durchaus denkbar, aber zumindest fragwürdig.

nimmt an diesem Punkt die Perspektive der Wählerinnen und Wähler ein und fragt, welche Informationen über die Kandidierenden die Wählerinnen sich wünschen. Der Fokus auf den Wunsch der Wählerinnen soll freilich nicht ausblenden, dass es anders gelagerte Gründe für oder gegen die Nennung bestimmter Informationen geben mag.

2.2 Das Layout von Stimmzetteln

Das Layout des Stimmzettels kann sich auf vielfältige Weise auf das politische Verhalten auswirken. Da das konkrete Layout eines Stimmzettels erheblich durch das Wahlsystem beeinflusst wird (Reynolds und Steenbergen 2006) und es zahllose Möglichkeiten zur grafischen Gestaltung eines Stimmzettels gibt, ist es deutlich schwieriger, Merkmale und/oder Kriterien für das Layout von Stimmzetteln aus der Forschung abzuleiten. Das heißt jedoch nicht, dass die bestehende Forschung sich nicht mit dem Einfluss des Layouts von Stimmzetteln auf politisches Verhalten befasst hat. Von Interesse ist hierbei in der Regel der Effekt des Layouts auf den Anteil ungültiger Stimmen oder die getroffene Wahlentscheidung.

Zum einen existieren Studien, in denen der Einfluss von Veränderungen des Layouts auf das Wahlergebnis analysiert wird. Diese Studien sind zumeist dadurch motiviert, dass ein bestimmtes Merkmal im Verdacht steht, den Wahlausgang beeinflusst zu haben. Das bekannteste Beispiel ist der sogenannte Butterfly-Stimmzettel bei der US-Präsidentenwahl aus dem Jahr 2000 (Sinclair et al. 2000; Wand et al. 2001).³ Weniger prominente, aber ähnlich gelagerte Beispiele werden in den Studien von Carman et al. (2008) und Pachón et al. (2017) diskutiert. Carman und Kollegen zeigen, dass schlecht gestaltete Stimmzettel bei den schottischen Parlamentswahlen im Jahr 2007 zu einem erheblichen Anstieg der ungültigen Stimmen geführt haben. Pachón et al. (2017) machen ein ähnliches Argument, legen den Fokus aber auf die Komplexität von Stimmzetteln. Auf Basis eines Experimentes und kolumbianischen Wahldaten können die Autoren zeigen, dass komplexe Stimmzettel häufig zu fehlerhaften Stimmabgaben führen. Sie können außerdem zeigen, dass gerade weniger gebildete Wählerinnen von einem komplexen Stimmzetteldesign überfordert sind und das Design somit einen verzerrenden Effekt auf die Zusammensetzung der Wählerschaft hat (vgl. Nyhuis 2014; Schäfer und Schoen 2013).

Zum anderen befassen sich einige wenige Studien experimentell mit dem Effekt des Layouts von Stimmzetteln auf politisches Verhalten. Bei diesen Studien wird, ähnlich der vorliegenden Studie, das Layout des Stimmzettels zufällig variiert und der Effekt auf politisches Verhalten analysiert. So konnte gezeigt werden, dass schlecht gestaltete und komplexe Wahlzettel einen signifikanten Effekt auf den Anteil ungültiger Stimmen haben (Herrnson et al. 2012). In Wahlsystemen mit Personenstimmen hat zudem bereits die Reihung der Kandidierenden einen Einfluss

³ Der „Butterfly“-Stimmzettel kam bei der US-Präsidentenwahl 2000 in Florida zur Anwendung. Auf dem Stimmzettel wurden die Kandidatennamen abwechselnd auf der linken und rechten Seite abgedruckt; die Ankreuzflächen befanden sich in der Mitte des Stimmzettels. Auf der linken Seite stand an oberster Stelle George W. Bush, gefolgt von Al Gore. Auf der rechten Seite stand der chancenlose Pat Buchanan ganz oben. Dieses Layout führte zu Verwirrung bei zahlreichen Wählerinnen, die ungewollt für Pat Buchanan stimmten.

auf das Wahlergebnis. Das ist selbst dann der Fall, wenn die Reihung der Kandidierenden alphabetisch oder randomisiert ist (Kim et al. 2015; Reidy und Buckley 2015).

Für diese Arbeit ist die zentrale Schlussfolgerung aus der bestehenden Forschung, dass sich das Layout von Stimmzetteln eindeutig auf politisches Verhalten auswirkt. Da die Möglichkeiten zur grafischen Gestaltung jedoch außerordentlich vielfältig sind, betrachten wir im nächsten Abschnitt die empirischen Varianten bei deutschen Kommunalwahlen, um Faktoren für das Experiment zu ermitteln.

3 Die Varianten kommunaler Stimmzettel in Deutschland

Um für das Experiment konkrete Merkmale entlang der im vorangegangenen Abschnitt herausgearbeiteten Dimensionen abzuleiten, wird in diesem Abschnitt ein Blick auf die kommunalen Stimmzettel in Deutschland geworfen. Dabei werden die in der Praxis auftretenden Merkmale und ihre Ausprägungen auf den Ebenen *Informationen zum Wahlsystem*, *Informationen zu den Kandidierenden* und *Layout* herausgearbeitet. Die auf diese Weise identifizierten Varianten werden in einem nächsten Schritt in einem fiktiven Stimmzettel aufgegriffen und den Teilnehmerinnen des Survey-Experimentes zur Bewertung vorgelegt. So können systematische Rückschlüsse darüber gewonnen werden, welche Varianten kommunaler Stimmzettel sich aus Sicht der Wählerinnen besonders eignen.

Für die Kategorisierung werden die kommunalen Stimmzettel in den deutschen Flächenländern in den Blick genommen. Die Stadtstaaten werden aufgrund ihres besonderen Status nicht berücksichtigt. Der Fokus dieser Studie auf Kommunalwahlen lässt sich vor allem durch die Komplexität des deutschen Kommunalwahlrechts mit seinen Mehrfachstimmen in einem offenen Listensystem mit Kumulier- und Panaschiermöglichkeit begründen. In einem derart anspruchsvollen Wahlkontext ist die gelungene Gestaltung von Stimmzetteln besonders bedeutsam.

Der Klassifizierung seien einige Bemerkungen vorangestellt. Erstens ist darauf hinzuweisen, dass das kommunale Wahlrecht die Gestaltung von Stimmzetteln erheblich vorbestimmt. So hat es seit der Wiedervereinigung eine Reihe von kommunalen Wahlrechtsreformen gegeben, die in der Tendenz zu einer Angleichung des kommunalen Wahlrechts in Richtung der offenen Listenswahl geführt haben. Dennoch können nach wie vor einige Unterschiede zwischen den Kommunalwahlsystemen festgestellt werden, nicht zuletzt mit Blick auf die Zahl der zu vergebenden Stimmen. Besonderheiten sind hier die Kommunalwahlsysteme von Schleswig-Holstein (personalisierte Verhältniswahl mit einer Stimme pro Direktmandat), Nordrhein-Westfalen (personalisierte Verhältniswahl mit einer Stimme) und des Saarlandes (geschlossene Listenswahl). Die restlichen Kommunalwahlsysteme, auch die komplexen süddeutschen, beschränken die Zahl der möglichen Stimmen unabhängig von der Gesamtstimmenzahl auf drei pro Kandidierenden, sodass die verschiedenen Regelungen gut innerhalb eines Experimentaldesigns abgebildet werden können. Bei der Klassifizierung der verschiedenen Stimmzettel wird jeweils vermerkt, wenn eine Kategorie innerhalb eines Wahlsystems nicht sinnvoll genutzt werden kann.

Tab. 1 Merkmale kommunaler Stimmzettel in Deutschland

Land	Wahl-	Kandidierende			Design			
	system	Beruf	Adresse	Alter	Stimmen-	Listen-	Stimm-	Pfeile
	Infos				zahl	stimme	fläche	
Baden-Württemberg	Mittel	Ja	Spezifisch	Nein	Ja	–	Rechts	Nein
Bayern	Kurz	Ja	Nein	Nein	Nein	Zusätzlich	Links	Nein
Brandenburg	Mittel	Ja	Unspezifisch	Ja	Ja	–	Rechts	Nein
Hessen	Lang	Nein	Nein	Nein	Ja	Kopfzeile	Rechts	Nein
Mecklenburg-Vorpommern	Mittel	Ja	Unspezifisch	Nein	Ja	–	Rechts	Nein
Niedersachsen	Mittel	Ja	Spezifisch	Ja	Ja	Zusätzlich	Rechts	Nein
Nordrhein-Westfalen	Kurz	Ja	Spezifisch	Ja	Nein	–	Rechts	Ja
Rheinland-Pfalz	Lang	Nein	Nein	Nein	Ja	Kopfzeile	Rechts	Nein
Saarland	Lang	Ja	Nein	Nein	Nein	Kopfzeile	–	Nein
Sachsen	Mittel	Ja	Nein	Nein	Ja	–	Rechts	Nein
Sachsen-Anhalt	Mittel	Ja	Unspezifisch	Ja	Ja	–	Rechts	Nein
Schleswig-Holstein	Kurz	Ja	Spezifisch	Nein	Ja	–	Rechts	Ja
Thüringen	Lang	Nein	Nein	Nein	Ja	Kopfzeile	Rechts	Nein

Die Variable *Infos* erfasst, in welchem Detail die korrekte Stimmabgabe auf dem Stimmzettel erläutert wird. Die Kategorisierung der Erläuterungen wurde explorativ im Vergleich der Stimmzettel zueinander vorgenommen. Ein Vergleich der hier verwendeten Kategorien zu den Experimentalkategorien ist deshalb nur bedingt möglich. Die Variablen *Beruf*, *Adresse* und *Alter* beschreiben, ob diese Information auf dem Stimmzettel angegeben ist. In der Variable *Adresse* entspricht eine spezifische Adresse der genauen Adressinformation, während eine unspezifische Adresse beispielsweise nur den Wohnort oder den Ortsteil angibt. *Stimmenzahl* beschreibt die typografische Hervorhebung der abzugebenden Stimmen. *Listenstimme* erfasst, ob die Listenstimme in der Zeile des Parteinamens (Kopfzeile) oder in einer zusätzlichen Zeile (Zusätzlich) abgedruckt wird. Fehlende Werte in dieser Variable bedeuten, dass ein Listenkreuz nicht vorgesehen ist. *Stimmfläche* gibt die Position der Ankreuzmöglichkeiten links oder rechts vom Kandidatennamen an. Der fehlende Wert für das Saarland ergibt sich durch die geschlossene Listenwahl. *Pfeile* erfasst, ob die Spalten für die Stimmabgabe grafisch hervorgehoben sind. Als Beispiel für die Faktoren sei auf den fiktiven Stimmzettel in Abb. 2 hingewiesen

Für die Klassifizierung der kommunalen Stimmzettel wurden ausschließlich die Wahlen zu den Gemeinderäten, nicht aber zu den Kreistagen berücksichtigt. Die Gestaltung der Stimmzettel ist in den Kommunalwahlordnungen festgelegt. In der Praxis können jedoch teils Abweichungen von den Musterstimmzetteln festgestellt werden. Deshalb wurden neben den Musterstimmzetteln auch jeweils konkrete Stimmzettel zufällig ausgewählter Gemeinden kodiert. Es ist jedoch nicht auszuschließen, dass die Stimmzettel in einzelnen Gemeinden Merkmale aufweisen, die von der allgemeinen Länderkodierung abweichen.

Für die Kodierung wurde ein exploratives Kategorienschema verwendet. Dabei wurde in einem ersten Schritt das zu kodierende Material, also die Muster- und Beispielstimmzettel, gesammelt. In einem zweiten Schritt wurden die Merkmale der

Stimmzettel in den drei Kategorien (1) Informationen zum Wahlsystem, (2) Informationen zu den Kandidierenden und (3) Designmerkmale kodiert und auf Grundlage der empirischen Varianz zu Variablen zusammengefasst. Es sei bemerkt, dass in der Übersicht nicht alle Designvarianten dargestellt werden. Es werden lediglich diejenigen Variablen berücksichtigt, die auf Grundlage der explorativen Kodierung als bedeutsam identifiziert und aus diesem Grund im Experimentaldesign übernommen wurden.

Tab. 1 zeigt das Ergebnis der Klassifizierung. Es wurde eine Variable auf der Ebene Informationen zum Wahlsystem erhoben, drei Variablen zu den Kandidatendeninformationen und vier Variablen zu den Designmerkmalen. Die Definition der Variablen und ihre Ausprägungen sind in den Anmerkungen zu Tab. 1 erläutert. Bei den Informationen zum Wahlsystem ist zu bemerken, dass die in Deutschland verwendeten Stimmzettel sich auf Informationen über die korrekte Stimmabgabe beschränken; die zugrundeliegenden wahlsystemischen Mechanismen werden dagegen nicht erläutert. Dabei werden den Wählerinnen durchaus auch potenziell wahlentscheidende Informationen vorenthalten. Beispielhaft sei der fehlende Hinweis auf die Fünf-Prozent-Hürde auf den Stimmzetteln für die Bundestagswahl genannt. Im Sinne des gewählten Forschungsdesigns, das sich eng an der Verwaltungspraxis orientiert, beschränken wir unsere Analyse auf Informationen zur Stimmabgabe.

Insgesamt ist in jedem Faktor Varianz zu beobachten. Dabei können keine offensichtlichen Ländergruppen festgestellt werden, die sich beispielsweise durch ähnliche Regelungen im kommunalen Wahlrecht ergeben könnten. Daraus ist zu schließen, dass es keine *Best-Practice*-Regelungen zur Gestaltung von Stimmzetteln gibt und dass sich die gewählten Gestaltungsmerkmale vor allem durch gewohnte Praxis begründen. Eine empirische Untersuchung der Gestaltungsmerkmale erscheint vor diesem Hintergrund umso wichtiger.

4 Untersuchungsdesign

4.1 Beschreibung des Discrete-Choice-Experimentes

Bei der Gestaltung von Stimmzetteln handelt es sich um ein mehrdimensionales Problem, da Stimmzettel sich in zahlreichen Faktoren unterscheiden. Untersucht werden soll der Einfluss dieser Faktoren auf die Bewertung von Stimmzetteln. Experimentell lassen sich solche mehrdimensionalen Einflüsse am besten in Form eines *Discrete-Choice-Experimentes* (DCE) prüfen (Engst et al. 2020; Hainmueller et al. 2014; Jankowski et al. 2019; Leeper et al. 2020), da hier der Einfluss verschiedener Faktoren zugleich betrachtet werden kann. Bei einem DCE müssen sich Befragte zwischen mehreren Profilen entscheiden, die sich zufällig in verschiedenen Faktoren unterscheiden. Auf diese Weise kann verglichen werden, ob Profile mit bestimmten Ausprägungen systematisch häufiger oder weniger häufig ausgewählt werden.

Im vorliegenden Experiment wurden den Befragten Bilder von je zwei Stimmzetteln zum Vergleich vorgelegt. Die zu vergleichenden Stimmzettel unterscheiden sich in acht Merkmalen, die den Dimensionen *Informationen zum Wahlsystem*, *Information zu den Kandidierenden* und *Designmerkmale* zuzuordnen sind. Die acht

Tab. 2 Übersicht der Varianten des experimentellen Stimulus

Faktor	Ausprägungen
<i>Informationen zum Wahlsystem</i>	
Erläuterung der Stimmabgabe	(1) Kurz (2) Mittel (3) Lang
<i>Informationen zu den Kandidierenden</i>	
Beruf	(1) Vorhanden (2) Nicht vorhanden
Adresse	(1) Vorhanden (2) Nicht vorhanden
Alter	(1) Vorhanden (2) Nicht vorhanden
<i>Designmerkmale</i>	
Hinweis auf Stimmzahl	(1) Im Fließtext (2) Grafisch abgesetzt
Position der Listenstimme	(1) In Kopfzeile (2) Zusätzliche Zeile
Position der Stimmfläche	(1) Links (2) Rechts
Pfeile über den Stimmflächen	(1) Vorhanden (2) Nicht vorhanden

Die Faktoren sind in Abb. 2 dargestellt

Faktoren sind in Tab. 2 dargestellt. Es ist zu bemerken, dass sich die kommunalen Stimmzettel in der Praxis mit Blick auf die Genauigkeit der Adressangabe unterscheiden und entweder die vollständige Kandidatenadresse oder eine unspezifische Wohnortangabe bieten. Im Experiment wurden lediglich die vollständige Adresse oder keine Adresse berücksichtigt, weshalb keine Erkenntnisse über eine mögliche Wählerpräferenz für eine unspezifische Wohnortangabe gewonnen werden können.

Das Experiment unterstellt ein offenes Listenwahlsystem mit drei Stimmen bei uneingeschränkten Kumulier- und Panaschiermöglichkeiten, wie es gegenwärtig bei den Kommunalwahlen in den ostdeutschen Flächenländern und Niedersachsen zur Anwendung kommt. Die Stimmzettel wurden in einem sogenannten „Slider“ angezeigt. Hierbei wurde den Befragten zunächst ein Stimmzettel angezeigt und nach einem Mausklick der zweite. Auf diese Weise konnten die Befragten zwischen den beiden Stimmzetteln hin- und herwechseln und diese direkt miteinander vergleichen.⁴ Abb. 1 zeigt einen Screenshot des Experimentes.

Sieben der acht Faktoren haben zwei Ausprägungen. Die Erläuterung der Stimmabgabe wurde in drei Stufen unterteilt.⁵ Die Ausprägung der Faktoren wurde für jeden Stimmzettel randomisiert. Insgesamt ergeben sich somit 384 ($= 3 \times 2^7$) mögliche Stimmzettel. Dabei verfügen alle Stimmzettel über das gleiche Grunddesign, um die Vergleichbarkeit der Ergebnisse sicherzustellen. Für die nichtvariierenden Layoutmerkmale wurden möglichst solche ausgewählt, die den Standards modernen Grafikdesigns entsprechen. Insbesondere waren bei allen Stimmzetteln die Hintergrundfarbe, das Wappen der Mustergemeinde⁶ sowie die Partei- und Kandidatennamen identisch. Reale Partei- und Kandidatennamen wurden nicht verwendet, um die

⁴ Befragte, die den Slider nicht betätigten, wurden von der Analyse ausgeschlossen.

⁵ Die drei Varianten zur Erläuterung der Stimmabgabe sind im Online-Anhang wiedergegeben.

⁶ Als Wappen der Mustergemeinde wurde das Wappen der baden-württembergischen Kleinstadt Unterschneidheim gewählt. Maßgeblich für die Auswahl des Wappens war der Versuch, ein weitgehend unbekanntes und somit neutrales Wappen zu wählen.

26.06.2019, Gemeinderatswahl Mustergemeinde

Wahl zur Vertretung der Mustergemeinde



Sie haben drei Stimmen: ⊗ ⊗ ⊗
 Sie können die Stimmen verteilen, wie Sie wollen: Alles ist möglich.

Partei A (PA)	Partei B (PB)	Partei C (PC)
<input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/>
<p>Alexander Rössner Informatiker Rudolf-Weber-Platz 127</p>	<p>Gerhard Diers Tischlermeister Erich-Lampe-Weg 9</p>	<p>Susanne Wehting Lehrerin Baldenhainer Straße 56</p>
<p>Albert Thielen Werbefachmann Lehmsol 23</p>	<p>Janine Setzer Diplom-Ingenieurin Cillenstraße 32</p>	<p>Torben Niebach IT-Berater Sollingtor 9</p>
<p>Evelyn Kappes Bankkauffrau Oleanderstraße 2</p>	<p>Olaf Remmers Physiotherapeut Offenbergstraße 98</p>	<p>Nicolas Wiczorek Vertriebsleiter Pussallee 14</p>
<p>Lukas Lohse Unternehmensberater Darmer Esch 84</p>	<p>Ellen Brugger Steuerfachangestellte Emeserweg 24</p>	<p>Julia Siebert Architektin Grobburgring 6</p>
<p>Katharina Pfister Tourismuskauffrau Dükerstraße 35</p>	<p>Christin Dillmann Rechtsanwältin Gelderland-Ring 49</p>	<p>Astrid Zahn Bankkauffrau Lehnerberg 29</p>
<p>Tobias Wilms Landwirt Koblplatzweg 41</p>		<p>Nils Gläser Landschaftsarchitekt Voelmerer Straße 74</p>
<p>Karl Hagen Betriebswirt Kastanienweg 13</p>		

Wahlzettel 1

1

Wenn Sie die beiden Wahlzettel vergleichen: Auf welchen der Wahlzettel treffen die folgenden Aussagen eher zu?

	Wahlzettel 1	Wahlzettel 2
Der Wahlzettel ist informativer.	<input type="radio"/>	<input type="radio"/>
Der Wahlzettel ist leichter verständlich.	<input type="radio"/>	<input type="radio"/>
Der Wahlzettel ist übersichtlicher.	<input type="radio"/>	<input type="radio"/>
Der Wahlzettel gefällt mir insgesamt besser.	<input type="radio"/>	<input type="radio"/>

Abb. 1 Screenshot des Experiments

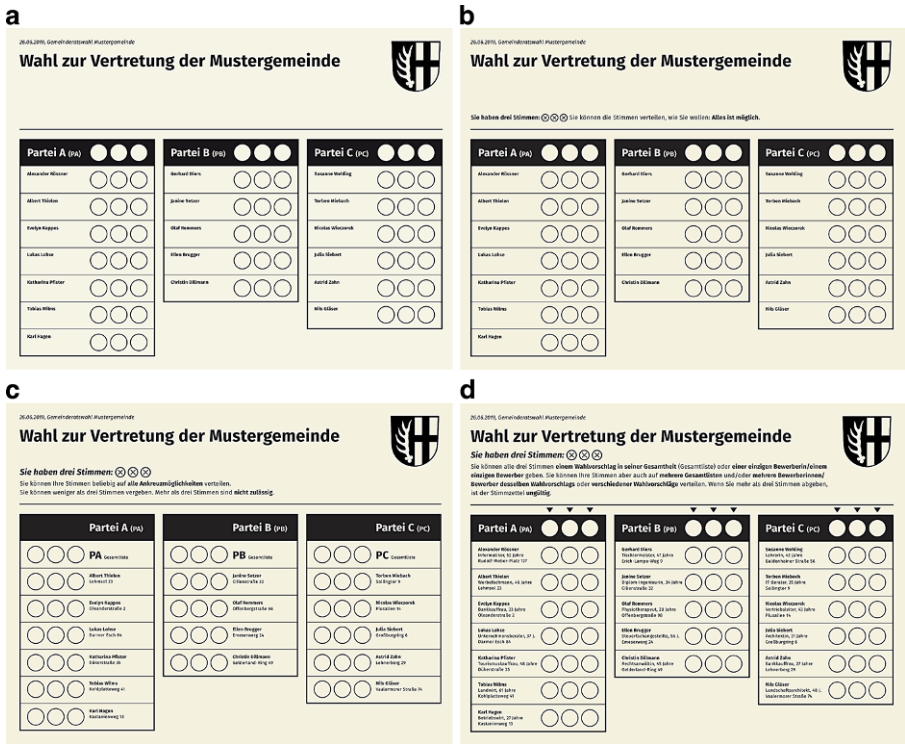


Abb. 2 a Grundvariante des Stimmzettels. b Stimmzettel mit kurzer Erläuterung. c Stimmzettel mit mittellanger Erläuterung. d Stimmzettel mit langer Erläuterung

Bewertungen der Stimmzettel nicht durch die politischen Präferenzen der Befragten zu beeinflussen. Die Grundvariante des Stimmzettels ist in Abb. 2a dargestellt.

Einen Eindruck der unterschiedlichen Varianten des Experimentaldesigns bietet Abb. 2.7 Abb. 2b zeigt den Stimmzettel mit der kürzesten Erläuterung der Stimmabgabe, die den Erläuterungen auf dem Stimmzettel des Landes Bremen nachempfunden sind. Ansonsten unterscheidet sich der Stimmzettel von der Grundvariante nicht.⁸ Stimmzettel 2c zeigt die Variante mit der mittellangen Erläuterung. Zudem sind die Stimmflächen links von den Kandidatennamen angeordnet und für das Listenkreuz ist eine zusätzliche Zeile vorgesehen. Abb. 2d stellt einen Stimmzettel mit der ausführlichen Erläuterung dar, der alle Kandidateninformationen enthält. Weiterhin zeigen Pfeile an, wo die Kreuze gesetzt werden können. Die Abb. 2b–d stellen

⁷ Die Stimmzettel wurden mit der Layout-Software *QuarkXPress* (Version 2018, Quark Software Inc., Denver, Colorado, USA) erstellt. Die Bilddateien sind auf Anfrage von den Autoren erhältlich.

⁸ Mit Blick auf die Faktoren in Tab. 2 setzt sich der Stimmzettel in Abb. 2b als eine Kombination der folgenden Faktorausprägungen zusammen: Erläuterung der Stimmabgabe (Kurz), Beruf (Nicht vorhanden), Adresse (Nicht vorhanden), Alter (Nicht vorhanden), Hinweis auf Stimmzahl (Im Fließtext), Position der Listenstimme (In Kopfzeile), Position der Stimmfläche (Rechts), Pfeile über den Stimmflächen (Nicht vorhanden).

somit drei der 384 möglichen Stimmzettel dar, wobei sie alle in Tab. 2 genannten Faktorausprägungen beinhalten.

Das Experiment wurde als Online-Survey im Zeitraum Ende Juni bis Anfang Juli 2019 mit knapp 1250 deutschen Befragten durchgeführt. Das im Hinblick auf Geschlecht, Alter und Bundesland quotierte Sample wurde von der Firma respondi bereitgestellt. Eine Beschreibung des Samples findet sich im Online-Anhang. Wenngleich das Sample quotiert ist, lassen sich auch bei Online-Befragungen gewisse Verzerrungen nicht vermeiden. Beispielsweise weist das Sample einen überdurchschnittlich hohen Anteil an Abiturientinnen und Abiturienten auf. Das Experiment wurde in der Umfrage-Software *Qualtrics* implementiert.

4.2 Abhängige Variablen

Allen Befragten wurden zweimal zwei zufällig ausgewählte Stimmzettel zur Bewertung vorlegt. Bei beiden Paarvergleichen sollten die Respondentinnen angeben, welchen Stimmzettel sie als (1) übersichtlicher, (2) verständlicher, (3) informativer und (4) insgesamt besser bewerten. Auf diese Weise kann der Effekt der Stimmzettelgestaltung auf unterschiedliche Bewertungsdimensionen betrachtet werden. Im Fokus der Analyse steht die allgemeine Bewertung. Diese Frage scheint am besten geeignet, um einen gelungenen Stimmzettel unter Abwägung der verschiedenen Faktoren auszumachen. Die Frage nach dem Informationsgehalt des Stimmzettels dagegen ist sowohl substanziell interessant, zugleich kann sie aber auch zum Zwecke der Manipulationsprüfung dienen. Manipulationsprüfungen werden in der Teilnehmerforschung eingesetzt, um sicherzustellen, dass ein Treatment von den Teilnehmerinnen tatsächlich wahrgenommen wird (Kane and Barabas 2019). Beispielsweise sollte ein Stimmzettel ohne die Altersangabe der Kandidierenden nicht als informativer wahrgenommen werden als ein Stimmzettel mit dieser Angabe.

Zugleich soll nicht unbemerkt bleiben, dass die gleichzeitige Abfrage mehrerer Bewertungsdimensionen zu Ausstrahlungseffekten zwischen den Fragen führen kann. So liegt es beispielsweise nahe, dass die Befragten bei der Gesamtbewertung der Stimmzettel aufgrund der weiteren Fragen insbesondere den Informationsgehalt, die Verständlichkeit und die Übersichtlichkeit der Stimmzettel berücksichtigt haben. Solche Effekte hätten durch die Abfrage der Teildimensionen auf unterschiedlichen Bildschirmen verringert werden können, was die Befragungszeit allerdings deutlich verlängert hätte.

4.3 Berechnung der Effekte

Wir folgen den Empfehlungen von Leeper et al. (2020) und analysieren das DCE mithilfe von Marginal Means. Marginal Means repräsentieren die durchschnittliche Wahrscheinlichkeit, mit der ein Stimmzettel ausgewählt wurde, der eine bestimmte Faktorausprägung enthält. Spielt eine Faktorausprägung bei der Bewertung der Stimmzettel keine Rolle, dann sollte sie gleich häufig in den ausgewählten und in den nicht ausgewählten Profilen auftauchen. Eine solche Faktorausprägung hätte einen Marginal Mean von 0,5. Wird eine Faktorausprägung von den Befragten bevorzugt, dann sollten Stimmzettel mit dieser Ausprägung überdurchschnittlich

häufig ausgewählt werden. Faktorausprägungen mit einem positiven Einfluss haben einen Marginal Mean über 0,5, negativ bewertete Faktorausprägungen einen Marginal Mean unter 0,5. Da alle Befragten zwei Paarvergleiche vornehmen, korrigieren wir die Standardfehler mit Blick auf die Clustering innerhalb der Befragten.

5 Analyse

5.1 Gesamteindruck

Wir konzentrieren uns zunächst auf die Auswertung des Gesamteindrucks der Stimmzettel. Die Marginal Means zusammen mit den 95 %-Konfidenzintervallen sind in Abb. 3 dargestellt. Die Effekte sind wie folgt zu interpretieren: Je größer der Wert einer bestimmten Faktorausprägung, desto wichtiger ist dieser Faktor für die Auswahl eines Stimmzettels und desto stärker wird diese Ausprägung durch die Befragten favorisiert und umgekehrt. So zeigt sich bei den Informationen zum Wahlsystem, dass die mittellange Erläuterung der Stimmabgabe deutlich positiver bewertet wird als die kurze oder die lange Erläuterung. Beispielsweise liegt die Auswahlwahrscheinlichkeit eines Stimmzettels mit einer mittellangen Erläuterung bei rund 0,55. Wenn also diese Ausprägung bei einem der zu vergleichenden Stimm-

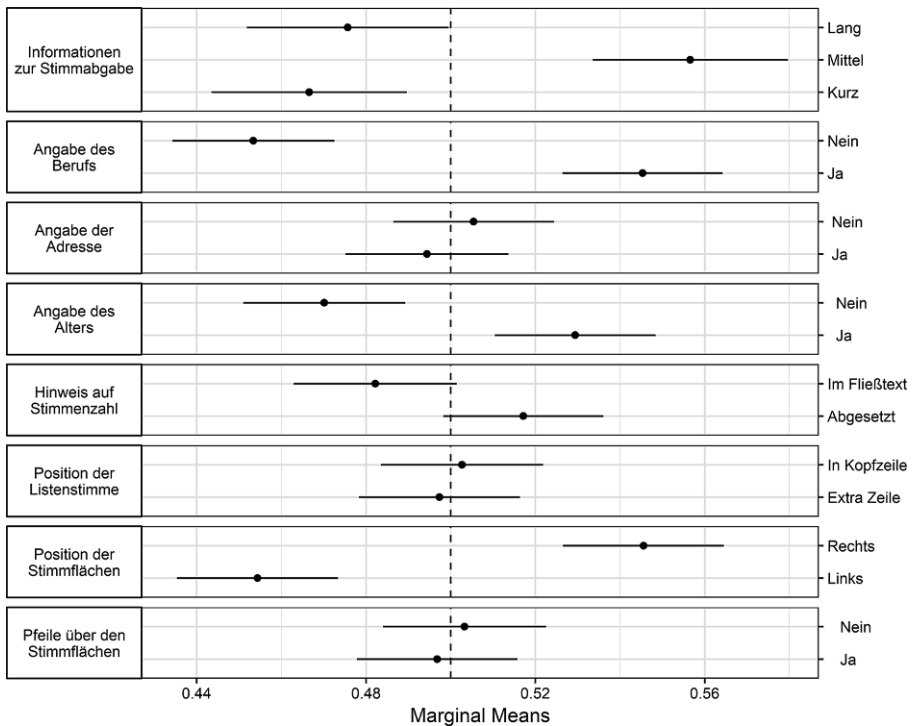


Abb. 3 Marginal Means der Faktoren für den Gesamteindruck

zettel vorliegt, dann beträgt die vorhergesagte Wahrscheinlichkeit für die Auswahl dieses Stimmzettels bei etwa 55 %. Die Werte für die kurze und die lange Erläuterung liegen dagegen unterhalb der 50 %-Schwelle. Trotz der häufig festgestellten Wissenslücken von Wählerinnen über Wahlsysteme, scheinen die Wählerinnen selbst also kein übermäßiges Interesse an detailreichen Erläuterungen zu haben.

Bei den Kandidateninformationen lassen sich klare Unterschiede feststellen. So zeigt sich, dass die Nennung des Berufs eindeutig gewünscht wird. Der Anstieg in der Auswahlwahrscheinlichkeit ist vergleichbar mit dem Anstieg, der bei der mittellangen Erläuterung des Wahlsystems zu beobachten ist. Dieser starke Effekt deutet an, dass der Beruf als entscheidungsrelevantes Merkmal angesehen wird. Dies deckt sich mit den Befunden von McDermott (2005) und Mechtel (2014), die zeigen, dass der Beruf eine wichtige Entscheidungsheuristik darstellt. Unsere Ergebnisse ergänzen diese Forschung dahingehend, dass Informationen über den Beruf der Kandidierenden von den Wählerinnen nicht nur genutzt, sondern auch aktiv gewünscht werden. Auch die Angabe des Alters wird positiv bewertet, wenngleich nicht so stark wie die Berufsangabe.

Sichtlich anders liegen die Befunde für die Angabe der Kandidatenadressen. Hier zeigen sich keine signifikanten Unterschiede zwischen den Marginal Means. Die Punktschätzer deuten gar eine Präferenz gegen die Nennung der Adresse an. Diese Beobachtung überrascht vor dem Hintergrund einer umfangreichen Literatur, die dem Wohnort ebenfalls eine wichtige heuristische Funktion bei der Wahlentscheidung zuschreibt (Campbell und Cowley 2014; Campbell et al. 2019; Jankowski 2016; Tavits 2010).

Der Unterschied zwischen der bestehenden Forschung und den Ergebnissen des vorliegenden Experiments kann mehrere Gründe haben. Erstens wird die Nennung der Adresse von vielen Wählerinnen womöglich als zu sensible Information angesehen, obwohl diese Angabe auf Stimmzetteln durchaus üblich ist. Es mag daher einen Unterschied machen, ob die spezifische Adresse oder eine unspezifische Wohnortsangabe auf dem Stimmzettel abgedruckt wird. Eine Reihe von offenen Antworten in einem Kommentarfeld am Ende des Surveys deutet darauf hin, dass die konkrete Adresse von den Teilnehmerinnen tatsächlich als zu sensibel wahrgenommen wird.⁹ Zweitens ist nicht auszuschließen, dass die Wählerinnen die Angabe von Adressen zwar prinzipiell als problematisch bewerten, von dieser Information aber dennoch Gebrauch machen. In diesem Falle stünden die Befunde der bisherigen Forschung und die hier vorgestellten Ergebnisse nicht im Widerspruch. Drittens muss bemerkt werden, dass die Befunde zur Adressennennung eventuell eine geringere externe Validität aufweisen als andere Experimentalfaktoren. Da für das Experimentaldesign hypothetische Adressen verwendet wurden, können die Befragten anders als bei

⁹ Exemplarisch: „Ich fand es nicht so gut, dass die Adressen der fiktiven Kandidaten angegeben waren, Alter ist ok, Beruf wäre auch gut“, „Auch wenn es nur fiktive Wahlzettel waren, die Angabe der Adresse der Kandidaten finde ich nicht gut – wie sähe es hier denn mit dem Datenschutz aus? Beruf und Alter reichen m. E. vollkommen aus“, „Mich hat gestört, das auf der einen Sorte der Wahlzettel, die kompletten Adressen gestanden haben, das würde ich auf keinen Fall machen“, „Ich finde, auf Wahlzettel gehören leider keine Adressdaten, damit die aggressiven Menschen zumindest kurz ihr nicht vorhandenes Hirn nutzen müssten, um Drohungen zu versenden“.

26.06.2019, Gemeinderatswahl Mustergemeinde

Wahl zur Vertretung der Mustergemeinde



Sie haben drei Stimmen: ⊗ ⊗ ⊗

Sie können Ihre Stimmen beliebig auf **alle Ankreuzmöglichkeiten** verteilen.
 Sie können weniger als drei Stimmen vergeben. Mehr als drei Stimmen sind **nicht zulässig**.

Partei A (PA)	Partei B (PB)	Partei C (PC)
Alexander Rössner Informatiker, 52 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>	Gerhard Diers Tischlermeister, 47 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>	Susanne Wehling Lehrerin, 42 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>
Albert Thielen Werbefachmann, 46 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>	Janine Setzer Diplom-ingenieurin, 34 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>	Torben Niebach IT-Berater, 35 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>
Evelyn Kappes Bankkauffrau, 33 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>	Olaf Remmers Physiotherapeut, 28 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>	Nicolas Wiczorek Vertriebsleiter, 43 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>
Lukas Lohse Unternehmensberater, 37 J. <input type="radio"/> <input type="radio"/> <input type="radio"/>	Ellen Brugger Steuerfachangestellte, 56 J. <input type="radio"/> <input type="radio"/> <input type="radio"/>	Julia Siebert Architektin, 31 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>
Katharina Pfister Tourismuskaufrfrau, 48 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>	Christin Dillmann Rechtsanwältin, 45 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>	Astrid Zahn Bankkauffrau, 27 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>
Tobias Wilms Landwirt, 61 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>		Nils Gläser Landschaftsarchitekt, 48 J. <input type="radio"/> <input type="radio"/> <input type="radio"/>
Karl Hagen Betriebswirt, 27 Jahre <input type="radio"/> <input type="radio"/> <input type="radio"/>		

Abb. 4 Stimmzettel mit den meist präferierten Merkmalskombinationen

den Angaben zu Beruf und Alter aus der Adresse keine Informationen über die Kandidierenden ableiten.

Abschließend widmen wir uns den Designelementen des Stimmzettels. Unsere Ergebnisse zeigen, dass weder die Darstellung der Listenstimme noch die Hinweispeile einen Einfluss auf die Bewertung der Stimmzettel haben. Die abgesetzten Ankreuzhinweise zeigen einen positiven Effekt gegenüber der Variante, bei der die Ankreuzhinweise in die Erläuterung des Wahlsystems eingebettet sind. Die Effektstärke ist insgesamt aber eher gering, sodass dieser Faktor kein entscheidender Aspekt bei der Stimmzettelgestaltung ist.

Ein deutlicher Effekt zeigt sich dagegen bei der Positionierung der Stimmflächen. Die Anordnung mit der Ankreuzoptionen rechts von den Kandidatennamen wird von den TeilnehmerInnen klar bevorzugt. Dies kann einerseits auf die Leserichtung von links nach rechts zurückgeführt werden, weshalb es für die TeilnehmerInnen natürlicher wirken mag, zunächst die Kandidatennamen und dann die Stimmflächen zu betrachten. Eine weitere Möglichkeit ist, dass sich in diesem Effekt die Häufigkeit von Rechtshändigkeit ausdrückt, die Mandal und Dutta (2001) mit rund 90% beziffern. Denkbar wäre, dass RechtshänderInnen die rechtsseitige Positionierung der Stimmflächen bevorzugen, während LinkshänderInnen eher die gegenteilige Anordnung auswählen. So legen die Ergebnisse von Kim et al. (2015) beispielsweise nahe, dass die Anordnung der Ankreuzoption einen Effekt auf das Wahlverhalten von LinkshänderInnen und RechtshänderInnen hat. Da wir die dominante Hand der Teil-

nehmenden nicht abgefragt haben, können wir zwischen diesen beiden Mechanismen nicht unterscheiden. Zumindest deuten vereinzelte Antworten im Kommentarfeld am Ende der Umfrage darauf hin, dass Befragte die linke Anordnung der Stimmflächen aufgrund der Leserichtung ablehnen.¹⁰ In zukünftiger Forschung sollte diese Frage weiter verfolgt und zwischen den möglichen Erklärungen unterschieden werden, da sie mit unterschiedlichen Handlungsempfehlungen einhergehen. Falls sich der Befund durch die typische Leserichtung erklärt, dann würde dieses Ergebnis für die Gestaltung von Stimmzetteln sprechen, die den Lesegewohnheiten in unterschiedlichen Ländern sprechen. Im Falle von Effekten der dominanten Hand könnten gar Stimmzettel für Linkshänderinnen und Rechtshänderinnen angezeigt sein.

Zusammenfassend kann jedoch festgehalten werden, dass insbesondere die Informationen zur Stimmabgabe und zu den Kandidierenden bei der Bewertung der Stimmzettel eine wichtige Rolle spielen. Mit Blick auf die Gestaltung der Stimmzettel ist die Positionierung der Stimmflächen entscheidend. Basierend auf diesen Ergebnissen haben wir in Abb. 4 den am besten gelungenen Stimmzettel dargestellt – also den Stimmzettel, bei dem die Punktschätzer der Marginal Means am größten sind.

5.2 Informationsgehalt, Übersichtlichkeit und Verständlichkeit

Neben dem Gesamteindruck wurde auch die wahrgenommene *Übersichtlichkeit*, *Verständlichkeit* und der *Informationsgehalt* der Stimmzettel abgefragt. Die Ergebnisse dieser Bewertungskriterien sind in Abb. 5 dargestellt.

Zunächst kann beobachtet werden, dass die Befragten den unterschiedlichen Informationsgehalt der Stimmzettel korrekt wahrgenommen haben. Dies unterstreicht zum einen die interne Validität des Experimentes. Zum anderen ergeben sich aus dem Vergleich der Bewertungen des Informationsgehalts mit den anderen Bewertungsdimensionen bemerkenswerte Folgen. So zeigt sich insbesondere, dass die Befragten die lange Erläuterung der Stimmabgabe als am informativsten wahrnehmen. Offensichtlich legen die Befragten aber keinen Wert auf derart detailreiche Erläuterungen. Dies zeigt sich sowohl beim Gesamteindruck und noch deutlicher bei der Übersichtlichkeit und der Verständlichkeit der Stimmzettel. Bei den beiden letztgenannten Aspekten wird die lange Erklärung am schlechtesten bewertet. Offenbar sind zu ausführliche Erläuterungen also problematisch, da die Befragten sich nicht mehr kompetent genug fühlen, um die Informationen aufzunehmen. Dieser Befund ist im Kontext der Debatte um die Verständlichkeit von Wahlsystemen zu sehen (Jankowski et al. 2020; Schmitt-Beck 1993). So stehen komplexe Wahlsysteme im Verdacht, Wählerinnen von der Wahlteilnahme abzuschrecken und den Anteil ungültiger Stimmen zu erhöhen (Nyhuis 2014; Schäfer und Schoen 2013). Unsere Befunde weisen in die gleiche Richtung. Zu umfassende Erläuterungen sind sichtbar unerwünscht und führen zu Unverständnis bei den Wählerinnen und Wählern.

Auch bei der Adressangabe und dem wahrgenommenen Informationsgehalt lassen sich klare Unterschiede zu den anderen Bewertungsdimensionen feststellen. So

¹⁰ Ein Teilnehmenden schrieb beispielsweise: „In Deutschland lesen wir von links nach rechts. Ich weiß nicht, warum sie das ändern wollen“.

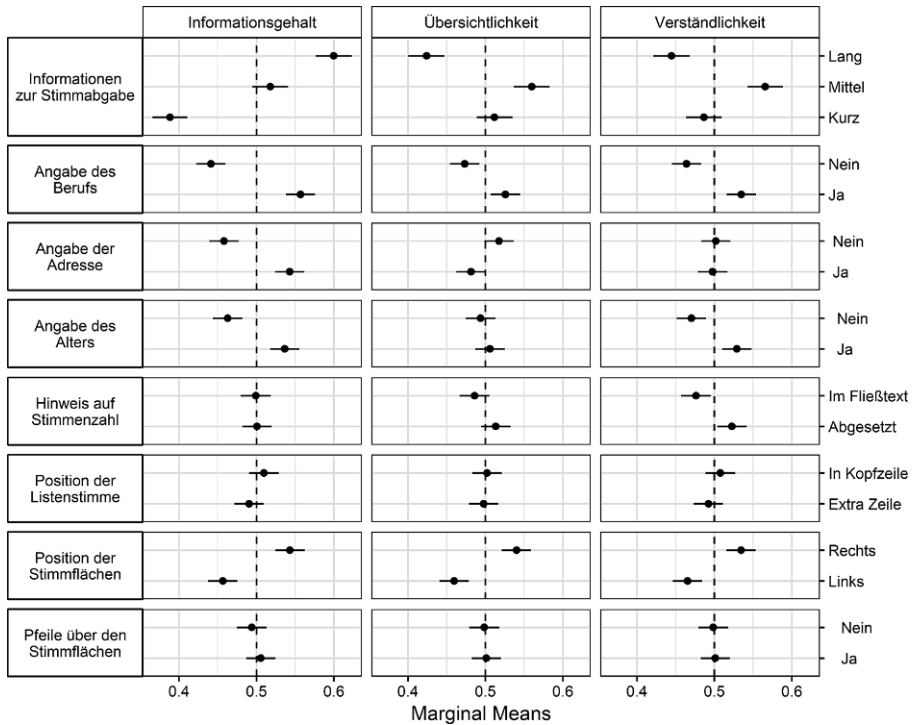


Abb. 5 Marginal Means der Faktoren auf den Dimensionen *Übersichtlichkeit*, *Verständlichkeit* und *Informationsgehalt*

wird die Adressangabe als informativ wahrgenommen, hat aber keinen Einfluss auf die Verständlichkeit und sogar einen leicht negativen Effekt auf die allgemeine Bewertung und die Übersichtlichkeit.

Auch die Ausrichtung der Stimmflächen wurde als unterschiedlich informativ wahrgenommen. Diese Beobachtung überrascht auf den ersten Blick, da es sich bei diesem Faktor lediglich um ein Layout-Element handelt, durch das sich der objektive Informationsgehalt des Stimmzettels nicht ändert. Eine mögliche Erklärung für diesen Effekt ist, dass Befragte sich bei objektiv gleichem Informationsgehalt für den übersichtlicheren Stimmzettel auf der Dimension Informationsgehalt entscheiden.¹¹ Diese Vermutung wird durch die Daten gestützt. Berechnet man die Effekte nur für Vergleiche zwischen Stimmzetteln mit objektiv gleichem Informationsgehalt, dann zeigt sich der Effekt der Stimmflächenposition auf den wahrgenommenen Informationsgehalt. Betrachtet man hingegen nur Paarvergleiche mit objektiv unterschiedlichem Informationsgehalt, dann findet sich ein solcher Effekt nicht (ohne Abbildung). Ebenso zu berücksichtigen sind mögliche Ausstrahlungseffekte zwischen den Fragen. Da alle Bewertungsdimensionen gleichzeitig erhoben wurden, ist

¹¹ Es war im Experiment nicht möglich, keinen der beiden Stimmzettel zu bewerten. Wenn der objektive Informationsgehalt der beiden Stimmzettel identisch war, mussten die Befragten also trotzdem eine Entscheidung über den wahrgenommenen Informationsgehalt treffen.

davon auszugehen, dass die Bewertungen der Einzeldimensionen sich gegenseitig beeinflussen. Auch so könnte sich erklären, weshalb die Position der Ankreuzflächen einen Effekt auf den Informationsgehalt hat.

5.3 Zusammenhang zwischen den Teildimensionen und dem Gesamteindruck

In diesem Abschnitt untersuchen wir, wie die drei abgefragten Teildimensionen *Informationsgehalt*, *Verständlichkeit* und *Übersichtlichkeit* mit dem Gesamteindruck des Stimmzettels zusammenhängen. Eine solche Analyse kann Aufschluss darüber geben, ob einer der drei Teilaspekte eine besondere Rolle für den Gesamteindruck spielt.

Zunächst kann festgestellt werden, dass die Mehrheit der Befragten den Stimmzettel mit dem besseren Gesamteindruck auch auf allen drei Teildimensionen besser bewertet hat. Dies gilt in 66,4% der Fälle. In 25,3% der Fälle gilt das für zwei der drei Teildimensionen. In immerhin 7,8% der Fälle wurde nur eine der drei Teildimensionen am besten eingeschätzt und trotzdem der Gesamteindruck als besser angesehen. Fälle, in denen zwar der Gesamteindruck besser ist, aber keine der Teildimensionen besser bewertet wurden, treten mit 0,4% praktisch nicht auf.

Um zu untersuchen, welche der Teildimensionen den stärksten Einfluss auf die Gesamtbewertung hat, prüfen wir zunächst, wie häufig ein Stimmzettel insgesamt als besser und zugleich als verständlicher, informativer oder übersichtlicher bewertet wurde. Diese Zusammenhänge sind in Tab. 3 dargestellt. Sie legen nahe, dass Verständlichkeit der wichtigste Faktor für die Gesamtbewertung ist. In über 90% der Fälle werden Stimmzettel, die als verständlicher eingeschätzt werden, auch insgesamt als besser bewertet. Hingegen wird in nur knapp 81% der Fälle der informativere Stimmzettel auch als der bessere eingeschätzt. Die Übersichtlichkeit liegt mit 86% zwischen den beiden Teildimensionen.

Um diesen Zusammenhang systematischer zu betrachten, wurde der Einfluss der Teildimensionen auf die Gesamtbewertung geschätzt. Dabei berücksichtigt das Modell alle möglichen Interaktionseffekte, dass also die Teildimensionen einen gemeinsamen Einfluss auf die Gesamtbewertung ausüben. Auf diese Weise ergeben sich acht ($= 2^3$) mögliche Kombinationen der drei Teildimensionen. Abb. 6 zeigt die Ergebnisse dieser Analyse. Auf der y-Achse ist die Wahrscheinlichkeit angegeben, dass ein Stimmzettel insgesamt als besser bewertet wird. Auf der x-Achse sind die acht möglichen Kombinationen dargestellt. Wenig überraschend zeigt sich zunächst, dass die positive/negative Bewertung aller Teildimensionen eindeutig mit einer positiven/negativen Gesamtbewertung einhergeht. Interessanter sind Fälle, bei denen eine oder zwei Teildimensionen als schlechter bewertet wurden. Insgesamt macht es

Tab. 3 Zusammenhang zwischen Gesamtbewertung und den drei Teildimensionen

	Der Stimmzettel ist ...		
	Verständlicher (%)	Informativer (%)	Übersichtlicher (%)
Der Gesamteindruck ist schlechter	9,5	19,1	13,6
Der Gesamteindruck ist besser	90,5	80,9	86,4

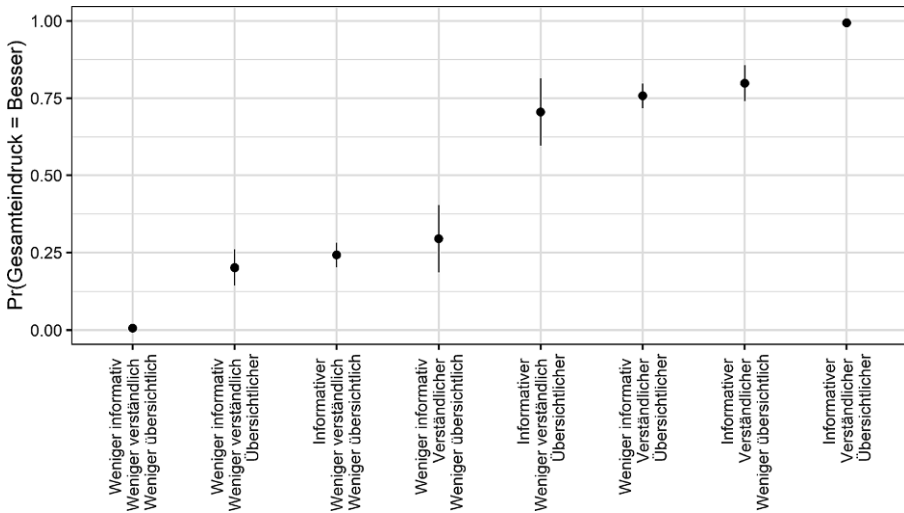


Abb. 6 Einfluss der Teildimensionen auf den Gesamteindruck

zwar einen eher geringen Unterschied, welche Teildimension/en als schlechter bewertet wurden. Dennoch hat die Übersichtlichkeit einen etwas geringeren Einfluss auf die positive Gesamtbewertung. Folglich zeigen auch diese Ergebnisse, dass die Verständlichkeit den stärksten Einfluss auf die Gesamtbewertung der Stimmzettel hat.

5.4 Effektheterogenität

In den letzten drei Abschnitten wurden allgemeine Präferenzen der Teilnehmerinnen für bestimmte Stimmzettelmerkmale herausgearbeitet. In diesem letzten Abschnitt der Analyse wird nun die Frage gestellt, ob bestimmte gesellschaftliche Gruppen identifiziert werden können, die systematisch anders gelagerte Präferenzen im Hinblick auf die Gestaltung von Stimmzetteln haben. So ist beispielsweise denkbar, dass der Bildungsgrad der Teilnehmerinnen mit unterschiedlichen Erwartungen an die Gestaltung der Stimmzettel einhergeht. Die Auswertung hat einen explorativen Charakter, sodass eine Vielzahl von möglichen und üblichen Gruppierungsfaktoren als Moderatoren geprüft wird.

Effektheterogenität liegt dann vor, wenn sich die Marginal Means zwischen Gruppen mit verschiedenen Ausprägungen eines bestimmten Gruppierungsmerkmals unterscheiden. Zur Prüfung heterogener Effekte kann die Gruppierungsvariable bei der Berechnung der Marginal Means einbezogen werden. Dabei wird ein Modell geschätzt, welches die Experimentalfaktoren und die Gruppenzugehörigkeit interagiert. Mittels eines F-Tests lässt sich prüfen, ob ein Modell mit der Gruppierungsvariable die Daten besser beschreibt als ein Modell ohne (Leeper et al. 2020).

Tab. 4 Test für Gruppenunterschiede

Moderator	<i>p</i> -Wert
Wahlabsicht: B'90/Die Grünen	0,079
Wahlabsicht: SPD	0,276
Abitur (Ja/Nein)	0,332
Wahlabsicht: AfD	0,332
Alter	0,397
Parteiidentifikation (Ja/Nein)	0,528
Wahlabsicht: FDP	0,659
Politisches Interesse	0,672
Wahlabsicht: CDU/CSU	0,721
Wahlabsicht: Die Linke	0,788
Geschlecht	0,988
Politisches Wissen	0,996

Die Wahlabsicht vergleicht Befragte mit einer bestimmten Wahlabsicht mit allen anderen Teilnehmerinnen. *Alter* wurde in vier Kategorien eingeteilt: 18–29 Jahre, 30–45 Jahre, 46–60 Jahre und älter als 60 Jahre. *Parteiidentifikation* misst, ob eine Parteiidentifikation vorliegt. *Politisches Interesse* wurde auf Basis der Selbsteinstufung als binär kategorisiert (hoch/niedrig). *Politisches Wissen* misst die Anzahl richtiger Antworten auf drei politische Wissensfragen

Die Ergebnisse des F-Tests sind in Tab. 4 für zwölf potenzielle Moderatoren wiedergegeben.¹² Für keine der betrachteten Gruppen lässt sich ein statistisch signifikanter Unterschied zwischen den Modellen feststellen. Immerhin liegt der *p*-Wert für den Vergleich von Anhängerinnen der Grünen mit den anderen Befragten relativ nahe am üblichen Signifikanzniveau von 0,05 (vgl. Abb. A11 im Online-Anhang). Unterschiede zeigen sich hier vor allem bei der Länge der Erläuterung der Stimmabgabe. Im Gegensatz zu den anderen Befragten bewerten die Anhängerinnen der Grünen eine lange Beschreibung der Stimmabgabe positiver, eine kurze Beschreibung dagegen negativer. Dieser Befund erscheint vor dem Hintergrund des sozioökonomischen Profils der grünen Wählerschaft plausibel (Dolezal 2010). Dennoch kann festgehalten werden, dass die mittellange Erläuterung auch unter den Anhängerinnen der Grünen die höchste Zustimmung erfährt.¹³

Insgesamt legen die Befunde der Subgruppenanalysen somit nahe, dass es ein hohes Maß an Übereinstimmung in der Bevölkerung über einen gelungenen Stimm-

¹² Die Marginal Means für die Subgruppen sind in den Abb. A3 bis A14 im Online-Anhang dargestellt.

¹³ Die Vertrautheit mit Merkmalen der fiktiven Stimmzettel aufgrund der kommunalen Wahlpraxis könnte ebenfalls einen Einfluss auf die Bewertung der Stimmzettel haben. Trotz einer allgemeinen Angleichung des kommunalen Wahlrechts gibt es nach wie vor Unterschiede in den kommunalen Wahlrechtsregelungen, insbesondere in Nordrhein-Westfalen, dem Saarland und Schleswig-Holstein. Zudem sind auch die Stimmzettel durch nennenswerte Unterschiede gekennzeichnet. Beispielsweise stellt Bayern als einziges Bundesland die Stimmfläche auf der linken statt auf der rechten Seite des Namens dar (siehe Tab. 1). Um entsprechende Effekte auszuschließen, haben wir Subgruppenanalysen auf der Ebene der Bundesländer durchgeführt (nicht dargestellt). Dabei konnten keine Präferenzunterschiede zwischen den Befragten in den verschiedenen Bundesländern festgestellt werden.

zettel gibt. Es ist folglich möglich, einen Stimmzettel zu gestalten, der allgemein als gut geeignet wahrgenommen wird.

6 Schlussfolgerungen

Auf Grundlage eines Discrete-Choice-Experimentes wurde in diesem Beitrag dargelegt, wie Stimmzettel nutzerfreundlicher gestaltet werden können. Zu diesem Zweck wurde zwischen drei Dimensionen der Gestaltung von Stimmzetteln unterschieden: Informationen zum Wahlsystem, Informationen zu den Kandidierenden und Designmerkmale. Unsere Ergebnisse zeigen, dass Befragte sowohl einer kurzen als auch einer ausführlichen Erläuterung der Stimmabgabe kritisch gegenüberstehen. Zwar empfinden Befragte eine detailreiche Erläuterung durchaus als informativ. Diese wird jedoch insgesamt negativer bewertet, da sie als unübersichtlich und unverständlich wahrgenommen wird. Die praktische Herausforderung bei der Gestaltung von Stimmzetteln besteht somit darin, alle für den Wahlakt notwendigen Informationen möglichst kompakt darzustellen.

Darüber hinaus bewerten die Teilnehmerinnen die Alters- und vor allem die Berufsangabe der Kandidierenden als positiv. Grundlegende Kandidateninformationen werden von den Wählerinnen also eindeutig gewünscht. Stimmzettel mit den Adressen der Kandidierenden wurden von den Teilnehmerinnen dagegen nicht systematisch bevorzugt. Es lässt sich sogar eine leicht ablehnende Tendenz feststellen. Auf den ersten Blick überrascht dieser Befund vor dem Hintergrund der bekannten Zusammenhänge zwischen Wohnort und Wahlerfolg. Er lässt sich jedoch vermutlich auf Bedenken um die Privatsphäre der Bewerberinnen und Bewerber zurückführen. An diesem Punkt ist weitere Forschung wünschenswert. So stellt sich zum einen die Frage, ob entsprechende Vorbehalte bei bestimmten Bevölkerungsgruppen besonders ausgeprägt sind und inwiefern die Debatte um zunehmende Angriffe auf deutsche Kommunalpolitikerinnen hier zu einem Einstellungswandel geführt hat. Zum anderen könnte die Präzision der Wohnortsangabe in einem Folgeexperiment variiert werden, beispielsweise durch Beschränkung auf den Wohnstadtteil, die zu einer besseren Bewertung der Adressangabe führen mag.

Mit Blick auf die Designelemente lässt sich festhalten, dass die meisten betrachteten Merkmale einen geringen Einfluss auf die Bewertung haben. Einzig die Ausrichtung der Stimmflächen hat einen Effekt – und zudem einen ganz erheblichen. Gerade vor dem Hintergrund der empirischen Vielfalt deutscher Stimmzettel mahnen die Ergebnisse also durchaus zu einer sorgfältigen Gestaltung von Stimmzetteln und weisen über den reinen Informationsgehalt der Stimmzettel hinaus. Insbesondere muss berücksichtigt werden, dass eine Reihe von denkbaren Designmerkmalen im vorliegenden Experiment nicht variiert wurde, um ein klares und modernes Grunddesign sicherzustellen.

Weiterhin konnte festgestellt werden, dass die Verständlichkeit das wichtigste Kriterium für die Bewertung der Stimmzettel ist, während der wahrgenommene Informationsgehalt eine eher untergeordnete Rolle spielt. Schließlich konnte gezeigt werden, dass die Präferenzen für die Gestaltung der Stimmzettel über verschiedene

Bevölkerungsgruppen hinweg stabil sind – es ist somit möglich, einen Stimmzettel zu entwerfen, der allgemein als gelungen angesehen wird.

Die Ergebnisse erlauben die Einordnung der Gestaltungsmerkmale, die gegenwärtig von den deutschen Länderverwaltungen für die Durchführung von Kommunalwahlen genutzt werden (siehe Tab. 1). Es gelingt den Länderverwaltungen unterschiedlich gut, Stimmzettel zu entwerfen, die von unseren Befragten als geeignet erachtet würden. Ohne die Ergebnisse zu überzeichnen, kann doch zumindest ein grober Zusammenhang in den Daten erkannt werden. So weisen die Stimmzettel in den neuen Bundesländern in der Tendenz bessere Merkmalskombinationen auf als die Stimmzettel in den alten Bundesländern. Eine naheliegende Interpretation für diesen Befund ist, dass Länderverwaltungen Wahlrechtsreformen zum Anlass nehmen, um die Gestaltung ihrer Stimmzettel zu hinterfragen und zu optimieren. Umgekehrt kann dieses Ergebnis auch als Appell an Bundesländer mit einem vergleichsweise stabilen Wahlsystem dienen, ihre Musterstimmzettel in regelmäßigen Abständen zu überdenken und zu modernisieren.

Darüber hinaus sei angemerkt, dass selbst Stimmzettel mit einer gelungenen Merkmalskombination einem modernen Grafikdesigns häufig nicht entsprechen. Zwar wurden viele grafische Elemente in der vorliegenden Arbeit vorausgesetzt, sodass keine Erkenntnisse über die Präferenzen der Wählerinnen abgeleitet werden können. Dennoch liegt es nahe, dass die Verletzung von Standards eines modernen Grafikdesigns der Nutzerfreundlichkeit von Stimmzetteln abträglich ist.¹⁴ Auch an dieser Stelle könnten und sollten die Länderverwaltungen für Nachbesserungen sorgen – etwa durch Nutzung moderner Softwarelösungen und/oder durch Vergabe von Gestaltungsaufträgen an spezialisierte Agenturen.

Trotz der relativ klaren Befunde sollen Einschränkungen der Studie nicht ungenannt bleiben. Zunächst ist zu bemerken, dass die Ergebnisse vermutlich zeit- und kontextabhängig sind. Präferenzen können sich sowohl mit Blick auf die ästhetischen Merkmale der Stimmzettel als auch mit Blick auf die verfügbaren Informationen wandeln. So liegt es beispielsweise nahe, dass sensible Kandidateninformationen in Zeiten eines wachsenden Bewusstseins für Datenschutzfragen kritischer gesehen werden.

Wie bei den meisten Discrete-Choice-Experimenten ist auch bei der vorliegenden Studie zu hinterfragen, ob alle relevanten Dimensionen berücksichtigt wurden. Zwar bilden die analysierten Faktoren viele der empirisch zu beobachtenden Varianten deutscher Stimmzettel ab. Bestimmte Elemente wurden jedoch nicht berücksichtigt und insbesondere wurden viele der denkbaren Designvarianten durch das feste Grunddesign des fiktiven Stimmzettels ausgeschlossen. Zudem besteht ein gewisser

¹⁴ Beispielhaft seien einige typische Designmängel kommunaler Stimmzettel genannt. Häufig werden allgemeine typografische Grundregeln verletzt wie etwa die Mischung von Mittel- und Flattersatz. Zudem werden Informationen oftmals unstrukturiert und für eine schnelle Aufnahme unvorteilhaft dargestellt. Nicht selten fehlen Kontraste, beispielsweise durch Fettdruck oder Trennlinien, zur Verdeutlichung von Hierarchie und Zusammengehörigkeit. Auch die typografische Gestaltung ist der Lesefreundlichkeit in vielen Fällen abträglich. Der Zeilenabstand, der Abstand zwischen Buchstaben sowie zwischen Buchstaben und Trennlinien ist häufig zu gering und/oder nicht konstant. Schließlich nutzen die meisten Stimmzettel kein Gestaltungsraster, weshalb sich Textblöcke nicht auf den gleichen vertikalen und horizontalen Linien befinden und deshalb unsystematisch platziert wirken.

Status Quo Bias, da in der vorliegenden Arbeit nur Merkmale aufgegriffen wurden, die in der aktuellen Verwaltungspraxis auftreten. Gerade weitergreifende Reformpotenziale für Stimmzettel bleiben durch das gewählte Verfahren unberücksichtigt. Schließlich besteht besonders mit Blick auf die Informationen zum Wahlsystem noch viel Forschungsbedarf. So wurden in der vorliegenden Arbeit drei mögliche Erläuterungen der zulässigen Stimmabgabe auf ihre Nutzerfreundlichkeit geprüft. Die beobachtete Präferenz der Teilnehmerinnen für eine Erläuterung mittlerer Länge lässt sicher noch kein abschließendes Fazit über eine ideale Erläuterung zu. Stattdessen versteht sich dieser Beitrag eher als Auftakt für weitere Forschung in diesem Bereich sowie zu anderen Faktoren der Stimmzettelgestaltung. Dabei eignet sich gerade die Experimentalforschung hervorragend für die systematische Erweiterung unserer Erkenntnisse über die nutzerfreundliche Gestaltung von Stimmzetteln. Mit einem vergleichbaren Forschungsdesign können leicht weitere Varianten der Stimmzettelgestaltung auf ihre Nutzerfreundlichkeit geprüft werden, um schrittweise zu einem idealen Stimmzettel zu gelangen.

Die Frage nach dem idealen Stimmzettel unterstreicht abermals die Notwendigkeit eines Bewertungsmaßstabes. Es wurde dargelegt, dass Stimmzettel mit Blick auf verschiedene Kriterien optimiert werden können, etwa den Informationsgehalt, die Effektfreiheit oder die Nutzerfreundlichkeit, zwischen denen mitunter Zielkonflikte bestehen. Während diese Arbeit argumentiert, dass die Nutzerfreundlichkeit und der Wille der Wählerinnen und Wähler ein zentrales Kriterium bei der Gestaltung von Stimmzetteln sein sollte, bleibt es eine Abwägungsfrage, welchem Kriterium bei der Stimmzettelgestaltung der Vorrang eingeräumt wird. Die zentrale Botschaft dieser Arbeit ist vielmehr, dass Stimmzettel mit Blick auf verschiedene Kriterien optimiert werden können und dass sie gegenwärtig in keine Richtung optimiert sind.

Die Notwendigkeit eines normativen Maßstabs zur Bewertung von Stimmzetteln bedeutet jedoch nicht, dass ein Stimmzettel trotz bestehender Zielkonflikte nicht auch im Hinblick auf mehrere Kriterien optimiert werden kann. So sollte sich zukünftige Forschung ebenso der Frage zuwenden, ob und wie sich die Gestaltung der Stimmzettel auf den Wahlakt auswirkt. Beispielsweise sollten Stimmzettel den Anteil ungültiger Stimmen möglichst verringern (Blom-Hansen et al. 2016; Kim et al. 2015; Koppell und Steen 2004). Die vorliegende Studie kann nicht zu der Frage sprechen, welchen Einfluss die hier getesteten Faktoren auf ungültige Stimmen haben. Solche Fragen wären feldexperimentell zweifellos besser zu prüfen als im Rahmen eines Survey-Experimentes, gehen aber auch mit einem erheblichen Aufwand einher – von forschungsethischen Einwänden ganz zu schweigen.

Mit Blick auf die Verwaltungspraxis scheint die Sensibilität für die Notwendigkeit guter Stimmzettelgestaltung zu wachsen. In einer Reihe von Bundesländern mussten die Stimmzettel aufgrund der vielfachen kommunalen Wahlrechtsreformen umgestaltet werden (vgl. Vetter 2009). Eine besonders radikale Umkehr hat dabei das Land Bremen vollzogen. Im Zuge der Einführung des Mehrstimmenwahlrechts (vgl. Probst und Gattig 2012) wurden die Stimmzettel komplett neu gestaltet und die Erläuterung der Stimmabgabe wurde drastisch verkürzt. Die kurze Erläuterung im vorliegenden Experiment ist eng an die Beschreibung des Stimmverfahrens auf dem Bremer Stimmzettel angelehnt. Während die Bemühungen der Bremer Verwaltung um die Vereinfachung des Wahlaktes somit vorbehaltlos zu begrüßen sind,

deuten unsere Befunde doch darauf hin, dass die Verwaltung an dieser Stelle über das Ziel hinausgeschossen hat und die Wählerinnen sich etwas mehr Informationen wünschen würden.

Zunächst verdeutlicht dieses Resultat den Konflikt zwischen der Informiertheit und der Überforderung der Wählerinnen. Vor allem aber unterstreicht das Ergebnis auch den Wert von empirischer Sozialforschung, um Verwaltungshandeln durch wissenschaftliche Erkenntnisse zu untermauern. So konnte es in der vorliegenden Arbeit mit vergleichsweise geringem finanziellen und personellen Aufwand gelingen, systematische Einsichten über einen zentralen Baustein der repräsentativen Demokratie zu gewinnen. Wir hoffen, dass diese Arbeit dazu beitragen kann, den Dialog zwischen der Verwaltung und der empirischen Sozialwissenschaft zu stärken. Ohne Zweifel gibt es zahllose Aspekte von Verwaltungstätigkeit, bei denen die empirische Sozialforschung – nicht zuletzt durch Rückgriff auf experimentelle Verfahren – wichtige Hinweise zur Verbesserung des Austauschs zwischen Bürgerinnen und Bürgern und Verwaltung geben kann.

Danksagung Wir danken Paul Jahnke, Dave Köhler, Paul Kruse und Hauke Meisner für ihre Unterstützung. Wir danken Frédéric Ranft für seine Vorschläge zur Modernisierung des Designs von Stimmzetteln, die wir in dieser Arbeit mit seiner Erlaubnis aufgreifen und fortführen. Wir danken den Teilnehmenden des fünften *Norddeutschen Kolloquiums Sozialwissenschaften* und insbesondere Jan Berz für hilfreiche Kommentare zu diesem Beitrag. Schließlich danken wir den Gutachterinnen und Gutachtern und der Redaktion der Politischen Vierteljahresschrift für ihre umsichtigen und konstruktiven Kommentare, die zur Verbesserung dieser Arbeit sichtlich beigetragen haben.

Förderung Finanzielle Unterstützung wurde durch das Institut für Politikwissenschaft und die Philosophische Fakultät der Leibniz Universität Hannover zur Verfügung gestellt.

Funding Open Access funding provided by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Ansolabehere, Stephen, und Charles Stewart III. 2005. Residual votes attributable to technology. *Journal of Politics* 67(2):365–389.
- Arzheimer, Kai, und Jocelyn Evans. 2012. Geolocation and voting: candidate-voter distance effects on party choice in the 2010 general election in England. *Political Geography* 31(5):301–310.
- Banducci, Susan A., Jeffrey A. Karp, Michael Thrasher, und Colin Rallings. 2008. Ballot photographs as cues in low-information elections. *Political Psychology* 29(6):903–917.

- Behnke, Joachim. 2015. Der Einfluss der Kenntnis des Wahlsystems auf das Wahlverhalten: Weil sie nicht wissen, was sie tun, tun sie, was sie nicht wollen? *Zeitschrift für Parlamentsfragen* 46(3):588–607.
- Blom-Hansen, Jens, Jørgen Elklit, Søren Serritzlew, und Louise Riis Villadsen. 2016. Ballot position and election results: evidence from a natural experiment. *Electoral Studies* 44:172–183.
- Bullock, Charles S., III, und M.V. Hood III. 2002. One person—no vote; one vote; two votes: voting methods, ballot types, and undervote frequency in the 2000 presidential election. *Social Science Quarterly* 83(4):981–993.
- Calvo, Ernesto, Marcelo Escobar, und Julia Pomares. 2009. Ballot design and split ticket voting in multiparty systems: experimental evidence on information effects on a vote choice. *Electoral Studies* 28(2):218–231.
- Campbell, Rosie, und Philip Cowley. 2014. What voters want: reactions to candidate characteristics in a survey experiment. *Political Studies* 62(4):745–765.
- Campbell, Rosie, Philip Cowley, Nick Vivyan, und Markus Wagner. 2019. Why friends and neighbors? Explaining the electoral appeal of local roots. *Journal of Politics* 81(3):937–951.
- Carman, Christopher, James Mitchell, und Robert Johns. 2008. The unfortunate natural experiment in ballot design: the Scottish parliamentary elections of 2007. *Electoral Studies* 27(3):442–459.
- Carnes, Nicholas, und Noam Lupu. 2016. Do voters dislike working-class candidates? Voter biases and the descriptive underrepresentation of the working class. *American Political Science Review* 110(4):832–844.
- Dolezal, Martin. 2010. Exploring the stabilization of a political force: the social and attitudinal basis of Green parties in the age of globalization. *West European Politics* 33(3):534–552.
- Engst, Benjamin G., Thomas Gschwend, und Sebastian Sternberg. 2020. Die Besetzung des Bundesverfassungsgerichts: Ein Spiegelbild gesellschaftlicher Präferenzen? *Politische Vierteljahresschrift* 61:39–60.
- Hainmueller, Jens, Daniel J. Hopkins, und Teppei Yamamoto. 2014. Causal inference in conjoint analysis: understanding multidimensional choices via stated preference experiments. *Political Analysis* 22(1):1–30.
- Hanmer, Michael J., Won-Ho Park, Michael W. Traugott, Richard G. Niemi, Paul S. Herrnsen, Benjamin B. Bederson, und Frederick G. Conrad. 2010. Losing fewer votes: the impact of changing voting systems on residual votes. *Political Research Quarterly* 63(1):129–142.
- Herrnsen, Paul S., Michael J. Hanmer, und Richard G. Niemi. 2012. The impact of ballot type on voter errors. *American Journal of Political Science* 56(3):716–730.
- Jankowski, Michael. 2016. Voting for locals: voters' information processing strategies in open-list PR systems. *Electoral Studies* 43:72–84.
- Jankowski, Michael, Eric Linhart, und Markus Tepe. 2019. Welches Wahlsystem wollen die Wähler? Evidenz von einem Conjoint-Experiment. *Politische Vierteljahresschrift* 60:221–243.
- Jankowski, Michael, Eric Linhart, und Markus Tepe. 2020. Keep it simple! German voters' limited competence to evaluate electoral systems' functions. *German Politics* <https://doi.org/10.1080/09644008.2020.1726322>.
- Kane, John V., und Jason Barabas. 2019. No harm in checking: Using factual manipulation checks to assess attentiveness in experiments. *American Journal of Political Science* 63(1):234–249.
- Karp, Jeffrey A. 2006. Political knowledge about electoral rules: comparing mixed member proportional systems in Germany and New Zealand. *Electoral Studies* 25(4):714–730.
- Kelley, Jonathan, und Ian McAllister. 1984. Ballot paper cues and the vote in Australia and Britain: alphabetic voting, sex, and title. *Public Opinion Quarterly* 48(2):452–466.
- Kim, Nuri, Jon Krosnick, und Daniel Casasanto. 2015. Moderators of candidate name-order effects in elections: an experiment. *Political Psychology* 36(5):525–542.
- Kimball, David C., und Martha Kropf. 2005. Ballot design and unrecorded votes on paper-based ballots. *Public Opinion Quarterly* 69(4):508–529.
- Koppell, Jonathan G.S., und Jennifer A. Steen. 2004. The effects of ballot position on election outcomes. *Journal of Politics* 66(1):267–281.
- Leeper, Thomas, Sara B. Hobolt, und James Tilley. 2020. Measuring subgroup preferences in conjoint experiments. *Political Analysis* 28(2):207–221.
- Mandal, Manas K., und Tanusree Dutta. 2001. Left handedness: Facts and figures across cultures. *Psychology and Developing Societies* 13(2):173–191.
- Matson, Marsha, und Terri Susan Fine. 2006. Gender, ethnicity, and ballot information: ballot cues in low-information elections. *State Politics and Policy Quarterly* 6(1):49–72.
- McDermott, Monika L. 2005. Candidate occupations and voter information shortcuts. *Journal of Politics* 67(1):201–219.

- Mechtel, Mario. 2014. It's the occupation, stupid! Explaining candidates' success in low-information elections. *European Journal of Political Economy* 33:53–70.
- Moehler, Devra, und Jeffrey Conroy-Krutz. 2016. Eyes on the ballot: priming effects and ethnic voting in the developing world. *Electoral Studies* 42:99–113.
- Nyhuis, Dominic. 2014. Mehr Stimmen gleich bessere demokratische Qualität? Negative Effekte der Komplexität des Wahlrechts auf die Zusammensetzung der Wählerschaft. *Zeitschrift für Politikwissenschaft* 24(3):255–280.
- Pachón, Mónica, Royce Carroll, und Hernando Barragán. 2017. Ballot design and invalid votes: evidence from Colombia. *Electoral Studies* 48:98–110.
- Piliavin, Jane Allyn. 1987. Age, race, and sex similarity to candidates and voting preference. *Journal of Applied Social Psychology* 17(4):351–368.
- Portmann, Lea, und Nenad Stojanović. 2019. Electoral discrimination against immigrant-origin candidates. *Political Behavior* 41(1):105–134.
- Probst, Lothar, und Alexander Gattig (Hrsg.). 2012. *Das neue Wahlsystem in Bremen: Auswertung und Analyse der Kommunikationskampagne und der Wirkungen des neuen Wahlsystems*. Bremen: Bremische Bürgerschaft.
- Reidy, Theresa, und Fiona Buckley. 2015. Ballot paper design: evidence from an experimental study at the 2009 local elections. *Irish Political Studies* 30(4):619–640.
- Reynolds, Andrew, und Marco Steenbergen. 2006. How the world votes: the political consequences of ballot design, innovation and manipulation. *Electoral Studies* 25(3):570–598.
- Schäfer, Armin, und Harald Schoen. 2013. Mehr Demokratie, aber nur für wenige? Der Zielkonflikt zwischen mehr Beteiligung und politischer Gleichheit. *Leviathan* 41(1):94–120.
- Schmitt-Beck, Rüdiger. 1993. Denn sie wissen nicht, was sie tun... Zum Verständnis des Verfahrens der Bundestagswahl bei westdeutschen und ostdeutschen Wählern. *Zeitschrift für Parlamentsfragen* 24(3):393–415.
- Sigelman, Lee, und Carol K. Sigelman. 1982. Sexism, racism, and ageism in voting behavior: an experimental analysis. *Social Psychology Quarterly* 45(4):263–269.
- Sinclair, Robert C., Melvin M. Mark, Sean E. Moore, Carrie A. Lavis, und Alexander S. Soldat. 2000. An electoral butterfly effect. *Nature* 408:665–666.
- Tavits, Margit. 2010. Effect of local ties on electoral success and parliamentary behaviour. *Party Politics* 16(2):215–235.
- Vetter, Angelika. 2009. Citizens versus parties: explaining institutional change in German local government, 1989–2008. *Local Government Studies* 35(1):125–142.
- Wand, Jonathan N., Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane Jr., Michael C. Herron, und Henry E. Brady. 2001. The butterfly did it: the aberrant vote for Buchanan in Palm Beach County, Florida. *American Political Science Review* 95(4):793–810.
- Webster, Steven W., und Andrew W. Pierce. 2019. Older, younger, or more similar? The use of age as a voting heuristic. *Social Science Quarterly* 100(3):635–652.