**Supplementary Material for**

**Populists’ Use of Nostalgia: A Supervised Machine Learning Approach**

**Validation via Manual Content Analysis**

We applied manual coding for three reasons: To (1) scrutinize our ground-truth data construction, (2) test the validity of our classifier, and (3) explore themes of personal and collective nostalgia in the examined texts.

**Analytical Approach**

We employed a structuring qualitative content analysis (Mayring & Fenzl, 2014) for the manual coding of the data. In such an analysis, the communication content that is analyzed is systematically assigned to pre-defined categories in a codebook. The codebook is developed based on the literature. For each category, it includes a description together with prototypical examples for the respective category and coding rules guiding the assignment. The codebook can be refined in a pilot-phase to include categories or coding rules that emerge from the material. Before the actual coding, the intersubjective comprehensibility of the codebook must be established via intercoder reliability checks. To this end, different reliability coefficients exist that measure agreement between at least two coders on a randomly sampled subset of the material (Lacy et al., 2015).

**Database**

We manually coded two data sets. The essay data set included all 520 essays that were used to train and evaluate the classifier. The Facebook data set included 125 randomly sampled political Facebook posts from the classifier application study. These posts had been classified as nostalgic versus non-nostalgic by the classifier with a probability of > 70%.

**Codebook Development**

Our codebook had two aims. First, we aimed at obtaining an overall classification of the texts as being nostalgic or not – a human judgement comparable to the one by the supervised machine learning algorithm employed in this study. Second, we aimed to explore personal and collective nostalgia in the texts. This part of the codebook was guided by prior content analyses of nostalgia in everyday and political communication (Holak & Havlena, 1992; Menke & Wulf, 2021; Szabó & Kiss, 2022). We coded for the subject of the nostalgic memory (personal or collective) and a set of nostalgic objects (e.g., social relationships, media experiences) and restorative appeals (Table S1).

We established intercoder reliability based on an iterative refinement procedure. First, two independent coders coded 40 essays that were randomly selected to represent both nostalgic essays (i.e., written by participants whose self-reported state nostalgia was above the scale mean) and non-nostalgic essays (i.e., written by participants who reported state nostalgia below the scale mean). Then we checked the intercoder reliability via the tidycomm shinyapp (<https://joone.shinyapps.io/icr_web/>). Given that we expected some categories to be infrequent in the essays, we used Brennan-Predigers’ κ (Brennan & Prediger, 1981) as a measure for intercoder reliability. We chose this coefficient because it outperforms Krippendorfs’ alpha in cases of imbalanced data (i.e., when the characteristic of interest is infrequent in the data). We considered values >.75 satisfactory.

Seven categories had κ values below this threshold. We discussed them in an extensive coding session in which we jointly coded five so-far unseen posts with a revised version of the codebook that included more detailed coding rules. Next, two independent coders coded another 40 posts with the revised categories resulting in a high intercoder-agreement for all categories, all Brennan-Prediger’s κ > .89. Table S1 shows the final categories and coding rules. We provide the German version of the codebook including the examples per category at the Open Science Framework (https://osf.io/gu92j/?view\_only=2f165bd0cb22423fa658abc8131e0778).

**Table S1.** *Codebook for the Manual Content Analysis*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **Definition** | **Coding rules** | **%** | **κ** |
| ID | Participant number of the essay respectively URL of the Facebook post |  |  |  |
| Coder Id | Initials of the coder |  |  |  |
| Theme | Keywords summarizing the content of the post | Open description of the content in a few keywords |  |  |
| Perceived nostalgia | Global evaluation of the text as expressing or potentially eliciting nostalgia. | Nostalgia is the bittersweet memory, the sentimental longing for a past, which is perceived as positive(er), beautiful(er), familiar(er), and possibly more glorious than the present.  That is, the text must refer to an emotionally positive(er) past that seems desirable. This impression can also be conveyed by lamenting the comparably worse present. | .95 | .90 |
| Past oriented | The text refers to the past | Code when the text refers to the past | .90 | .80 |
| Subjects: personal | The text entails personal memories or experiences | Code when the text describes personal memories or experiences | 1 | 1 |
| Subjects: A younger self | The text refers to a younger version of oneself | Coded when the text refers to a younger version of the author | .75 | .93 |
| Objects: Social relationships | The text refers to social interactions with close others | Coded when one focus of the text is the interaction with close others such family, friends, partners | .78 | .94 |
| Objects: Setting | The text refers to specific settings | Coded when one focus of the text is the settings (e.g., because they are particularly positive, impressive, awe inducing or negative) | .73 | 1 |
| Objects: Animals | The text refers to pets or other animals | Coded when one focus of the text is on animals such as pets. | .92 | .85 |
| Objects: Non-vivid objects | The text refers to non-vivid objects such as vehicles or food | Coded when one focus of the text are specific non-vivid objects (e.g., because they are particularly positive, impressive, awe inducing or negative), the category is not coded when these objects are only used (e.g., sitting in the car) but when they valuable (e.g., grandfathers’ old car). | .78 | 1 |
| Objects: Media | The text refers to specific media experiences | Coded when one focus of the text are specific media-related experiences, for instance specific devices, TV shows, games, beloved media characters and so on. | .90 | .80 |
| Subject: Collective | The text entails collective memories | Coded when the text describes memories of the past of a stable collective/social identity (i.e., a social category such as “the Germans", "the Berlin people", "the students"). It is about identities one could use to describe him- or herself, e.g., I am XYZ. For instance, I am a fan of …, I am German, I am Muslim etc. | .78 | .89 |
| Collective Identity | Which collective is addressed? | Open description of the collective identity addressed. | 1 |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **Definition** | **Coding rules** | **%** | **κ** |
| Restorative nostalgia | Texts that present the past as an ideal time that should be restored | Coded when the text contains references to an ideal world ("heartland"), which―as opposed to utopian ideas―is constructed retroactively from the past. In other words, a vision derived from the past that is projected onto the present as "what has been lost". Often also: lamenting the present perceived as more negative. | 1 | 1 |
| Conservative nostalgia | Texts that transmit restorative nostalgia for a more conservative time | Coded when the restorative nostalgia refers to a conservative past (e.g., more traditional gender roles, a more homogenous society with less diversity (e.g., less migration)). | 1 | 1 |
| Liberal nostalgia | Texts that transmit restorative nostalgia for a more liberal time | Coded when the restorative nostalgia refers to a liberal past (e.g., less social inequality, better nature, more liberal sexual norms (e.g., in the 70ths) | 1 | 1 |
| Explicit nostalgia expression | Texts that express bittersweet emotions about the past | Coded when the text contains descriptions of (a) positive (joy, pleasure, love, well-being) AND (b) negative (anger, sadness, disappointment, longing) emotions? | .98 | .95 |
| Explicit nostalgia: Word | Texts that include the word nostalgia | Coded when the text includes the word nostalgia | 1 | 1 |
| Miscellaneous | Open category for further comments |  |  |  |
| *Notes*. Translated for publication purposes and slightly edited for consistency in the descriptions. | | | | |

**Index construction**

To test the prevalence of personal and collective nostalgia in our data, we calculated two indices based on the human coding that summarized personal and collective nostalgia, respectively. Both indices used the following formula:

= the single themes of either personal or collective nostalgia

= references to the past

= the human perception of the text as being nostalgic (versus not).

Thus, the *personal nostalgia index* counted personal nostalgia when the text referred to the past in a nostalgic manner *and* entailed any of the themes of personal nostalgia. Themes of personal nostalgia were counted, when the subject of the nostalgic memory was a personal experience or memories of a younger self or when typical objects of personal nostalgia—such as social relationships, animals, and pets, inanimate objects, or media experiences—were mentioned. The personal nostalgia index was zero when the respective text either did not refer to the past in a nostalgic manner or did not entail any of these personal nostalgia themes.

The *collective nostalgia index* counted collective nostalgia when the text referred to the past in a nostalgic manner *and* the subject of the text was a collective identity or the text expressed restorative appeals. The collective nostalgia index was zero when the texts either did not refer to the past in a nostalgic manner or did not include any of the themes of collective nostalgia.

**Results**

**S.1 Comparing Self-Reported and Manually Coded Nostalgia in the Essay Data**

Our first analysis used self-reported nostalgia as label for the ground-truth data. As detailed in the main text, we labeled as nostalgic essays that were written by participants with self-reported state nostalgia above the scale midpoint, and we labeled as non-nostalgic essays that were written by participants with self-reported state nostalgia below the scale midpoint. We excluded essays written by participants with self-reported nostalgia corresponding to the scale midpoint. We compared these labels to the perception of the essays as nostalgic (vs. non-nostalgic) by the human coders. Essays labeled as nostalgic (vs. non-nostalgic) based on participants’ self-report were more likely to be perceived as nostalgic (vs. non-nostalgic) by the coder, χ²(1)=38.31, *p*<.001. Essays that were written by someone reporting above-mean levels of nostalgia had 2.12-times higher odds to be perceived as nostalgic (vs. not-nostalgic) by the human coder. Essays that were written by someone reporting below-mean levels of nostalgia had 1.63-times higher odds to be perceived as non-nostalgic (vs. nostalgic) by the human coder. Overall, the label based on self-reported nostalgic affect and manual coding agreed in 66% of the essays (Table S2).

**Table S2**

*Confusion Matrix for Essays Labeled as Nostalgic or Non-Nostalgic Based on Self-Report Versus the Manual Coding of the Essays*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Manually coded as non-nostalgic | Manually coded as nostalgic | Σ |
| Low self-reported nostalgia | 96 | 59 | 155 |
| High self-reported nostalgia | 113 | 239 | 352 |
| Σ | 209 | 298 | 507 |

**S.2 Comparing Machine Classified Nostalgia and Manually Coded Nostalgia in the Evaluation Data**

This analysis compared the performance of our classifier to the human coding of the essays in the hold-out evaluation dataset. Human and algorithmic evaluation were marginally associated, in a Fischer’s test, *p* = .07. Essays that were deemed nostalgic by the classifier had 1.9-times higher odds to be perceived as nostalgic (vs. non-nostalgic) by the human coder. Essays that were deemed non-nostalgic by the classifier had 1.6-times higher odds to be perceived as non-nostalgic (vs. nostalgic) by the human coder (Table S3).

**Table S3**

*Confusion Matrix for Essays Classified as Nostalgic and Non-Nostalgic in the Evaluation Data Versus the Manual Coding of the Essays*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Manually coded as non-nostalgic | Manually coded as nostalgic | Σ |
| Classified as non-nostalgic | 8 | 5 | 13 |
| Classified as nostalgic | 30 | 58 | 88 |
| Σ | 38 | 63 | 101 |

**S.3. Exploring Themes of Personal and Collective Nostalgia in the Essays**

To test whether essays classified as nostalgic (vs. non-nostalgic) by the random-forest classifier differed in their coverage of personal and collective nostalgia, we used the classification of the essays as nostalgic (vs. non-nostalgic) as independent variable and our indices of personal and collective nostalgia as dependent variables in two independent *t*-tests with Welch-correction.

Essays classified as nostalgic scored higher on the index of personal nostalgia (*M* = 2.21, *SD* = 2.05) than essays classified as non-nostalgic (*M* = 0.47, *SD* = 1.25), *t*(406.82) = 13.11, *p* < .001, *d* = 1.01. Essays classified as nostalgic also entailed more collective nostalgia (*M* = 0.13, *SD* = 0.39) than essays classified as non-nostalgic (*M* = 0.06, *SD* = 0.27), *t*(734.72) = 25.48, *p* < .001, *d* = 0.27.

**S.4. Exploring Themes of Personal and Collective Nostalgia in the Facebook Posts**

To test whether political Facebook posts classified as nostalgic (vs. non-nostalgic) by the random-forest classifier covered more themes of personal and collective nostalgia, we used the classification of the posts as nostalgic (vs. non-nostalgic) as independent variable and our indices of personal and collective nostalgia as dependent variables in two independent *t*-tests with Welch-correction.

Facebook posts classified as nostalgic included more personal nostalgia (*M* = 0.41, *SD* = 0.87) than posts classified non-nostalgic (*M* = 0.02, *SD* = 0.15), *t*(245.46) = 2.32, *p* = .02, *d* = 0.61. Facebook posts classified as nostalgic tended to include more collective nostalgia (*M* = 0.44, *SD* = 0.76) than posts classified as non-nostalgic (*M* = 0.05, *SD* = 0.23), although the difference was only marginally significant, *t*(247.47) = 1.83, *p* = .07, *d* = 0.69.

**S.5. Distribution of Facebook Posts Across Parties**

Most posts in our database were uploaded by the governing conservative parties CDU and CSU (Table S4). The Green party was least active on Facebook during the examined time period.

**Table S4**

*Absolute Number of Facebook Posts per Party*

|  |  |  |
| --- | --- | --- |
| **Party Ideology** | **Party Name** | ***n*** |
| Right-wing populist, far right | Alternative for Germany (AfD) | 405 |
| Green politics, center left | Alliance 90/The Greens (Bündnis 90/Die Grünen) | 334 |
| Christian democratic, liberal conservative | Christian Democratic Union (CDU) | 637 |
| Christian democratic, conservative | Christian Social Union (CSU) | 991 |
| Liberal, free democratic | Free Democratic Party (FDP) | 805 |
| Democratic socialism, left-wing populism | The Left (Die Linke) | 413 |
| Social-democratic | Social Democratic Party of Germany (SPD) | 437 |
| *Notes*. *N* = 4,022 Facebook posts uploaded to the official party pages between January 1st and December 31st, 2019. | | |