



*Journal of Articles in Support of the Null Hypothesis*

Vol. 18, No. 2

Copyright 2022 by Reysen Group. 1539–8714

www.jasnh.com

# Does word order predict ethnocentric helping behavior? Results of a ‘lost letter’ field experiment in Berlin

---

Désirée Meuthen  
**Berlin**

Aileen Oeberst  
**FernUniversität Hagen**

Word order in intergroup labels may strengthen ethnocentrism as people tend to name their in-group first and perceive it as more important than the second. To test for corresponding helping behavior in a realistic intergroup setting, a lost letter field experiment was conducted. In a  $2 \times 2$  between-subject design, 680 apparently lost letters were labeled with German-Turkish vs. Turkish-German ordered indications and dispersed throughout Berlin, Germany, in a district with low vs. high Turkish population rates. Descriptively, German-Turkish letters were returned more often than Turkish-German ones. Yet, these differences were not statistically significant, neither was the interaction of word order and district. Thus, we did not obtain evidence for ethnocentric bias in helping behavior due to mere word order.

*Keywords:* word order, ethnocentrism, lost letter technique, helping behavior

---

*Corresponding author:* Prof. Dr. Aileen Oeberst; FernUniversität Hagen; Universitätsstraße 47, 58097 Hagen, Germany; Phone: +49 2331 987-4613;  
E-mail: aileen.oeberst@fernuni-hagen.de

We provide this study's data, syntax, and supplemental material online on the Open Science Framework: <https://osf.io/8p6rx>

*Author's Contributions:* Author 1: Conceptualization, Methodology, Formal Analysis, Investigation, Resources, Data Curation, Writing—Original Draft.  
Author 2: Conceptualization, Methodology, Writing—Review & Editing, Supervision.

Language is sequential. At that very moment, we put one word after the other in order to tell you about our research. And it is also at that very moment that you discover word by word what we found out. That is, words need to be ordered, and even though there are several linguistic rules guiding this process, many degrees of freedom remain. Interestingly, however, word order is rarely a coincidence. Research has shown that people and groups tend to put themselves into the first place (Cooper & Ross, 1975). The same war in 1864, for instance, is called *Danish-German War* in Danish, but *German-Danish War* in German. Even more importantly, the way groups are ordered affects our perceptions of these groups. Those mentioned first are perceived to be more important (Oeberst & Matschke, 2017). Taken together, word order may foster ethnocentrism as groups tend to put themselves first and, by doing so, grant more attention to themselves and nurture their own importance over others. But how far does this go? Does mere word order affect altruistic intergroup behavior; for instance? To examine this question, we made use of Milgram's (1977) lost letter technique. Specifically, we addressed letters to a German-Turkish versus Turkish-German Community Center, distributed them throughout two different districts of Berlin, Germany, with higher vs. lower rates of Turkish residents, and compared the return rates. Before presenting our study and results in detail, we will briefly outline relevant linguistic research on word order and link it to social psychological research on groups. Next, we elaborate on the consequences of ordering groups and introduce Milgram's lost letter paradigm as a suitable test for our research question.

### Origins of Binomial Word Order: From Worlds Into Words

At first glance, notations like *law and order*, *fish and chips*, or even *German and Turkish* do not have much in common. We know them as a name for a TV show, as one of the British' (supposed) favorite food and, last but not least, as indications of two groups. In linguistic terms, however, all of the above examples are *binomials*—two-word sequences that share the same syntactical level and are typically linked by the word *and* (Malkiel, 1959). And while their order is not predefined by any necessity, their arrangement is typically very stable and all but random (Mollin, 2014). Cooper and Ross (1975) were the first to show that word order follows rules grounded in human life. The authors identified more than 20 semantic principles that share a distinctive commonality, the major rule under which they all can be subsumed—the “me-first” principle, which implies that people tend to name those things first that are linked to them personally (see also Benor & Levy, 2006; Sobkowiak, 1993).

According to these principles, the ordering of other people is largely determined by the link to oneself: The person perceived as more similar to oneself tends to get the pole position: More closely related family members are placed first when it comes to kin binomials (McGuire & McGuire, 1992), and women refer to the female part of a familiar couple first, whereas men mention

the male name first (Hegarty, 2014; Hegarty et al., 2011). In the same vein, the order of groups depends on social identity and group membership of the speaker: People put their own group in the first place. Oeberst and Matschke (2017) demonstrated this tendency not only under controlled experimental conditions but also for actual historical intergroup conflict titles from more than 40 languages: In more than 85% of the 172 cases, the own group was referred to first—as already illustrated in the above-mentioned example of the *German-Danish War* vs. *Danish-German War*. In a linguistic sense, such intergroup conflict titles are *double adjective premodifications*—the combination of two adjectives (that have to be ordered) preceding a noun (Hegarty et al., 2016). Although less frequently researched, it had been suggested their order is likewise affected by semantic factors (e.g., Wulff, 2003).

Whether people give precedence to their own group *deliberately*, is not yet clear. That is, mentioning one's group first does not necessarily imply an underlying *ethnocentric motivation* (Oeberst & Matschke, 2017). What is clear, however, is that ethnocentrism may *result* from practicing the “we-first”-order, as we will outline next.

### Consequences of Word Order: From Words Into Worlds

Language shapes social reality (Fiedler, 2008), by affecting the way we think about the world (Boroditsky, 2003). Consequently, one may ask whether the perception of ordered groups determines how we relate to these groups. Numerous researchers from various psychological fields found order effects in form of a primacy bias—by generally granting more importance to the first position in an order (e.g. Asch, 1946; Bettinsoli et al., 2015; Dean, 1980; Stevens & Duque, 2019). Thus, what comes first is given more weight (Gundel, 1988; Johnson-Laird, 1968). This is, in fact, what Oeberst and Matschke (2017) found when analyzing *perceptions* of the two groups mentioned in a conflict title depending on their word order: The group mentioned first was generally perceived as more important and powerful than the second one. That is, the effect was independent of the specific groups involved and, thus, also independent of the fact whether it was about participants' ingroup or outgroups—or even entirely unknown groups, for that matter. When it is mostly the own group that is mentioned first, however, ethnocentrism may result. After all, if the first place is generally perceived as more important, then coming across the own group first could nourish peoples' perception that it is more important than other groups.

Crucially, the general effect of word order on group perceptions—without a pronounced effect for the ingroup—obtained by Oeberst and Matschke (2017) could have been due to the paradigm of their study: First, participants were always asked to rate both groups involved or even had to provide relative ratings (i.e., with both groups as anchors on a scale), which explicitly directed participants' attention to both groups. This might have effectively countered biases as it could have invited participants to monitor their responses (in order to not come across as biased,

e.g., Nederhof, 1985) due to a possible sensitization for the role of their own group. Second, the combinations used with regard to the ingroup (e.g., *German-Danish* vs. *Danish-German*) were neither very prevalent ones, nor did they involve a salient outgroup (see Asbrock et al., 2014, who found that Danes were not mentioned when asking German respondents about groups associated with the term “foreigners”). While it was for these reasons that the authors chose that combination—in order to minimize the influence of stereotypes as well as familiarity effects due to the high prevalence of one word order over the other—it raises the question of whether the results generalize to other group combinations.

The present study therefore set out to provide a more critical test of the consequences of word order by (1) assessing actual helping behavior in the real world and (2) combining the ingroup (Germans) with the largest foreign national group in Germany: Turkish people<sup>1</sup> (Bundesamt für Migration und Flüchtlinge, 2019) who are not only highly prevalent in Berlin (Amt für Statistik Berlin-Brandenburg, 2020b), where the study was conducted but who are also highly salient in peoples’ minds: When asked about foreigners living in Germany a large amount of the respondents named Turkish people first within different studies and national surveys (e.g., Asbrock et al., 2014; GESIS, 2017). Additionally, research showed strong prejudices and negative stereotypes toward Turkish people as a stigmatized ethnic group in Germany (e.g., Asbrock, 2010).

With regard to intergroup helping behavior, previous research clearly documented in-group favoritism: Helping someone in need increased with a common group membership (e.g., Levine et al., 2002; Levine et al., 2005). These results are in line with a meta-analysis of Balliet et al. (2014) who found that people, in general, show more cooperative behavior for in-group versus out-group members. The authors identified this pattern even in the case that prosocial behavior was linked to personal costs for the helping person. Thus, helping behavior and the choice of a cooperative action are more likely when it comes to a member of the own group. Applied to word order, the following question arises: Does this effect generalize to cases where the own group comes *first* in a combination of two (e.g., *German-Turkish* vs. *Turkish-German*)? We tested this question with a lost-letter study (Milgram, 1977), in which stamped letters were labeled either with a fictional German-Turkish or Turkish-German sender and recipient and dispersed in Germany’s capital Berlin. Helping behavior in this paradigm thus consisted of posting the letter into a letterbox. And if helping behavior was, in fact, sensitive to word order, one would expect more letters with German-Turkish labels being returned (i.e., having been posted) as Germans are the dominant national in-group in Berlin (Amt für Statistik Berlin-Brandenburg, 2020b).

However, another explanation for such an effect could also be mere familiarity rather than in-group favoritism, as the expression “German-Turkish” is much more common than

“Turkish-German”, both in German as well as Turkish.<sup>2</sup> While this is reminiscent of the possibility that reasons other than an ethnocentric motivation may underlie the word order effect (e.g., alphabetical order explaining both cases perfectly), it provides an alternative explanation for higher return rates of letters with German-Turkish labels that is not group-related (but based on familiarity instead). In order to take this into account, we added another factor to our design and made use of the fact that the proportion of Turkish people varies substantially between Berlin’s districts (Amt für Statistik Berlin-Brandenburg, 2020b). Thus, we selected the district with the highest (Neukölln, 11%) and one of the lowest rates (Lichtenberg, 1%) of Turkish people (Amt für Statistik Berlin-Brandenburg, 2020b). Despite further differences (e.g., unemployment rate, Arbeitsagentur, 2020), the two districts were chosen as they are comparable with respect to (a) population density, and (b) the proportion of urban and rural areas. If helping behavior was sensitive to *group* order, that is, the position of the ingroup, one would expect the return rate to differ as a function of group order and district—with German-Turkish letters returning disproportionately more frequently when dispersed in Lichtenberg but less so when dispersed in Neukölln. If it was rather *word* order and its mere familiarity, in contrast, that affected helping behavior, one would expect a main effect of word order that is not qualified by the district as the combination of German-Turkish is generally much more prevalent. Consequently, we tested the following two preregistered hypotheses:

(H1) Overall, significantly more letters with a German-Turkish label are returned than letters with a Turkish-German label.

(H2) A significant interaction of label and district is expected, with relatively more letters with a German-Turkish label being returned from a district with a low (vs. high) proportion of Turkish people, and the reversed pattern for letters with a Turkish-German label.

## Method

In line with Simmons et al. (2012), we report how sample size, all data exclusions, manipulations, and measures including the underlying design, material, and procedure for this study were determined. The current study was preregistered (see <https://aspredicted.org/vj778.pdf>).

### The Lost Letter Technique

We employed the lost letter technique developed by Milgram

<sup>1</sup> Consistent with the classification of the Amt für Statistik [Department for Statistics] Berlin-Brandenburg (2020a), the terms “Turkish people” resp. “Turks” in this paper refer to people who have (1) a Turkish citizenship or (2) a Turkish migration background (defined by native country, second citizenship, naturalization, or parental background).

<sup>2</sup> A research on Google showed distinct numbers in the results for the expressions “Deutsch-Türkisch“ [German-Turkish] vs. “Türkisch-Deutsch“ [Turkish-German]: 3,5 million vs. 1,5 million (including all declinations and spellings). When searching for the respective Turkish terms the version with the German group first was also the predominant one: “Almanca Türkçe“ [German-Turkish] 1,5 million vs. “Türkçe Almanca“ [Turkish-German] 1 million.

et al. (1965): Addressed and stamped but unposted letters were dispersed throughout a certain area, city, or location. When a finder comes across such a letter, he or she must decide what to do: ignoring, destroying, or posting (Milgram, 1977). By varying the ostensible sender, recipient, or both, the letters' return rates can be attributed to the indications on the envelope, and thus interpreted as a positive attitude or prosocial behavior towards the indicated group, organization, person, or issue (Milgram et al., 1965). In the past decades the lost letter technique has been applied in a wide range of studies with various research foci, mainly to analyze possible discrimination against ethnic, national, or religious groups (e.g., Ahmed, 2010; Berger & Berger, 2019; Hellmann et al., 2015; Hellmann et al., 2020; Klink & Wagner, 1999; Kremer et al., 1986; Petrykowski et al., 2010; Phillips et al., 2016).

## Design

Our predictions were tested in a 2 (order of indications: German-Turkish vs. Turkish-German)  $\times$  2 (urban district: low vs. high population rates of Turkish people) between-subject design, resulting in four experimental conditions.

## Sample

In order to be able to detect small to medium-sized effects ( $OR = 1.3$ ), an a priori power analysis with G\*Power (Faul et al., 2009) suggested a sample size of  $N = 568$  letters (i.e., 142 letters per experimental condition, alpha level = .05, 1 - beta = .95,  $df = 1$ ) for the logistic regression analyses. Yet, due to two special local and situational circumstances, we opted for a higher sample size: First, a previous lost letter study that was realized in Berlin found very different return rates depending on the specific area of dispersion (Koopmans & Veit, 2014). Second, we expected return rates to be generally diminished due to the Corona pandemic: Fewer people than usually might pick up the letters based on fear of a Corona infection or a general sensitization for viral diseases. Therefore, we increased the number of letters per experimental condition to  $n = 170$ , resulting in a total of  $N = 680$  letters.

## Material

All the dispersed letters consisted of a white standard envelope and were franked with a neutral 0.80 Euro flower motive stamp. Only the ostensible sender and recipient varied: 340 letters were labeled with *Deutsch-Türkische Begegnungsstätte* [German-Turkish community center] as sender and *Deutsch-Türkisches Kulturzentrum* [German-Turkish cultural center] as the recipient. The other 340 letters were labelled with the reversed order in form of *Türkisch-Deutsche Begegnungsstätte* [Turkish-German community center] and *Türkisch-Deutsches Kulturzentrum* [Turkish-German cultural center].<sup>3</sup>

The recipient's details were printed on the letters' inlay and visible through the envelopes' front window, the sender's details on the back via a seal. These two different kinds of printing made the envelopes look more official, and therefore reputable. The

ostensible sender and recipient were printed in bold letters to enhance their visibility. Intentionally, both national indications were written in capital letters at the beginning of the words to prevent a bias in the groups' perception of importance only because of the orthographical notations. The address details were the same in both experimental conditions and contained information of a post office box in Berlin for the sender and a private address in North Rhine-Westphalia for the recipient.<sup>4</sup> When selecting the addresses, it was made sure that the corresponding postcodes (10125 and 52249) looked as differently and widely separated as possible to increase plausibility. We refrained from using real names as a postal conjunct (like e.g., "c/o Müller") as names may signal ethnicity or social background and therefore evoke stereotypes and discrimination (e.g., Carpusor & Loges, 2006).

In contrast to former lost letter studies, not only the sender or the recipient on the envelopes were manipulated but both, for two reasons: First, to strengthen the manipulation. This was done due to the fact that Oeberst and Matschke (2017) found smaller effect sizes in their online studies when the order of the two groups in the conflict titles was inconsistent with the order presented in the response scales. Second, to enhance the odds that the nationality indications could actually be seen and read by a potential finder—regardless of potentially factors (e.g., weather) that might affect a letter's position.

In case that someone would open a letter, it contained a short note in German that could not be seen through the envelope. The note implied a pretended postponement of an ostensible weekend-workshop in the cultural community center some weeks later. With regard to the Corona pandemic, this scenario would seem plausible, although the word *Corona* was not used in the note to prevent evoking any feelings and rather keep it as neutral as possible. No specific location was mentioned to prevent the appearance of possible attendants. Additionally, the respective date and district were printed on the note, and the particular areas of dispersion were documented daily to make the experimental conditions of a returned letter traceable. An example for the dispersed letters and the wording of the letters' inlay are illustrated in the Supplemental Material (see <https://osf.io/8p6rx/>).

## Procedure

The field experiment was conducted between August 4 and August 27, 2020. The prepared letters were dispersed throughout Berlin in parks, in front of administrative offices, ticket machines, cycle tracks, and bottle banks, at stops of subways, trams, and city trains, and on sidewalks and benches, so that they appeared to be lost. To minimize the risk that a dispersed letter would be visible from the different dispersion places, a distance of about 200 meters was ensured. Also, every place of dispersion was used only once during the whole study to reduce the chances that someone

<sup>4</sup> Especially, for legal and ethical reasons it may be critical touching and using other people's private property, even when doing it with no ill intent but for research reasons. Additionally, some authors abstained completely from dispersing letters on parked cars' windscreens as research assistants reported negative experiences in the respective environments (e.g., Reid, 2018).

might find more than one letter. The letters were dispersed only during the daytime and not on rainy or stormy days to ensure their findability and indications' readability. Places near traffic lights or heavily traveled street corners were avoided for safety reasons. Additionally, the letters were dispersed only in public places and not on private grounds. In contrast to a majority of former lost letter studies, parked cars were excluded as dispersion places because by dispersing a letter on a car's windscreen only one person would have been addressed: the respective car owner. As only every third person in Berlin owns a car (Kraftfahrtbundesamt, 2019) two-third of the population would have been excluded from the study in general.<sup>5</sup> Similarly, industrial business areas were excluded as a maximum of 1,000 meters distance to the next letterbox (which, for residential areas, is regulated by the legal framework in Germany, Bundesnetzagentur, 2017) could not be guaranteed.

Several procedures were implemented to avoid any confounds: For every day and every of the 34 dispersion areas, a stack of letters was prepared that contained the same amounts of letters with the two kinds of manipulated word order. Before the dispersion, the stacks were shuffled to arrange the letters in a random order for an unsystematic distribution. When choosing areas for dispersion, social indices (social structural analyses of the Senatsverwaltung für Gesundheit und Soziales [Senate Department for Health and Social Issues], 2014, and the Senatsverwaltung für Stadtentwicklung und Wohnen [Senate Department for Urban Development and Housing Issues], 2019a; 2019b) were taken into account and held as constant as possible per day.

A detailed documentation of the dispersion including the respective date, district, planning area, and socioeconomic index plus the numbers of dispersed and returned letters for the experimental conditions is provided in the Supplemental Material: <https://osf.io/8p6rx/>

## Measurements

In line with the lost letter technique (Milgram, 1977), helping behavior was operationalized as returning a lost letter to the supposed recipient or sender. Consequently, the main dependent variable was the return rate. As preregistered, letters that did not arrive until three weeks after their dispersion were coded as not returned (none did). This time frame was adapted from previous studies (e.g., two weeks, Hellmann et al., 2020, study 1; four weeks, Hellmann et al., 2020, study 2; 1–20 days of return time, Reid, 2018). Furthermore, it might be assumed that a letter not returning after three weeks of dispersion was either destroyed, disposed or had become illegible.

## Results

<sup>5</sup> Especially, for legal and ethical reasons it may be critical touching and using other people's private property, even when doing it with no ill intent but for research reasons. Additionally, some authors abstained completely from dispersing letters on parked cars' windscreens as research assistants reported negative experiences in the respective environments (e.g., Reid, 2018).

Table 1. Absolute numbers and rounded percentages of the letters' return rates depending on word order and urban district

Urban district of dispersion	Word order on the letters' labels		Total
	German-Turkish	Turkish-German	
Lichtenberg (low Turkish rates)	85 (50%)	76 (45%)	161 (47%)
Neukölln (high Turkish rates)	76 (45%)	74 (44%)	150 (44%)
Total	161 (47%)	150 (44%)	311 (46%)

Note. Dispersed letters  $N = 680$ . The percentages in the four experimental conditions refer to the number of dispersed letters per cell ( $n = 170$ ), the total percentages to the overall number of dispersed letters per factor ( $n = 340$ ).

Overall, 311 of the 680 (i.e., 46%) dispersed letters were returned. In order to test whether the return rate varied as a function of word order (H1) and/or the interaction of word order and urban district (H2), we conducted logistic regression analyses (e.g., Osborne, 2006).

## Preregistered Analyses

### Return Rates Depending on Word Order

Hypothesis 1 proposed that letters with a German-Turkish label were more often returned than letters with a Turkish-German label. Descriptively, this was true, as 47% (161 out of 340) letters labelled with the German-Turkish order were returned while only 44% (150 out of 340) with the Turkish-German label got back (see Table 1). However, the logistic regression model including word order as a single predictor was not significant (see also Table 2). Consequently, the return rate was not affected by word order,  $b = 0.13$ ,  $OR = 1.14$ ,  $p = .397$ , 95% CI [0.84, 1.54]. Accordingly, the data does not support hypothesis 1.

### Return Rates Depending on District and Word Order

Taking the different districts into account, Hypothesis 2 stated that letters with a German-Turkish label were more likely to be returned from the district with a lower (vs. higher) population of Turkish people (Lichtenberg vs. Neukölln), while the reverse pattern was expected for letters with a Turkish-German label. When additionally entering both district and the interaction term of word order and district into the logistic regression analyses, the predicted interaction was not obtained,  $b = 0.17$ ,  $OR = 1.18$ ,  $p = .593$ , 95% CI [0.64, 2.16]. Furthermore, none of the models was significant (see also Table 2). As presented in Table 1, the return rates were descriptively slightly higher for letters with German-Turkish labels that had been lost in the district with rather a lower rate of Turkish residents (50%) when compared to all other experimental conditions (44–45%), but the difference was not statistically significant. Thus, hypothesis 2 was likewise not supported by the data. The interplay of the district and word order did not affect the return rates. Or put differently, helping behavior was rather stable and did not vary as a function of word order and population composition.

Table 2. Logistic regression predicting the likelihood of returning a lost letter depending on word order, urban district, and the respective interaction

Models and predictors	Model $\chi^2(df)$	Nagelkerke $R^2$	95% CI					
			$b$	$SE$	$p$	$OR$	$LL$	$UL$
Model 1	0.72 (1)	.00			.397			
Word order			0.13	0.15	.397	1.14	0.84	1.54
Model 2	1.44 (2)	.00			.488			
Word order			0.13	0.15	.397	1.14	0.84	1.54
Urban district			0.13	0.15	.397	1.14	0.84	1.54
Model 3	1.72 (3)	.00			.632			
Word order			0.38	0.49	.438	1.46	0.56	3.79
Urban district			0.38	0.49	.438	1.46	0.56	3.79
WO $\times$ UD			0.17	0.31	.593	1.18	0.64	2.16

Note.  $N = 680$ . Reference category for word order: Turkish-German, for the urban district: Neukölln.  $OR$  = odds ratio,  $CI$  = confidence interval,  $LL$  = lower limit,  $UL$  = upper limit,  $WO$  = word order,  $UD$  = urban district. The identical values in models 1 and 2 are due to the identical distribution of returned letters (160 vs. 151) among word order and urban district.

## Exploratory Analyses

### Return Rates Depending on SES and Word Order

As low income and economic levels can reduce return rates as former lost letter studies showed for London (Holland et al., 2012), the Netherlands (Volker et al., 2015), Australia (Westlake et al., 2019), and Italy (Baldassarri, 2020), a second logistic regression was run in which the SES (1 = very low, 2 = low, 3 = medium, 4 = high) was added as three dummy variables with *very low* as the reference category in the first, the word order in the second, and the interaction terms in the third model. Table 3 shows that return rates increased with increases in SES. Letters dispersed in areas with a low SES,  $b = 0.86$ ,  $OR = 2.37$ ,  $p = .002$ , 95%  $CI$  [1.37, 4.08], a medium SES,  $b = 1.26$ ,  $OR = 3.53$ ,  $p < .001$ , 95%  $CI$  [2.19, 5.69], and a high SES,  $b = 2.03$ ,  $OR = 7.58$ ,  $p < .001$ , 95%  $CI$  [4.00, 14.34], were more likely to be returned than letters from very low SES areas. Again, word order was not significantly predictive of return rates,  $ORs = 0.44 - 1.15$ ,  $ps > .200$  (see also Table 4). Thus, helping behavior increased with increasing SES. However, SES did not interact with word order.

### Return Time Depending on Word Order, District, and SES

In addition to the return rates we analyzed the return time, that is the interval between distributing and posting a letter. After all, one could argue that letters with a longer return time were passed by and ignored more frequently, thus potentially indicating lower helping behavior. The time for returning a lost letter was computed by the difference between the dispersion date (documented on the letters' inlay) and the postmark date on

Table 3. Absolute numbers and rounded percentages of the letters' return rates depending on word order and socioeconomic status

SES index	Dispersed letters	Returned letters	Word order on the letters' labels		
			German-Turkish	Turkish-German	Total
Very low	120	27	11 (9%)	16 (13%)	22%
Low	140	57	29 (21%)	28 (20%)	41%
Medium	340	172	92 (27%)	80 (24%)	51%
High	80	55	29 (36%)	26 (33%)	69%

Note. Dispersed letters  $N = 680$ , returned letters  $n = 311$ . The rounded percentages of returned letters for both word order conditions refer to the respective table row, that is, relational to the total number of dispersed letters in each of the four SES levels. The uneven numbers of dispersed letters among the SES indices were due to the compliance of preferably constant socio-structural parameters during the experiment's conduction for the test of Hypothesis 2 (see also Supplemental Material for a detailed documentation including the SES index for all areas of dispersion).

the envelopes.<sup>6</sup> Envelopes with a missing ( $n = 6$ ) or unreadable ( $n = 8$ ) postmark were excluded, unreturned letters ( $n = 369$ ) were treated as missing data, resulting in a final sample of  $N = 297$  letters. The average return time was 1.59 days ( $SD = 1.56$ , range: 0–11 days). The means and standard deviations of return time for all experimental conditions among word order and urban district are presented in Table 5. A bootstrapped linear regression analysis basically replicated our previous findings.<sup>7</sup> The return time was neither significantly predicted by word order,  $b = -0.15$ ,  $t(296) = -0.84$ ,  $p = .401$ , nor by urban district,  $b = 0.10$ ,  $t(296) = 0.52$ ,  $p = .603$ , nor their interaction,  $b = -0.24$ ,  $t(296) = -0.65$ ,  $p = .517$  (see also Table 6).

<sup>6</sup> One could argue that the number of days between dispersion and postmark may not illustrate the exact return time due to postal mistakes, extended processing time, or partly no letterbox emptying on Sundays. As such incidents would have been centered among the whole study's conduction and all experimental conditions, this procedure yet seemed adequate, also in respect of no other possible measurement.

<sup>7</sup> Due to the continuous dependent variable of return time, a hierarchical linear regression was performed. The Shapiro-Wilk test assessed that none of the groups was normally distributed ( $p < .001$ ). The homogeneity of variances was not met as assessed by the Levene's test ( $p = .027$ ,  $.002$ ). Conducting a parametric procedure could then lead to an incorrect estimation of the standard errors and, subsequently, to less reliable significance tests (e.g., Wright et al., 2011). Such issues might be minimized by applying non-parametric bootstrapping (Field, 2013). Especially, when parametric assumptions like normality and homogeneity are violated and the sample size is 50 at a minimum (which was fulfilled) bootstrapping has shown to be a very robust procedure (e.g., Hesterberg, 2015). For that reason and based on statistical recommendations (e.g., Fitrianto & Cing, 2014; Rousselet et al., 2019), all linear estimates were based on 10,000 bootstrap samples from the observed data.

Table 4. Hierarchical logistic regression predicting the likelihood of returning a lost letter depending on word order, socioeconomic status, and the respective interactions

Models and predictors	Model $\chi^2$ (df)	Nagelkerke $R^2$						95% CI	
			$b$	$S$	$p$	OR	LL	UL	
Model 1	49.88 (3)	.10			< .001				
Low SES			0.86	0.28	.002	2.37	1.37	4.08	
Medium SES			1.26	0.24	< .001	3.53	2.19	5.69	
High SES			2.03	0.33	< .001	7.58	4.00	14.34	
Model 2	50.65 (4)	.10			< .001				
Low SES			0.86	0.28	.002	2.37	1.37	4.09	
Medium SES			1.26	0.24	< .001	3.53	2.19	5.70	
High SES			2.03	0.33	< .001	7.60	4.01	14.38	
Word order			0.14	0.16	.380	1.15	0.84	1.57	
Model 3	53.33 (7)	.10			< .001				
Low SES			1.15	0.41	.005	3.15	1.40	7.07	
Medium SES			1.66	0.37	< .001	5.25	2.56	10.80	
High SES			2.46	0.49	< .001	11.74	4.53	30.48	
Word order			-0.48	0.44	.277	0.62	0.26	1.47	
L SES $\times$ WO			-0.54	0.56	.335	0.58	0.19	1.75	
M SES $\times$ WO			-0.77	0.49	.121	0.47	0.18	1.23	
H SES $\times$ WO			-0.83	0.66	.205	0.44	0.12	1.58	

Note.  $N = 680$ . Reference category for SES: very low. OR = odds ratio, CI = confidence interval, LL = lower limit, UL = upper limit, L = low, M = medium, H = high, WO = word order.

When additionally including SES as a predictor, the results' pattern was identical to return rates with SES being the only significant predictor and no effects of word order by itself nor in combination with SES (see <https://osf.io/8p6rx/> for the full results).

## Discussion

The current study set out to test for potential consequences of ordered groups in the real world, as previous research (Oeberst & Matschke, 2017) came along with some relevant limitations (see above). In a nutshell, return rates did not differ for letters labeled with a German-Turkish or Turkish-German addressee and addressor (and neither did return times). Likewise, return rates and times did not differ as a function of group order and district. That is, differences in the German-Turkish population rates also did not interact with word order. Consequently, we did not obtain any evidence for a pronounced reception effect for the ingroup. Of course, this conclusion needs to be treated with caution as we simply do not know anything about the people who decided to mail the letters. Germans, however, still represent the largest group of all inhabitants in Berlin (Amt für Statistik Berlin-

Table 5. Means and standard deviations of the letters' return time depending on word order and urban district

Return time	German-Turkish word order			Turkish-German word order		
	Lichtenberg ( $n = 78$ )	Neukölln ( $n = 75$ )	Total ( $n = 153$ )	Lichtenberg ( $n = 74$ )	Neukölln ( $n = 70$ )	Total ( $n = 144$ )
$M$	1.56	1.77	1.67	1.53	1.50	1.51
$SD$	1.30	1.93	1.63	1.09	1.82	1.49

Note.  $N = 297$ . Lichtenberg: low rate of Turkish people, Neukölln: high rate. The return time was measured in days, smaller values indicate a faster return time.

Table 6. Hierarchical linear regression predicting the letters' return time depending on word order, urban district, and the respective interaction

Models and predictors	$R^2$	$b$	$SE$	$\beta$	$t$	$p$	95% CI	
							LL	UL
Model 1	.00					.401		
Word order		-0.15	0.18	-.05	-0.84	.401	-0.51	0.20
Model 2	.00					.614		
Word order		-0.15	0.18	-.05	-0.84	.403	-0.51	0.19
Urban district		0.10	0.19	.03	0.52	.603	-0.26	0.46
Model 3	.01					.707		
Word order		0.20	0.50	.06	0.35	.728	-0.79	1.22
Urban district		0.45	0.59	.14	0.78	.436	-0.72	1.67
WO $\times$ UD		-0.24	0.37	-.16	-0.65	.517	-0.94	0.46

Note.  $N = 297$ . CI = confidence interval, LL = lower limit, UL = upper limit, WO = word order, UD = urban district. Standard errors and confidence intervals based on 10,000 Beas bootstrap samples.

Brandenburg, 2020b) and it is, thus, a matter of mere probability, that mostly Germans came across the letters – even more so in the district of Lichtenberg. Still, this did not lead to higher return rates for German-Turkish-labelled letters in general or Lichtenberg, in particular.

Basically, our results are in line with the general effect obtained by Oeberst and Matschke (2017): If the first group was *generally* perceived as more important, regardless of the specific groups involved, then one would not expect any differences in return rates. Importantly, this does not preclude ethnocentrism to be nourished by word order: After all, it is the *combination* of giving (verbal) precedence to the ingroup and associating the first position with higher importance, that may foster ethnocentrism. And as we have outlined above, one order—“German-Turkish”—is much more prevalent in real life than the other one (“Turkish-German”). But in effect, this did not translate into differential helping behavior in our study.

As previous studies in Germany did find significant differences in return rates between Turkish and German labels (Hellmann et al., 2015; Klink & Wagner, 1999; but see Koopmans & Veit, 2014, for a null-effect), it might be speculated that our manipulation of

mere group *order* was too subtle to exert any effects. After all, each label always included both groups (vs. only one of the groups as typically realized in lost letter studies, Volker et al., 2015). It might even be argued that the labels in our study (G-T/T-G cultural or community center) could be recategorized into a common superordinate group, namely organizations that are concerned with intercultural exchange. If this was the case, one would not expect people to make any differences between the order in which groups are mentioned (Gaertner et al., 2011).

Another aspect is worth mentioning, however. Note, that the original paper found a reception effect not only concerning relevance but also for power (Oeberst & Matschke, 2017). That is, the group mentioned first was not only perceived as more important but also as more powerful. And while this effect might be limited to the context of inter-group conflicts (as tackled in all studies of Oeberst & Matschke, 2017), it could also have contributed to the present results. After all, being powerful is about the opposite of being in need of help. In other words, if the group mentioned first was perceived as more powerful, it might also be regarded as less needy. Consequently, a political context (e.g., international relations) or even a more competitive context as present in the original conflict studies, might have provided a more adequate test than examining a pronounced reception effect for the ingroup in a *helping* context.

In sum then, the present results might support the notion of a general reception effect as obtained by Oeberst and Matschke (2017). That is, impressions of the group mentioned first may, in fact, be independent of the recipients' own group. On the other hand, the present study might not have provided the crucial hypothesis test that we had aimed at. Consequently, it might be worthwhile to once more examine potentially pronounced reception effects for ingroup members in a more competitive setting (vs. helping) and in a way that allows for insight into participants' own group memberships, their social identifications as well as their perceptions and motivations.

### Limitations

As with every scientific approach, the lost letter technique has its strengths and weaknesses (which was already noted in the first study by Milgram et al., 1965). Some of its weaknesses, however, clearly limit the conclusions that may be possibly drawn from our study. One major disadvantage is doubtlessly the lack of knowledge about participants. Their age, gender, cultural background, and social identification naturally remained completely unknown. For the present purpose, this made it impossible to test empirical relations between potential correlates and ethnocentric behavior or possible mediators or moderators related to the finder's characteristics. Especially, the finders' nationalities and their social identifications with the German and Turkish populations would have been of particular interest. Instead, we only have mere probabilities that are based on the population rates of Berlin's districts. But of course, the fact that Germans represent the largest group of inhabitants in Berlin and the district of Lichtenberg, in particular, does not *necessarily* come along with a higher probability of coming across a lost letter. After all, this was only the case if

nationality was not confounded with other relevant variables (e.g., mobility behavior, attention to the environment), which we simply do not know. Note, however, that all lost letter studies share the problem of unidentifiable participants. It is of particular relevance, however, in the case of null results as the paradigm precludes in-depth analyses. The same holds for information about how many people passed a letter *without* returning it (which might have been different for the different experimental conditions without being visible in the data, although return *time* should be positively correlated with it).

A study-specific limitation could have been the Corona pandemic. Of course, one can only speculate, but it cannot be denied that at the time of the experiment's conduction people were highly sensitized regarding a viral infection and the Corona virus' lifespan on different surfaces due to the media coverage at that time and the months before. Note, however, that this parameter was constant for all conditions and, thus, should not have affected our hypothesis test. Moreover, the overall return rate in our study was not at all below the range of those previously obtained (see Volker et al., 2015, for a review, and Hellmann et al., 2020, for the most recent lost letter study conducted in Germany before the Corona pandemic). Consequently, these special circumstances unlikely exerted any special effects in our study.

### Conclusion

The present study tested whether the order of two groups on allegedly lost letters affected the letters' return rates. It did not. Putting the German or the Turkish group first on an envelope did not yield different helping behavior (in terms of returning that envelope). While this is in line with the previously found *general* reception effect (Oeberst & Matschke, 2017) the limitations of the lost letter paradigm preclude more precise and deeper insights into potential alternative explanations. Also, the choice of a helping context may have countered potentially pronounced reception effects for ingroup members. Therefore, it is still too early for final conclusions. Still, the obtained null-results are also a valid possibility, particularly as they are in line with prior research. And with regard to helping behavior in the real world, however, our null results are actually good news: People posted the letters regardless of the way that groups were ordered on the envelope. The missing differences in return rates are highly desirable. After all, they indicate a *lack* of discrimination.

### References

- Ahmed, A. M. (2010). Muslim discrimination: Evidence from two lost-letter experiments. *Journal of Applied Social Psychology*, 40, 888–898. <https://doi.org/10.1111/j.1559-1816.2010.00602.x>
- Amt für Statistik Berlin-Brandenburg (2020a). Einwohnerregisterstatistik Berlin. Bestandsdaten Einwohnerregister Berlin [Population index statistics. Inventory data population index Berlin]. Retrieved from [https://download.statistik-berlin-brandenburg.de/cc2b96a4e11ca6b2/6bde1fe55f26/MD\\_19211\\_2019.pdf](https://download.statistik-berlin-brandenburg.de/cc2b96a4e11ca6b2/6bde1fe55f26/MD_19211_2019.pdf)



- Amt für Statistik Berlin-Brandenburg (2020b). Statistischer Bericht. Einwohnerinnen und Einwohner im Land Berlin am 31. Dezember 2019 [Statistical report. Residents of the federal state Berlin on December 31, 2019]. Retrieved from [https://www.statistik-berlin-brandenburg.de/publikationen/stat\\_berichte/2020/SB\\_A01-05-00\\_2019h02\\_BE.pdf](https://www.statistik-berlin-brandenburg.de/publikationen/stat_berichte/2020/SB_A01-05-00_2019h02_BE.pdf)
- Arbeitsagentur (2020). Arbeitsmarkt im Überblick – Berlin Bundesland [Labour market at a glance – Federal state of Berlin]. Retrieved from [https://statistik.arbeitsagentur.de/Auswahl/raeumlicher-Geltungsbereich/Politische-Gebietsstruktur/Bundeslaender-SGBII/Berlin-SGBII.html?nn=25568&year\\_month=202007](https://statistik.arbeitsagentur.de/Auswahl/raeumlicher-Geltungsbereich/Politische-Gebietsstruktur/Bundeslaender-SGBII/Berlin-SGBII.html?nn=25568&year_month=202007)
- Asbrock, F. (2010). Stereotypes of social groups in Germany in terms of warmth and competence. *Social Psychology*, 41, 76–81. <https://doi.org/10.1027/1864-9335/a000011>
- Asbrock, F., Lemmer, G., Becker, J. C., Koller, J., & Wagner, U. (2014). “Who are these foreigners anyway?” The content of the term foreigner and its impact on prejudice. *Sage Open*, 4, 1–8. <https://doi.org/10.1177/2158244014532819>
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258–290. <https://doi.org/10.1037/h0055756>
- Baldassarri, D. (2020). Market integration accounts for local variation in generalized altruism in a nationwide lost-letter experiment. *PNAS*, 117, 2858–2863. <https://doi.org/10.1073/pnas.1819934117>
- Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, 140, 1556–1581. <https://doi.org/10.1037/a0037737>
- Benor, S., & Levy, R. (2006). The chicken or the egg? A probabilistic analysis of English binomials. *Language*, 82, 233–278. <https://doi.org/10.1353/lan.2006.0077>
- Berger, R., & Berger, J. (2019). Islamophobia or threat to secularization? Lost letter experiments on the discrimination against Muslims in an urban area of Switzerland. *Swiss Journal of Sociology*, 45, 83–105. <https://doi.org/10.2478/sjs-2019-0005>
- Bettinsoli, M. L., Maass, A., Kashima, Y., & Suitner, C. (2015). Word-order and causal inference: The temporal attribution bias. *Journal of Experimental Social Psychology*, 60, 144–149. <https://doi.org/10.1016/j.jesp.2015.05.011>
- Boroditsky, L. (2003). Linguistic relativity. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (pp. 917–922). London, UK: McMillan.
- Bundesamt für Migration und Flüchtlinge (2019). Minas. Atlas über Migration, Integration und Asyl [Minas. Atlas for migration, integration, and asylum]. Retrieved from [https://www.bamf.de/SharedDocs/Anlagen/DE/Statistik/Migrationsatlas/migrationsatlas-2019-11.pdf?\\_\\_blob=publicationFile&v=3](https://www.bamf.de/SharedDocs/Anlagen/DE/Statistik/Migrationsatlas/migrationsatlas-2019-11.pdf?__blob=publicationFile&v=3)
- Bundesnetzagentur (2017). Tätigkeitsbericht Post 2016/2017 [Activity report German Post 2016/2017]. Retrieved from [https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Allgemeines/Bundesnetzagentur/Publikationen/Berichte/2017/TB\\_Post20162017.pdf?\\_\\_blob=publicationFile%20&v=1](https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Allgemeines/Bundesnetzagentur/Publikationen/Berichte/2017/TB_Post20162017.pdf?__blob=publicationFile%20&v=1)
- Carpusor, A. G., & Loges, W. E. (2006). Rental discrimination and ethnicity in names. *Journal of Applied Social Psychology*, 36, 934–952. <https://doi.org/10.1111/j.0021-9029.2006.00050.x>
- Cooper, W. E., & Ross, J. R. (1975). World order. In R. E. Grossman, L. J. San, & T. J. Vance (Eds.), *Papers from the parasession on functionalism* (pp. 63–111). Chicago Linguistic Society.
- Dean, M. L. (1980). Presentation order effects in product taste tests. *Journal of Psychology*, 105, 107–110. <https://doi.org/10.1080/00223980.1980.9915138>
- Dudenredaktion (n.d.). Deutsch-türkisch, deutschtürkisch [German(-) Turkish]. Retrieved May 17, 2020, from Duden website: <https://www.duden.de/node/134131/revision/134167>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fiedler, K. (2008). Language: A toolbox for sharing and influencing social reality. *Perspectives on Psychological Science*, 3, 38–47. <https://doi.org/10.1111/j.1745-6916.2008.00060.x>
- Field, A. (2013). *Discovering statistics using IBM SPSS Statistics*. Sage.
- Fitrianto, A., & Cing, N. M. (2014). Empirical distributions of parameter estimates in binary logistic regression using bootstrap. *International Journal of Mathematical Analysis*, 8, 721–726. <https://doi.org/10.12988/ijma.2014.4394>
- Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (2011). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology*, 4, 1–26. <https://doi.org/10.1080/14792779343000004>
- GESIS–Leibniz Institute for the Social Sciences (2017). ALLBUS/GGSS 2016. Allgemeine Bevölkerungsumfrage der Sozialwissenschaften [German general social survey 2016] (ZA5250 data file version 2.1.0). <https://doi.org/10.4232/1.12796>
- Gundel, J. K. (1988). Universals in topic-comment structure. In M. Hammond, E. A. Moravcsik, & J. Wirth (Eds.), *Studies in syntactic typology* (pp. 209–240). John Benjamins. <http://dx.doi.org/10.1075/tsl.17.16gun>
- Hegarty, P. (2014). Ladies and gentlemen: Word order and gender in English. In G. G. Corbett (Ed.), *The expression of gender* (pp. 69–86). de Gruyter.
- Hegarty, P., Mollin, S., & Foels, R. (2016). Binomial word order and social status. In H. Giles & A. Maass (Eds.), *Advances in intergroup communication* (pp. 118–135). Peter Lang.
- Hegarty, P., Watson, N., Fletcher, L., & McQueen, G. (2011). When gentlemen are first and ladies are last: Effects of gender stereotypes on the order of romantic partners’ names. *British Journal of Social Psychology*, 50, 21–35. <https://doi.org/10.1348/014466610X486347>
- Hellmann, J. H., Berthold, A., Rees, J. H., & Hellmann, D. F. (2015). “A letter for Dr. outgroup”: On the effects of an indicator of competence and chances for altruism toward a member of a stigmatized out-group. *Frontiers in Psychology*, 6, Article 1422. <https://doi.org/10.3389/fpsyg.2015.01422>
- Hellmann, J. H., Forthmann, B., Knausenberger, J., Hellmann, D. F., Rees, J. H., Gansel, E., ... Echterhoff, G. (2020). Support for refugee integration in West and East Germany: Results from two lost letter studies. *Social Psychology*, 51, 106–115. <https://doi.org/10.1027/1864-9335/a000397>
- Hesterberg, T. C. (2015). What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum. *The American Statistician*, 69, 371–386. <https://doi.org/10.1080/00031305.2015.1089789>
- Holland, J., Silva, A. S., & Mace, R. (2012). Lost letter measure of variation in altruistic behaviour in 20 neighborhoods. *PLoS ONE*, 7, Article e43294. <https://doi.org/10.1371/journal.pone.0043294>
- Johnson-Laird, P. N. (1968). The interpretation of the passive voice. *The Quarterly Journal of Experimental Psychology*, 20, 69–73. <http://dx.doi.org/10.1080/14640746808400129>
- Klink, A., & Wagner, U. (1999). Discrimination against ethnic minorities in Germany: Going back to the field. *Journal of Applied Social Psychology*, 29, 402–423. <https://doi.org/10.1111/j.1559-1816.1999.tb01394.x>

- Koopmans, R., & Veit, S. (2014). Cooperation in ethnically diverse neighborhoods: A lost-letter experiment. *Political Psychology*, 35, 379–400. <https://doi.org/10.1111/pops.12037>
- Kraftfahrtbundesamt (2019). Bestand an Kraftfahrzeugen und Kraftfahrzeuganhängern nach Gemeinden, 1. Januar 2019 (FZ 3) [Stock of vehicles and trailers by registration districts, January 1, 2019]. Retrieved from [https://www.kba.de/SharedDocs/Publikationen/DE/Statistik/Fahrzeuge/FZ/2019/fz3\\_2019\\_xlsx.xlsx?\\_\\_blob=publicationFile&v=4](https://www.kba.de/SharedDocs/Publikationen/DE/Statistik/Fahrzeuge/FZ/2019/fz3_2019_xlsx.xlsx?__blob=publicationFile&v=4)
- Kremer, J., Barry, R., & McNally, A. (1986). The misdirected letter and the quasi-questionnaire: Unobtrusive measures of prejudice in Northern Ireland. *Journal of Applied Social Psychology*, 16, 303–309. <https://doi.org/10.1111/j.1559-1816.1986.tb01142.x>
- Levine, M., Cassidy, C., Brazier, G., & Reicher, S. (2002). Self-categorization and bystander non-intervention: Two experimental studies. *Journal of Applied Social Psychology*, 32, 1452–1463. <https://doi.org/10.1111/j.1559-1816.2002.tb01446.x>
- Levine, M., Prosser, A., Evans, D., & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin*, 31, 443–453. <https://doi.org/10.1177/0146167204271651>
- Malkiel, Y. (1959). Studies in irreversible binomials. *Lingua*, 8, 113–160. [https://doi.org/10.1016/0024-3841\(59\)90018-x](https://doi.org/10.1016/0024-3841(59)90018-x)
- McGuire, W. J., & McGuire, C. V. (1992). Psychological significance of seemingly arbitrary word-order regularities: The case of kin pairs. In G. R. Semin & K. Fiedler (Eds.), *Language, interaction and social cognition* (pp. 214–236). Sage.
- Milgram, S. (1977). *The individual in a social world. Essays and experiments* (2nd ed.). McGraw-Hill.
- Milgram, S., Mann, L., & Harter, S. (1965). The lost-letter technique: A tool of social research. *Public Opinion Quarterly*, 29, 437–438. <https://doi.org/10.1086/267344>
- Mollin, S. (2014). *The (ir)reversibility of English binomials. Corpus, constraints, developments*. Amsterdam, The Netherlands: John Benjamins.
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: a review. *European Journal of Social Psychology*, 15, 263–280. <https://doi.org/10.1002/ejsp.2420150303>
- Oeberst, A., & Matschke, C. (2017). Word order and world order: Titles of intergroup conflicts may increase ethnocentrism by mentioning the in-group first. *Journal of Experimental Psychology: General*, 146, 672–690. <https://doi.org/10.1037/xge0000300>
- Osborne, J. W. (2006). Bringing balance and technical accuracy to reporting odds ratios and the results of logistic regression analyses. *Practical Assessment, Research, and Evaluation*, 11, Article 7. <https://doi.org/10.7275/z7we-dg46>
- Petrykowski, K., Davis, K., Brown, T., Hastings, S. O., Butler, J., & Pryor, B. (2010). Perceived Arab ethnicity on willingness to help: An application of the lost-letter technique. *North American Journal of Psychology*, 12, 389–400. Retrieved from [http://www.academia.edu/download/58279804/Perceived\\_Arab\\_Ethnicity\\_North\\_American\\_Journal\\_of\\_Psychology\\_2010.pdf](http://www.academia.edu/download/58279804/Perceived_Arab_Ethnicity_North_American_Journal_of_Psychology_2010.pdf)
- Phillips, W., Gharib, A., & Davis, M. (2016). An international lost letter study: Measuring attitudes toward Middle-Easterners. In G. T. Papanikos (Ed.), *An anthology of social themes* (pp. 183–192). Retrieved from [https://www.academia.edu/41545976/An\\_Anthology\\_of\\_Social\\_Themes?auto=download](https://www.academia.edu/41545976/An_Anthology_of_Social_Themes?auto=download)
- Reid, M. (2018). Lost letters: Using the lost-letter technique to teach social research methods. *Teaching Sociology*, 46, 356–363. <https://doi.org/10.1177/0092055X18776728>
- Rousseeuw, P. J., Pernet, C. R., & Wilcox, R. R. (2019). A practical introduction to the bootstrap: A versatile method to make inferences by using data-driven simulations. *PsyArXiv*. <https://doi.org/10.31234/osf.io/h8ft7>
- Senatsverwaltung für Gesundheit und Soziales (2014). *Handlungsorientierter Sozialstrukturatlas Berlin 2013 [Action orientated social structure atlas Berlin 2013]*. Retrieved from [https://www.berlin.de/sen/gessoz/\\_assets/publikationen/gesundheits-und-sozialberichterstattung/gesundheitsberichterstattung-epidemiologie/spezialberichte/gbe\\_spezial\\_2014\\_1\\_ssa2013.pdf](https://www.berlin.de/sen/gessoz/_assets/publikationen/gesundheits-und-sozialberichterstattung/gesundheitsberichterstattung-epidemiologie/spezialberichte/gbe_spezial_2014_1_ssa2013.pdf)
- Senatsverwaltung für Stadtentwicklung und Wohnen (2019a). *Gesamtindex Soziale Ungleichheit (Status/Dynamik-Index) 2019 auf Ebene der Planungsräume (31.12.2018) [Overall index of social inequality (status/dynamic index) 2019 based on planning areas (December 31, 2018)]*. Retrieved from [https://www.stadtentwicklung.berlin.de/planen/basisdaten\\_stadtentwicklung/monitoring/download/2019/tab/1.SDI\\_MSS2019.pdf](https://www.stadtentwicklung.berlin.de/planen/basisdaten_stadtentwicklung/monitoring/download/2019/tab/1.SDI_MSS2019.pdf)
- Senatsverwaltung für Stadtentwicklung und Wohnen (2019b). *Monitoring Soziale Stadtentwicklung 2019 [Monitoring social urban development 2019]*. Retrieved from [https://www.stadtentwicklung.berlin.de/planen/basisdaten\\_stadtentwicklung/monitoring/download/2019/MSS\\_Fortschreibung2019\\_Langfassung.pdf](https://www.stadtentwicklung.berlin.de/planen/basisdaten_stadtentwicklung/monitoring/download/2019/MSS_Fortschreibung2019_Langfassung.pdf)
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. Retrieved from SSRN website: <https://ssrn.com/abstract=2160588>
- Sobkowiak, W. (1993). Unmarked-before-marked as a freezing principle. *Language and Speech*, 36, 393–414. <https://doi.org/10.1177/002383099303600403>
- Stevens, J. R., & Duque, J. F. (2019). Order matters: Alphabetizing in-text citations biases citation rates. *Psychonomic Bulletin & Review*, 26, 1020–1026. <https://doi.org/10.3758/s13423-018-1532-8>
- Volker, B., Mollenhorst, G., Steenbeek, W., Schutjens, V., & Flap, H. (2015). Lost letters in Dutch neighborhoods: A field experiment on collective efficacy. *Social Forces*, 94, 953–974. <https://doi.org/10.1093/sf/sov106>
- Westlake, G., Coall, D., & Grueter, C. C. (2019). Educational attainment is associated with unconditional helping behaviour. *Evolutionary Human Sciences*, 1, Article e15. <https://doi.org/10.1017/ehs.2019.16>
- Wright, D. B., London, K., & Field, A. (2011). Using bootstrap estimation and the plug-in principle for clinical psychology data. *Journal of Experimental Psychopathology*, 2, 252–270. <https://doi.org/10.5127/jep.013611>
- Wulff, S. (2003). A multifactorial corpus analysis of adjective order in English. *International Journal of Corpus Linguistics*, 8, 245–282. <https://doi.org/10.1075/ijcl.8.2.04wul>