People's Beliefs About Pronouns Reflect Both the Language They Speak and Their Ideologies

April H. Bailey[1]*, Robin Dembroff[2], Daniel Wodak[3], Elif G. Ikizer[4], Andrei Cimpian[5]

[1]Department of Psychology, University of New Hampshire; Durham, NH, 03823, USA.
[2]Department of Philosophy, Yale University; New Haven, CT, 06520, USA.
[3]Department of Philosophy, University of Pennsylvania; Philadelphia, PA, 19104, USA.
[4]Department of Psychology, University of Wisconsin-Green Bay, Green Bay, WI, 54311, USA.
[5]Department of Psychology, New York University, New York, NY, 10012, USA.

* Corresponding author. Email: april.bailey@unh.edu

In press, *Journal of Experimental Psychology: General*

Abstract

Pronouns often convey information about a person's social identity (e.g., gender). Consequently, pronouns have become a focal point in academic and public debates about whether pronouns should be changed to be more inclusive, such as for people whose identities do not fit current pronoun conventions (e.g., gender non-binary individuals). Here, we make an empirical contribution to these debates by investigating which social identities lay speakers think that pronouns should encode and why. Across four studies, participants were asked to evaluate different types of real and hypothetical pronouns, including binary gender pronouns, race pronouns, and identity-neutral pronouns. We sampled participants from two languages with different pronoun systems: English (N = 1,120) and Turkish (N = 260). English pronouns commonly denote binary gender (e.g., *he* for men), whereas Turkish pronouns are identity-neutral (e.g., *o* for anyone). Participants' reasoning about pronouns reflected both a familiarity preference (i.e., participants preferred the pronoun type used in their language) and—critically—participants' *social ideologies*. In both language contexts, participants' ideological beliefs that social groups are inherently distinct (essentialism) and should be hierarchal (social dominance orientation) predicted relatively greater endorsement of binary gender pronouns and race pronouns. A preregistered experimental study with an English-speaking sample showed that the relationship between ideology and pronoun endorsement is causal: Ideologies *shape* attitudes toward pronouns. Together, the present research contributes to linguistic and psychological theories concerning how people reason about language and informs policy-relevant questions about whether and how to implement language changes for social purposes.

Keywords: language; gender; race; essentialism; social dominance orientation

Public Significance Statement

Because pronouns encode information about social identities such as gender, they are at the center of ongoing public debates. For example, should society change pronoun use to be more inclusive of people whose identity doesn't fit current pronoun conventions? We brought empirical evidence to bear on these debates by investigating beliefs about which social identities pronouns should encode (if any) and why. We recruited speakers of English and Turkish: English commonly uses binary gender pronouns (*he* for men, *she* for women), whereas Turkish uses identity-neutral pronouns (*o* for everyone). We found that pronoun endorsement reflected not just a familiarity preference (with English speakers preferring binary gender pronouns and Turkish speakers identity-neutral pronouns) but also *ideological commitments*. The more English- and Turkish-speakers believed social groups should be distinct and unequal, the more they endorsed binary gender pronouns and race pronouns. The link between ideology and pronoun endorsement was causal.

People's Beliefs About Pronouns Reflect Both the Language They Speak and Their Ideologies

Pronouns based on race would breed racism.

—Anonymous Participant

Similar to other function words (e.g., conjunctions, articles), pronouns are some of the most frequent lexical items in any language. However, unlike other function words, pronouns can encode social information, which—given how frequent pronouns are—may shape how people reason about their social world. In some cases, the social information encoded by pronouns conveys uncontroversial relationships between the speaker and the addressee. For instance, the Hindi second-person singular pronoun *aap* signals a respectful relationship and implies the speaker is younger than the addressee(s). In other cases, pronouns convey deeper commitments and imply the existence of certain social identities and the non-existence of others. For instance, English third-personal singular personal pronouns commonly denote binary gender identities[1] (*she* for women, *he* for men)—as do many other languages (e.g., *il* and *elle* in

[1] Our studies focus on third-person singular personal pronouns used to refer to a specific person (i.e., definite use). In English, the two most common pronouns in this category are pronouns that also denote binary gender identities (i.e., *he* for men and *she* for women; Corpus of Contemporary American English, Davies, 2008-2022). However, English third-person pronoun use is somewhat in flux (e.g., Bradley et al., 2019; Bradley, 2020). In addition to its typical use as a plural pronoun, *they* is a common third-person *singular* pronoun as well, but it is usually used to refer to an abstract individual (e.g., a person) as an alternative to masculine generic *he*; this is an indefinite use of singular *they*. Less commonly, *they* is also used as a third-person singular personal pronoun to denote someone with a non-binary gender identity; this is a definite use of singular *they*. *They* is also sometimes used to reject the idea that gender is relevant—for instance, a woman might use singular *they* as an identity-neutral personal pronoun while nevertheless identifying as a woman (definite use). There are also other, less common third-person neopronouns (e.g., *ze*, *hir*) that can be used to denote non-binary gender identities

French)—implying that gender is binary and apparent to others. Given such implied ontological commitments, pronouns are increasingly being debated by academics and the public alike (e.g., Dembroff & Wodak, 2018; Parker et al., 2019). These debates are also fueled by the fact that many other languages, such as Turkish and Finnish, use the same identity-neutral pronoun for everyone (e.g., the Turkish pronoun *o*[2]).

Whereas prior psychological research has investigated people's beliefs about various social identities, including gender (e.g., Dar-Nimrod & Heine, 2011; Eagly & Mladinic, 1994; Fiske et al., 2002; Glick & Fiske, 1997; Haslam et al., 2000), less is known regarding how people reason about which of these identities, if any, should be encoded into the structure of language. To begin to investigate these issues, we measured lay speakers' endorsement of various pronouns that are—or could plausibly be—used in a language, focusing in particular on third-person singular personal pronouns. Specifically, we compared endorsement of binary gender pronouns to pronouns denoting other identities (e.g., race pronouns) and pronouns that do not convey any information about identity (i.e., identity-neutral pronouns). We sampled speakers of two languages with different pronoun systems to assess both language-specific and language-general aspects of people's beliefs about pronouns. In particular, we recruited English-speaking (Studies 1, 2a, and 3) and Turkish-speaking (Study 2b) participants. Because English has binary gender pronouns but Turkish does not, a basic prediction here is that English-speaking participants would prefer binary gender pronouns, whereas Turkish-speaking participants would

(definite use; Baron, 2010). Despite these recent trends, in English binary gender pronouns are still by far the most common definite third-person singular personal pronouns, which is why we focus on them here. In Turkish, these pronouns are commonly identity-neutral.

[2] The Turkish pronoun *o* is used generically to refer to anyone or anything, including anyone of any gender identity and even animals and inanimate objects. As is the case for many generic words that are inclusive in theory (e.g., *person* in English; Bailey et al., 2022), there is evidence that *o* evokes male-defaulting biases (Renström et al., 2022).

prefer identity-neutral pronouns. Our crucial question was whether participants' reasoning about pronouns would show any systematic patterns *beyond* this preference for the familiar, language-specific status quo. Are there language-general factors that influence individual participants' relative endorsement of certain pronouns over others, regardless of which language they speak? In the present research, we explored three such factors: (a) reliance on low-effort cognitive heuristics, (b) certain ideological commitments about social groups, and (c) meta-linguistic beliefs. Before detailing each of these three factors, we provide some background on how languages encode social identities and the debates surrounding this relation.

Language and Social Identity

Language can shape how people think (Boroditsky, 2001; Whorf, 1956), drawing attention to or away from certain distinctions and influencing what concepts are readily accessible for reasoning (e.g., Regier & Kay, 2009; Winawer et al., 2007). Thus, the fact that basic aspects of many languages (e.g., pronouns) encode social identities may make certain social distinctions more salient and accessible than they would otherwise be.

More than any other social identity, binary gender tends to be built into the structure of languages across the world, albeit to different degrees depending on the language (Prewitt-Freilino et al., 2012; Stahlberg et al., 2007). For instance, in Hebrew third-person singular personal pronouns are gendered, but so are verbs, leading to literal English translations such as, "Jessica *woman-drove her* car." At the other extreme, Turkish does not have gendered verbs or even gender pronouns (e.g., the third-person singular pronoun *o* can refer to anyone and anything). English falls in between: It includes binary gender pronouns, but verbs are not marked for gender or any other social identity. These differences in how gender is marked across languages might matter for how people think about gender: People who speak languages that

mark gender more prominently tend to also be more gender-biased (DeFranza et al., 2020; Lewis & Lupyan, 2020; see also Sczesny et al., 2016) and to acquire a conception of binary gender earlier in life (Guiora et al., 1982).

This potential relationship between language and social biases have spurred inter-disciplinary debates about gendered language and particularly binary gender pronouns (Atir, 2022; Bigler & Leaper, 2015; Burgess & Plunkett, 2013; Gabriel et al., 2018; Hofstadter, 1985; Plunkett, 2015; Sczesny et al., 2016). Which pronouns *should* languages have? In English, some arguments call for eliminating gender pronouns and using identity-neutral pronouns for everyone (e.g., Dembroff & Wodak, 2018, 2021; Saguy & Williams, 2019). These changes are supported by evidence that binary gendered language brings up gender making it "omnirelevant" (Bigler & Leaper, 2015) and may thus contribute to gender stereotyping and discrimination against women (Bäck et al., 2015; Hilliard & Liben, 2010; Tavits & Pérez, 2019). Binary gender pronouns also marginalize gender non-binary individuals (Hyde et al., 2019; Morgenroth & Ryan, 2020) and contribute to negative experiences of misgendering for everyone (e.g., referring to a man using *she*; Kapusta, 2016). On the other side of the debate, proponents of binary gender pronouns point out that they can make it easier to communicate. Identity-neutral pronouns can create ambiguity about whether one or multiple individuals are being referred to (e.g., "Qi and Jessica went out, and *they* paid") and cause comprehension difficulties (Leventhal et al., 2020; Prasad & Morris, 2020). Binary gender pronouns resolve some (but not all) of these ambiguities.

Here, we contribute a key empirical element that is missing so far from these academic debates: evidence for what lay people think about pronouns and why. Lay speakers create and shape language. Any proposed change to a language requires buy-in from everyday speakers of that language (see Baron, 2010; Vergoossen et al., 2020). Thus, understanding how people

reason about the pronouns that exist or could exist in their language can inform debates about the feasibility of any changes. For example, we might find that English speakers endorse binary gender pronouns just because of familiarity. Or instead, we might find that speakers endorse binary gender pronouns not just because of familiarity but also because of their ideological beliefs about the nature of gender (e.g., that gender is binary). If so, this would suggest that merely increasing familiarity with neutral pronouns might not be sufficient to increase acceptance and use of neutral pronouns. Here, we provide a cross-linguistic examination of the language-specific and -general factors that influence why people prefer some types of pronouns over others.

Language-Specific and -General Influences on Beliefs about Pronouns

People tend to prefer to keep things as they are rather than change them, even when the status quo is unfair, inefficient, or personally harmful (e.g., Jost et al., 2004). People also like things that are familiar, all else equal (e.g., Bornstein, 1989; Zajonc, 1968). With respect to third-person singular personal pronouns, the (familiar) status quo in English is to denote binary gender (e.g., *he* or *she*) but not other social identities. (Notably, while some English speakers have started using pronouns such as *they* to denote non-binary gender identities, such usage is still limited.) In contrast, the status quo in Turkish is to use the same identity-neutral pronoun for everyone. Thus, we hypothesized that participants' pronoun preferences would be language-specific and mirror the status quo pronouns in their language, with English speakers preferring binary gender pronouns and Turkish speakers preferring identity-neutral pronouns. Given the extensive evidence for familiarity and status quo preferences in human psychology, we regard evidence for this hypothesis as a "sanity check"—a means of validating our methodology by producing an expected and theoretically grounded result.

The more novel contribution of the present work lies instead in investigating whether there are *language-general* factors that shape attitudes toward pronouns. Based on prior research and theory, we identified three possible language-general factors and corresponding measures. As we describe each, we also review any prior research and highlight the distinctive contribution to the present work.

Cognitive Heuristics. We investigated whether participants who are susceptible to heuristic reasoning, whether in general or about the status quo specifically, might show stronger endorsement of the types of pronouns in their own language (whatever they may be). Heuristics are "quick and dirty" judgments or decisions made on the basis of limited information and without careful thought (e.g., Kahneman, 2011). Heuristics predispose people to accept the status quo as natural and even good, in part because believing otherwise would require considerable effort (e.g., imagining counterfactual possibilities).

Here, we tested whether relative susceptibility to heuristic reasoning would relate positively to endorsement of status quo pronouns: binary gender pronouns in English and identity-neutral pronouns in Turkish. As our primary measure of susceptibility to heuristic reasoning, we administered the Cognitive Reflection Test (CRT; Frederick, 2005) in both language contexts. This measure uses performance on three simple word problems with intuitively appealing "lure" answers to assess individual differences in participants' reliance on heuristic (vs. analytic) reasoning. In some of our studies, we also assessed two heuristics that have been linked directly to endorsement of the status quo: the *longevity bias* and the *inherence bias*. The longevity bias is the tendency to view the status quo (that is, what is and what has been that way for a long time) as the way things *ought* to be (Blanchar & Eidelman, 2019). Related, the inherence bias leads people to explain the status quo (e.g., roses are a common gift for

Valentine's Day) as a product of inherent or intrinsic features (e.g., roses are beautiful), overlooking plausible but less accessible historical or contextual facts about the relevant phenomenon (e.g., marketing by profit-seeking corporations; Salomon & Cimpian, 2014; Tworek & Cimpian, 2016). The inherence heuristic often prompts people to acquiesce to the status quo (e.g., it's *good* that we give roses for Valentine's Day; Hussak & Cimpian, 2015), including with respect to language (Sutherland & Cimpian, 2015).

In terms of past work, we are not aware of any research that has examined the link between susceptibility to heuristic reasoning per se and attitudes toward pronouns. However, the general issue of cognitive effort has been raised in discussions of why individuals might resist changing language to make it more inclusive. Using new terms is sometimes cognitively taxing (for a review, see Gabriel et al., 2018), so language users might prefer to maintain the status quo. For example, when the definite use-case of the gender-neutral Swedish pronoun *hen* first drew public attention in 2012, reactions were generally negative, but attitudes toward *hen* became more positive with time as *hen* became more mainstream and arguably more cognitively accessible (Bäck et al., 2015; Gustafsson Sendén et al., 2015; Vergoossen et al., 2020). Building on this work, the present research provided a direct cross-linguistic investigation of the link between reliance on heuristic reasoning and relative endorsement of familiar (vs. unfamiliar) pronouns.

Social Ideologies. Reasoning about pronouns might also plausibly be influenced by individuals' social ideologies. Pronouns tend to make the social distinctions they encode salient even in circumstances where they are not relevant, which may in turn reinforce these distinctions and corresponding social hierarchies (Bigler & Leaper, 2015). Binary gender pronouns thus might reinforce sexism and existing hierarchies that disadvantage women. Thus, people whose

ideologies support such group-based social distinctions and/or hierarchies may also be relatively more supportive of identity-based pronouns, including binary gender pronouns.

One such ideology about social distinctions, *essentialism*, captures the belief that social groups are distinct at a "deep," often biological, level. For instance, people who endorse gender essentialism believe that there is an underlying quality or "essence" that makes women and men distinctly different kinds of people (Gelman, 2003; Medin & Ortony, 1989). Often, people think of this underlying essence as being discrete and binary (i.e., you're either a woman or a man), biological, and fixed at birth (Rhodes & Gelman, 2009). With respect to pronouns, individual differences in essentialism about a particular social identity might predict relative endorsement of pronouns that encode that identity. For example, greater gender essentialism might relate to greater endorsement of binary gender pronouns: one for women and one for men.

Another ideology, *social dominance orientation*, is a preference for hierarchal social arrangements that maintain disadvantages for already disadvantaged groups. This construct captures a large amount of the empirical variability in prejudice about disadvantaged groups (Ho et al., 2015; Pratto et al., 1994), including women and transgender people (Perez-Arche & Miller, 2021; Sibley et al., 2007). With respect to pronouns, people high in social dominance orientation might show relatively greater endorsement of pronouns for any hierarchy-relevant social identity, including binary gender pronouns but also pronouns for social identities such as race.

In terms of past work, we are not aware of any research that has examined essentialism and social dominance orientation as predictors of attitudes toward pronouns. However, indirect support for the role of social ideologies is provided by the well-documented link between various forms of sexism and resistance to the introduction of new, gender-egalitarian linguistic forms (including gender-neutral pronouns such as *hen*; e.g., Bradley, 2020; Gustafsson Senden et al.,

2015; 2021; Sarrasin et al., 2012; Sczesny et al., 2015; Vergoossen et al., 2020). In addition to extending the range of social ideologies beyond sexism and the range of language forms beyond gender-inclusive linguistic innovations, the present research provides the first test of whether social ideologies *causally shape* attitudes toward language denoting various social identities.

Meta-linguistic Beliefs. Third and finally, lay beliefs about pronouns may be influenced by people's meta-linguistic beliefs—that is, their ideas about how language works more generally. Specifically, our focus here is on the belief that words somehow "fit" with their referents, even though in reality the mapping between words and referents is largely arbitrary (Ćwiek et al., 2021; Köhler, 1947; Sutherland & Cimpian, 2015). Concerning pronouns, this general meta-linguistic belief that words fit reality (and, conversely, that reality is somehow reflected in linguistic forms) may increase endorsement of pronouns that already exist in participants' language because these pronouns may be assumed to be particularly fitting or appropriate.

In terms of prior work, we are again not aware of any research that has examined the meta-linguistic belief that language fits the world as a predictor of attitudes toward pronouns. However, there is recent evidence regarding a different meta-linguistic belief: linguistic prescriptivism, which views any use of language that departs from (current) established standards as incorrect. Bradley (2020) found that endorsement of prescriptivism predicted more negative attitudes toward the definite use of singular *they* (i.e., to refer to a specific individual of non-binary gender), which is inconsistent with current usage rules. This evidence implicates meta-linguistic beliefs in people's attitudes toward a particular use of a particular pronoun (*they*). Here, we advance this line of inquiry by focusing on the influential belief that words fit reality (Sutherland & Cimpian, 2015) and by expanding the range of pronouns investigated.

The Present Studies

The present studies investigated lay beliefs about a range of third-person singular personal pronouns. We assessed these beliefs among speakers of two languages that were chosen because of the differences in the third-person singular personal pronouns they use: English, which commonly uses binary gender pronouns, and Turkish, which uses identity-neutral pronouns. Our first, most basic hypothesis was that the status quo in each language would influence participants' attitudes toward pronouns, leading them to endorse familiar pronouns. The more novel contribution of the present work lies in our investigation of the social-cognitive factors that might shape participants' attitudes toward pronouns across languages. We examined three such social-cognitive factors: (a) susceptibility to cognitive heuristics, (b) social ideologies, and (c) meta-linguistic beliefs.

In addition to contributing to linguistic and psychological theories concerning how people reason about language, the present research is relevant to socio-ethical questions about whether and how to implement language changes for social purposes (e.g., to promote the well-being of gender minorities, to promote equity between women and men). Language change is driven in part by lay speakers' views about what is acceptable for their language to express. For instance, English speakers might think that using pronouns that encode racial identity would be problematic (see epigraph: "pronouns based on race would breed racism"). Academics have raised similar concerns that binary gender pronouns promote sexism (Lakoff, 2004; Saguy & Williams, 2019), yet lay attitudes towards these pronouns appear to remain positive (Bradley, 2020). The present research provides a cross-linguistic snapshot of speakers' endorsement of different types of pronouns and an examination of *why* attitudes toward these pronouns are positive or negative. Our studies provide valuable insights into whether and why any proposed

changes in pronouns might be embraced or resisted.

Transparency and Openness in Scientific Practices

Complete materials and analysis scripts for all studies are openly available on the Open Science Framework (OSF): https://osf.io/7gjau/?view_only=970ab55b3c9e4bcc92aba678997316b8. This repository also contains the preregistrations and anonymized data for Studies 2a, 2b, and 3. Study 1 was not preregistered, and the data cannot be publicly shared due to the stipulations of the ethics protocol under which it was conducted (which is different from that of the other studies). In all studies, we report all manipulations, measures, exclusions, and criteria for determining sample size.

Study 1

The present study compared English speakers' endorsement of binary gender pronouns and identity-neutral pronouns to a range of other plausible identity-based pronouns. Methodologically, this study allowed us to develop and validate a paradigm for assessing participants' reasoning about pronouns. We expected to find a basic familiarity preference, with participants showing the strongest endorsement of binary gender pronouns (the status quo in English).

Method

Participants

The sample consisted of 220 U.S. English-speaking participants. Prior to exclusions, 281 respondents completed the online survey using Amazon's Mechanical Turk (MTurk), and 61 were excluded as described in detail below. Due to a technical error, complete demographic information was not collected, but partial aggregated information was recovered from the platform administrators, indicating that participants were all U.S. adults, $M_{age} = 41.11$, 41%

women and 49% men, 10% unspecified. A sensitivity power analysis indicated that this sample size was large enough to detect small differences between endorsement of any two pronoun types (d = 0.19 on a two-tailed paired-samples t test, power = 80%, α = .05; Cohen, 1992; Faul et al., 2007).

Procedure and Materials

Rather than assessing beliefs about specific examples of pronouns (e.g., *they*), we instead assessed participants' beliefs about *types* of pronouns in the abstract (e.g., "neutral pronouns"). This approach enabled us to keep our methodology consistent across the English- and Turkish-speaking samples; given the substantial phonological and morphosyntactic differences between English and Turkish, eliciting participants' attitudes toward specific lexical items would have posed challenges for consistency. In addition, this methodology allowed us to measure participants' attitudes toward pronouns that do not exist in either language but could in principle exist (e.g., race pronouns), broadening the scope of our investigation.

All participants first completed a familiarization phase with a comprehension check. Participants were reminded that third-person singular personal pronouns often denote binary gender in English (e.g., *she* denotes a woman). They were told that it would be possible, in theory, for these types of pronouns to denote other information about individuals, such as their race. This familiarization phase served the critical function of making it clear to participants that our study focused on definite use of third-person singular personal pronouns, without using this jargon. To check comprehension, we showed participants an image of a person with an apparent binary gender and race and asked them to select the pronouns that could apply to the person in the image based on what they had just read. For this study, this check served as an initial screening, and all participants needed to answer this question correctly to proceed. Any

respondents who failed this initial check were immediately redirected to the end of the survey without completing any measures.

After this initial check, participants then responded to six statements about their endorsement of pronouns that encode each of the following identities: binary gender ($\alpha = .83$), race ($\alpha = .87$), age ($\alpha = .87$), wealth ($\alpha = .86$), and weight ($\alpha = .83$). Participants were also asked about their endorsement of identity-neutral pronouns—that is, pronouns that do not encode any identity ($\alpha = 0.84$). For instance, participants were asked, "Should each of the pronouns below be included in all languages?" where one of the "pronouns below" was written as "gender-based pronouns (e.g., males vs. females)." The full list of items is provided in Table 1, as is information about the order in which the items were administered. In total, participants answered 36 questions, six questions about six different pronoun types.

Table 1

Pronoun Endorsement Items in Studies 1–3

Item Label	Item	Scale
Useful	On a scale from 1 to 7, how *useful* are each of the pronouns below?	*1=not useful at all, 7=extremely useful*
Natural	On a scale from 1 to 7, how *natural* are each of the pronouns below?	*1=not natural at all, 7=extremely natural*
Offensive[a]	On a scale from 1 to 7, how *offensive* are each of the pronouns below?	*1=not offensive at all, 7=extremely offensive*
Difficult[a]	On a scale from 1 to 7, how *difficult to use* are each of the pronouns below?	*1=not difficult at all, 7=extremely difficult*
Build a Language	If you were building a language from scratch, how likely would you be to add the pronouns below?	*1=extremely unlikely, 7=extremely likely*
Should	Should each of the pronouns below be included in all languages?	*1=definitely should not, 7=definitely should*

[a]These items were reverse-scored.

Note. The first four items were presented in a random order, followed by the last two items in a random order. One of the "pronouns below," for example, was written as "gender-based pronouns (e.g., males vs. females)."

Next, participants completed two established measures of cognitive heuristics—the inherence bias (Salomon & Cimpian, 2014) and need for cognition (Cacioppo et al., 1984). Because this topic is addressed more fully in Study 2, these analyses are not discussed further in the main text but are reported in the supplementary materials for transparency (see section SI.1.1); the conclusions dovetail with those from Study 2. Relevant to participant exclusions, the inherence bias scale includes built-in exclusion criteria. That is, following the instructions for the scale administration (as in Salomon & Cimpian, 2014), respondents (n = 61) who gave incorrect responses to two or more of four catch items embedded in the scale (e.g., agreeing with "It seems right to kill other people for fun") were excluded.

At the end, participants were thanked for their participation and compensated $0.70. Median completion time was 6 minutes.

Results and Discussion

We conducted a linear mixed-effects model with pronoun type (a six-level categorical variable) predicting pronoun endorsement and a random intercept for participant. Because we had six types of pronouns, the main effect of pronoun type was represented by five different terms in the output of the model. (For any categorical predictor with n levels, its main effect in a regression is represented by $n - 1$ terms, each of which reflects the comparison between a particular level and the omitted level.) Thus, to test the overall statistical significance of the main effect, we used a likelihood-ratio test to compare this model to an intercept-only model. The full model fit significantly better, indicating a main effect of pronoun type, $\chi^2(5) = 678.01$, $p < .001$ (Figure 1, Table 2). In simple slopes analysis, we found that binary gender pronouns were endorsed the most and more than identity-neutral pronouns, $B = -0.77$, $SE = 0.12$, $p < .001$, $\beta =$

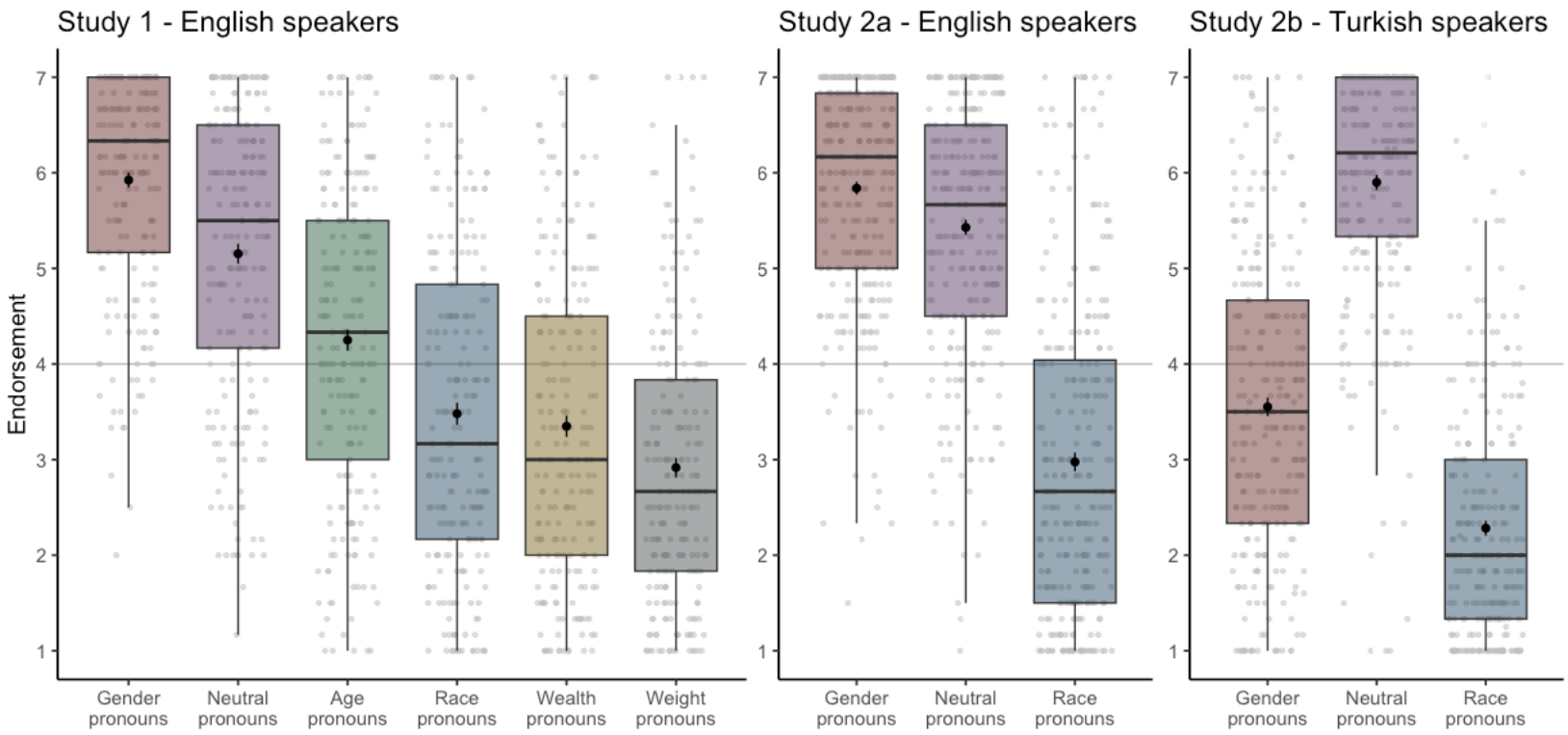
−0.41.[3] In turn, identity-neutral pronouns were endorsed more than age pronouns, $B = -0.90$, $SE = 0.12$, $p < .001$, $\beta = -0.49$, which were endorsed more than race pronouns, $B = -0.77$, $SE = 0.12$, $p < .001$, $\beta = -0.41$, which were endorsed non-significantly more than wealth pronouns, $B = -0.13$, $SE = 0.12$, $p = .276$, $\beta = -0.07$, which were endorsed more than weight pronouns, $B = -0.43$, $SE = 0.12$, $p < .001$, $\beta = -0.23$.

In summary, the results of Study 1 provide evidence for the reliability and validity of our method for assessing participants' reasoning about pronouns. We found that, as hypothesized, English-speaking participants endorsed binary gender pronouns (the familiar status quo) more than any other type of pronouns. Identity-neutral pronouns were also endorsed at relatively high levels, perhaps reflecting growing acceptance of gender-inclusive pronouns or participants' awareness that English has third-person *indefinite* pronouns that are identity-neutral (namely, singular *they*). Age pronouns were endorsed just above the scale midpoint; this level of endorsement could be due to the uncontroversial nature of age distinctions and/or to participants' familiarity with languages in which respect for older individuals is marked via pronoun use (e.g., Spanish, French). Endorsement of pronouns for racial groups, wealth groups, and weight groups were all below the scale midpoint. Notably, as Figure 1 makes clear, our approach also captures a substantial amount of variability among participants in their pronoun endorsement. This variability suggests that participants' reasoning about pronouns is multifaceted rather than being only a byproduct of their familiarity with English pronouns.

[3] The beta coefficient (β) is a measure of effect size derived from a model with standardized pronoun endorsement. This coefficient can be interpreted as the difference between endorsement of gender pronouns and identity-neutral pronouns in standard deviation units.

Figure 1

Pronoun Endorsement in Studies 1, 2a, and 2b



Note. Boxplots with means (black dots), medians (horizontal lines in each of the boxplots), and standard error bars. Raw data are underlaid (gray dots).

Table 2

Mean Pronoun Endorsement and Zero-order Correlations in Studies 1, 2a, and 2b

Pronoun Type	Study 1 – English speakers						Study 2a – English speakers			Study 2b – Turkish speakers		
	M (*SD*)	2.	3.	4.	5.	6.	*M* (*SD*)	2.	4.	*M* (*SD*)	2.	4.
1. Gender	5.92 (1.14)	.09	.17*	.20**	.01	.03	5.84 (1.15)	−.05	.11	3.55 (1.56)	−.27***	.57***
2. Neutral	5.15 (1.52)	—	.03	−.09	.01	−.07	5.43 (1.31)		−.09	5.90 (1.26)		−.33***
3. Age	4.25 (1.63)		—	.72***	.75***	.70***	—			—		
4. Race	3.48 (1.70)			—	.71***	.77***	2.98 (1.66)			2.28 (1.27)		
5. Wealth	3.35 (1.64)				—	.80***	—			—		
6. Weight	2.92 (1.49)					—	—			—		

*$p < .05$. **$p < .01$. ***$p < .001$

Study 2a

In Study 2a, our main goal was to test whether some of the variability in pronoun endorsement among English speakers would be explained by individual differences in the three hypothesized social-cognitive processes: cognitive heuristics (as measured by, e.g., the Cognitive Reflection Test), social ideologies (e.g., essentialism), and meta-linguistic beliefs. Because the additional measures included in this study lengthened it considerably, we narrowed our focus to a key subset of pronouns from Study 1: namely, binary gender, race, and identity-neutral pronouns. Study 2a also set up Study 2b, which was similar but with Turkish-speaking participants.

Method

Participants

As preregistered, the sample was 300 English-speaking participants: all U.S. adults; M_{age} = 39.02; 47% men, 43% women, 5% unspecified, and 2% gender non-binary; and 69% White, 10% Black, 6% East Asian, 5% Hispanic, 5% unspecified, 2% South Asian, 2% multi-racial, and ≤ 1% Middle Eastern or Native American. Prior to exclusions, we collected data from 550 participants from a combination of MTurk and Prolific platforms. Exclusion rates based on the preregistered criteria were unusually high for the MTurk sample (exclusion rate = 69%), which is why Prolific was also used (exclusion rate = 15%). Participants were excluded according to two preregistered criteria: if they failed the inherence bias attention check items described in Study 1 (n = 245; as in Salomon & Cimpian, 2014) or if they reported that they did not take the survey seriously when directly asked at the end of the survey (n = 5).

A simulation-based sensitivity power analysis indicated that our sample size (N = 300) was large enough to detect medium-sized simple slopes (minimum detectable β = 0.31) between

any given social-cognitive process and endorsement of any given pronoun type in a multilevel model structured as described below (number of simulations = 1000; power = 80%; α = .05; Cohen, 1992; Green & MacLeod, 2016).

Procedure and Materials

Using the same procedures as in Study 1, all participants first completed the familiarization and initial comprehension check portion of the survey. As in Study 1, participants needed to answer the comprehension check question correctly to proceed to the rest of the study. Next, participants reported their endorsement of binary gender (α = .83), race (α = .89), and identity-neutral (α = .83) pronouns using the same measures as in Study 1. Participants then completed the seven preregistered measures of social-cognitive processes, described in detail next, which we hypothesized might relate to pronoun endorsement (for details and a high-level summary of hypotheses, see Table 3). The seven measures were completed in a counterbalanced order across participants, except that the two measures of essentialism were always completed back-to-back. Finally, participants were thanked for their participation and paid $1.00-1.41. Median completion time was 13 minutes.

Measures of Cognitive Heuristics. We included three measures of cognitive heuristics. We administered the Cognitive Reflection Test (CRT), which is a well-established behavioral measure of susceptibility to domain-general heuristic thinking. Even though this measure is used widely on participant recruitment websites, including the ones used here (Prolific and MTurk), prior exposure to it does not substantially improve participants' performance or otherwise undermine its validity (Meyer et al., 2018), all the more so in the present case because we used a newer (and thus less familiar) version of the items (Baron et al., 2015). The CRT was scored as the proportion of correct responses (Frederick, 2005). Higher scores indicate more careful,

analytic thinking.

To measure heuristics about the status quo specifically, we also administered the longevity bias scale (Blanchar & Eidelman, 2019) and the inherence bias scale (also called the *inherence heuristic scale*; Salomon & Cimpian, 2014). The longevity bias scale had high face validity for our purposes—it was designed to capture heuristics about the status quo (i.e., what is and what has been that way for a long time) being good and the way things ought to be (Blanchar & Eidelman, 2019). The inherence bias scale captures the tendency to generate inherent explanations for various aspects of the status quo, explanations that in turn rationalize the status quo (Hussak & Cimpian, 2015; Tworek & Cimpian, 2016). Particularly relevant here, this explanation-driven rationalization process has been documented with respect to language (Sutherland & Cimpian, 2015).

Measures of Social Ideologies. We included three measures of social ideologies. Two of them assessed essentialism about gender and, separately, race (Rhodes & Gelman, 2009). These scales capture the belief that gender and race are discrete (and in the case of gender, binary), biologically deep, and fixed social distinctions. We also administered the updated social dominance orientation scale (SDO; Ho et al., 2015). This ideology captures a preference for hierarchies among social groups.

Measures of Meta-linguistic Beliefs. We created 10 items to measure the belief that language does and should reflect social reality (*language–metaphysics link*). These items formed a single factor and thus are treated as a single scale ($\alpha = .88$; see SI.2a.1 for additional details).

Table 3

Social-Cognitive Processes Measured in Studies 2a and 2b with Example Items, Hypotheses, Internal Consistencies, and Zero-order Correlations

Measure	Example Item and Hypothesis	Study 2a – English speakers							Study 2b – Turkish speakers			
		α	2.	3.	4.	5.	6.	7.	α	4.	5.	6.
1. CRT	"Soup and salad cost $5.50 in total. The soup costs a dollar more than the salad. How much does the salad cost?"	—	−.22 ***	−.27 ***	−.14 *	−.25 ***	−.09	−.30 ***	—	.04	−.02	.08
	We hypothesized that CRT score (higher = less heuristic thinking) would relate to less endorsement of status quo pronouns (e.g., binary gender pronouns in English).											
2. LB	"Timeworn ideas generally turn out to be right about things."	.93		.71 ***	.71 ***	.63 ***	.39 ***	.63 ***				
	We hypothesized that this heuristic would relate to more endorsement of status quo pronouns (e.g., binary gender pronouns in English).											
3. IB	"It seems ideal that there are 7 days in a week."	.90			.62 ***	.58 ***	.35 ***	.57 ***				
	We hypothesized that this heuristic would relate to more endorsement of status quo pronouns (e.g., binary gender pronouns in English).											
4. GE	"People that are the same gender have many things in common."	.91				.71 ***	.45 ***	.52 ***	.87		.59 ***	.32 ***
	We hypothesized that this ideology about gender distinctions would relate to higher endorsement of binary gender pronouns.											
5. RE	"People that are the same race have many things in common."	.87					.36 ***	.58 ***	.86			.24 ***
	We hypothesized that this ideology about race distinctions would relate to higher endorsement of race pronouns.											
6. SDO	"An ideal society requires some groups to be on top and others to be on the bottom."	.89						.62 ***	.80			
	We hypothesized that this ideology about social hierarchies would relate to higher endorsement of binary gender and race pronouns.											
7. LML	"If a language uses different words for groups that seem similar, there MUST be some deeper difference between those groups."	.88										
	We hypothesized that this belief would relate to more endorsement of status quo pronouns (e.g., binary gender pronouns in English).											

Note. CRT = Cognitive Reflection Test. LB = Longevity bias. IB = Inherence bias. GE = Gender essentialism. RE = Race essentialism. SDO = Social dominance orientation. LML = Language–metaphysics link. *$p < .05$. **$p < .01$. ***$p < .001$.

Results and Discussion

First, we tested which types of pronouns participants endorsed the most, using the same preregistered analytic strategy described in Study 1. We computed a linear mixed-effects model with pronoun type (binary gender, race, identity-neutral) predicting pronoun endorsement, including a random intercept for participant. We compared this model to an intercept-only model and found a significant difference in fit, indicating a main effect of pronoun type, $\chi^2(2) = 15.71$, $p < .001$ (Figure 1, Table 2). Binary gender pronouns were endorsed the most and more than identity-neutral pronouns, $B = -0.41$, $SE = 0.12$, $p < .001$, $\beta = -0.22$.[4] In turn, identity-neutral pronouns were endorsed more than race pronouns, $B = -2.45$, $SE = 0.12$, $p < .001$, $\beta = -1.31$. These findings replicate Study 1.

Our primary goal was to determine how the measures of social-cognitive processes relate to individual participants' pronoun endorsement. To do so, we conducted a series of preregistered linear mixed-effects models with each social-cognitive process (e.g., gender essentialism; a continuous variable), pronoun type (a three-level categorical variable), and their interaction terms predicting pronoun endorsement. Each model included a random intercept for participant.[5] Because we had three types of pronouns, the interaction between each social-cognitive process and pronoun type was represented by two terms in each model output.

To test the statistical significance of the overall process × pronoun type (gender, identity-

[4] This model exhibited singular fit, likely because the random intercept for participants did not explain any variance. Omitting the random intercept and conducting an ordinary least-squares regression (which is the recommended strategy in this situation; Muradoglu et al., 2023) gave us identical fixed effects estimates. Conclusions were also the same from paired-samples *t* tests comparing endorsement of binary gender pronouns to identity-neutral pronouns, $t(283) = 3.79$, $p < .001$, *Hedges' g* = 0.33, 95% CI [0.16, 0.50], and comparing identity-neutral pronouns to race pronouns, $t(281) = 18.82$, $p < .001$, *Hedges' g* = 1.63, 95% CI [1.42, 1.86].

[5] The model involving inherence bias also exhibited singular fit. Again, omitting the random intercept (Muradoglu et al., 2023) resulted in identical estimates.

neutral, race) interaction, we used a likelihood-ratio test to compare this model to an identical model without the two interaction terms. Using this approach, we found evidence for a significant interaction for all seven social-cognitive processes we investigated (Table 4).

Six of the seven social-cognitive measures (all except the CRT) showed the same pattern of simple slopes: Higher scores on the relevant social-cognitive process corresponded to higher relative endorsement of binary gender and race pronouns and lower endorsement of identity-neutral pronouns. Because higher scores on the CRT indicate *less* heuristic thinking, its relations with pronoun endorsement were almost the mirror image of the others': Higher scores on the CRT corresponded to lower endorsement of race pronouns and higher endorsement of identity-neutral pronouns. Scores on the CRT did not relate significantly with endorsement of binary gender pronouns. While some of these relations were expected (e.g., status quo heuristics predicting higher endorsement of binary gender pronouns), others were not (e.g., gender essentialism predicting higher endorsement of race pronouns). However, it is noteworthy that the seven measures of social-cognitive processes shared substantial variance (see Table 3). For instance, gender essentialism might predict higher endorsement of race pronouns simply because it is correlated at $r = .71$ with race essentialism. We also preregistered a "simultaneous" model where all predictors were entered in parallel to examine the *unique* variance in pronoun endorsement explained by each social-cognitive process. This model, described next, is likely the most informative given the high degree of intercorrelations we observed.

The preregistered "simultaneous" model included the seven measures of social-cognitive processes, the three-level pronoun-type variable, and their 14 interaction terms. Despite the large number of intercorrelated predictors, variance inflation factors were acceptable (range = 1.14 to 3.97), indicating multicollinearity was not a concern. In interpreting our results, we focus on the

simple slopes between each social-cognitive process and each pronoun type.

Social ideologies stood out as the key predictors of pronoun endorsement in this model. Stronger endorsement of gender and race essentialism predicted stronger endorsement of binary gender pronouns and race pronouns, respectively. In terms of magnitude, the positive relations of gender essentialism with endorsement of binary gender pronouns ($\beta = 0.37$) and race essentialism with endorsement of race pronouns ($\beta = 0.30$) were approximately 1.5 to 2 times as strong as the next strongest relation in the simultaneous model (see Table 4, rightmost column). Social dominance orientation also predicted stronger endorsement of race pronouns. Aside from social ideologies, CRT scores related to stronger endorsement of identity-neutral pronouns, and LML scores related to stronger endorsement of race pronouns.

Finally, we conducted a few robustness checks. First, we considered the possibility that including multiple measures of the same general construct (e.g., three measures of cognitive heuristics) in the simultaneous model might have artificially depressed the relations of these measures with pronoun endorsement—they may have "cancelled each other out." To explore this idea, we instead created a cognitive heuristic composite (averaging reverse-scored CRT score, longevity bias, and inherence bias) and a social ideology composite (averaging gender and race essentialism and social dominance orientation). The simultaneous model using these composite measures confirmed the key role of social ideologies (see SI.2a.2 in the supplementary materials). Second, because the exclusion rate was unusually high on the MTurk recruitment platform, we conducted exploratory analyses with the complete sample (see SI.2a.3). Conclusions do not change. Third, we explored whether participants' gender moderated the results; it did not (see SI.2a.4). Fourth, two sets of preregistered but secondary analyses are reported in supplementary materials for brevity, alongside additional details about the gender

essentialism items (see SI.2a.5 and SI.2a.6). None of these analyses alter the conclusions reported here in the main text.

In summary, as in Study 1 we again found that English-speaking participants showed a familiarity preference, endorsing binary gender pronouns the most when asked to consider different types of singular third-person personal pronouns. Crucially, we also found that participants' social ideologies explained some of the individual variability in pronoun endorsement. For instance, participants higher in (binary) gender essentialism and race essentialism endorsed binary gender pronouns and race pronouns relatively more, respectively. Similarly, participants higher in SDO were more likely to endorse race pronouns.

Table 4

Relationships Between Seven Social-Cognitive Processes and Endorsement of Three Pronoun Types in Study 2a, in Separate and Simultaneous Models

Measure	Term	Separate Models for Each Process			Simultaneous Model	
		χ^2 (*df*)	*B* (*SE*)	β	*B* (*SE*)	β
CRT	CRT × Pronoun Type	23.40 (2) ***				
	CRT × Neutral (= 1) vs. Gender (= 0) Pronouns		0.71 (0.31) *	0.14	0.42 (0.31)	0.08
	CRT × Race (= 1) vs. Gender (= 0) Pronouns		−0.81 (0.31) *	−0.16	−0.49 (0.31)	−0.10
	CRT × Neutral (= 1) vs. Race (= 0) Pronouns		1.52 (0.31) ***	0.30	0.91 (0.31) **	0.18
	Simple slope of CRT for Gender Pronouns		0.12 (0.22)	0.02	0.29 (0.23)	0.06
	Simple slope of CRT for Race Pronouns		−0.69 (0.22) **	−0.14	−0.20 (0.23)	−0.04
	Simple slope of CRT for Neutral Pronouns		0.83 (0.22) ***	0.17	0.71 (0.23) **	0.14
LB	LB × Pronoun Type	54.74 (2) ***				
	LB × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.61 (0.09) ***	−0.40	−0.18 (0.15)	−0.12
	LB × Race (= 1) vs. Gender (= 0) Pronouns		−0.02 (0.09)	−0.01	−0.06 (0.15)	−0.04
	LB × Neutral (= 1) vs. Race (= 0) Pronouns		−0.59 (0.09) ***	−0.38	−0.12 (0.15)	−0.08
	Simple slope of LB for Gender Pronouns		0.35 (0.07) ***	0.23	0.05 (0.11)	0.04
	Simple slope of LB for Race Pronouns		0.33 (0.07) ***	0.21	0.00 (0.11)	0.00
	Simple slope of LB for Neutral Pronouns		−0.26 (0.07) ***	−0.17	−0.12 (0.11)	−0.08
IB	IB × Pronoun Type	42.13 (2) ***				
	IB × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.43 (0.07) ***	−0.35	−0.02 (0.11)	−0.02
	IB × Race (= 1) vs. Gender (= 0) Pronouns		−0.03 (0.07)	−0.02	−0.02 (0.11)	−0.02
	IB × Neutral (= 1) vs. Race (= 0) Pronouns		−0.41 (0.07) ***	−0.33	0.00 (0.11)	0.00
	Simple slope of IB for Gender Pronouns		0.27 (0.05) ***	0.22	0.09 (0.08)	0.07
	Simple slope of IB for Race Pronouns		0.24 (0.05) ***	0.20	0.06 (0.08)	0.05
	Simple slope of IB for Neutral Pronouns		−0.16 (0.05) **	−0.13	0.07 (0.08)	0.05
GE	GE × Pronoun Type	69.49 (2) ***				
	GE × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.47 (0.06) ***	−0.49	−0.44 (0.09) ***	−0.45
	GE × Race (= 1) vs. Gender (= 0) Pronouns		−0.13 (0.06) *	−0.14	−0.55 (0.09) ***	−0.56
	GE × Neutral (= 1) vs. Race (= 0) Pronouns		−0.34 (0.06) ***	−0.35	0.11 (0.09)	0.11
	Simple slope of GE for Gender Pronouns		0.30 (0.04) ***	0.31	0.36 (0.07) ***	0.37
	Simple slope of GE for Race Pronouns		0.16 (0.04) ***	0.17	−0.19 (0.07) **	−0.19
	Simple slope of GE for Neutral Pronouns		−0.18 (0.04) ***	−0.18	−0.08 (0.07)	−0.08
RE	RE × Pronoun Type	59.22 (2) ***				
	RE × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.30 (0.06) ***	−0.30	0.17 (0.09)	0.17

	RE × Race (= 1) vs. Gender (= 0) Pronouns		0.16 (0.06) **	0.16	0.43 (0.09) ***	0.43
	RE × Neutral (= 1) vs. Race (= 0) Pronouns		−0.46 (0.06) ***	−0.46	−0.26 (0.09) **	−0.26
	Simple slope of RE for Gender Pronouns		0.16 (0.04) ***	0.16	−0.13 (0.07) *	−0.14
	Simple slope of RE for Race Pronouns		0.32 (0.04) ***	0.32	0.30 (0.07) ***	0.30
	Simple slope of RE for Neutral Pronouns		−0.14 (0.04) ***	−0.14	0.03 (0.07)	0.03
SDO	SDO × Pronoun Type	42.56 (2) ***				
	SDO × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.43 (0.09) ***	−0.29	−0.14 (0.10)	−0.10
	SDO × Race (= 1) vs. Gender (= 0) Pronouns		0.13 (0.09)	0.09	0.22 (0.10) *	0.15
	SDO × Neutral (= 1) vs. Race (= 0) Pronouns		−0.55 (0.09) ***	−0.38	−0.36 (0.10) ***	−0.25
	Simple slope of SDO for Gender Pronouns		0.14 (0.06) *	0.10	−0.06 (0.07)	−0.04
	Simple slope of SDO for Race Pronouns		0.27 (0.06) ***	0.18	0.16 (0.07) *	0.11
	Simple slope of SDO for Neutral Pronouns		−0.28 (0.06) ***	−0.20	−0.20 (0.07) **	−0.14
LML	LML × Pronoun Type	46.32 (2) ***				
	LML × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.36 (0.07) ***	−0.29	−0.02 (0.10)	−0.02
	LML × Race (= 1) vs. Gender (= 0) Pronouns		0.15 (0.07)	0.12	0.13 (0.10)	0.11
	LML × Neutral (= 1) vs. Race (= 0) Pronouns		−0.50 (0.08) ***	−0.41	−0.16 (0.10)	−0.13
	Simple slope of LML for Gender Pronouns		0.18 (0.05) ***	0.15	0.01 (0.07)	0.01
	Simple slope of LML for Race Pronouns		0.33 (0.05) ***	0.26	0.15 (0.07) *	0.12
	Simple slope of LML for Neutral Pronouns		−0.17 (0.05) ***	−0.14	−0.01 (0.07)	−0.01

Note. CRT = Cognitive Reflection Test. LB = Longevity bias. IB = Inherence bias. GE = Gender essentialism. RE = Race essentialism. SDO = Social dominance orientation. LML = Language–metaphysics link. The "Separate Models" column shows seven different models, each with a single process and its interaction with pronoun type predicting pronoun endorsement, with a random intercept for participant. The rightmost column displays the results of an eighth model, in which each process and its interactions with pronoun type were entered simultaneously. The beta coefficient (β) is a measure of effect size derived from a model with standardized process(es) and standardized pronoun endorsement. Note that for RE, the unstandardized and standardized coefficients are identical because RE and pronoun endorsement have similar standard deviations. *$p < .05$. **$p < .01$. ***$p < .001$.

Study 2b

In Study 2b, we investigated beliefs about pronouns among Turkish-speaking participants and with a subset of the social-cognitive process measures from Study 2a. We chose to focus on Turkish speakers because in Turkish, singular third-person personal pronouns are identity-neutral and can refer to people of any and all identities (e.g., *o*). Whereas in Studies 1 and 2a we hypothesized that English-speaking participants would favor binary gender pronouns because of the language-specific status quo in English, here we instead hypothesized that Turkish-speaking participants would favor identity-neutral pronouns because of the language-specific status quo in Turkish. Despite the substantial differences between the English and Turkish languages (and the differences between our two samples, e.g., in recruitment procedures), we also hypothesized that the same social ideologies would relate to participants' pronoun endorsement. That is, we expected that Turkish-speaking participants who endorsed gender essentialism, race essentialism, and social dominance orientation more would also show relatively stronger endorsement of binary gender pronouns and race pronouns. This hypothesized result would suggest that ideological processes may be a language-general factor underlying people's reasoning about pronouns.

Method

Participants

Following the preregistered recruitment plan, the sample consisted of 260 Turkish-speaking participants with the following characteristics: all adults; M_{age} = 30.69; 25% men, 57% women, 17% unspecified, and 1% gender non-binary. In terms of race/ethnicity, 47% self-identified as ethnically "Turkish" with the next largest groups being unspecified (29%), and White (8%) with a wide range of other identities listed (e.g., "Kurdish," "Armenian," "Jewish,"

"Black," and "mixed"). All participants were fluent in Turkish and most listed Turkish as their first language—only 3% listed another language other than Turkish as their first language, and only 5% listed that they were raised in another country besides Turkey.

To recruit this sample, we used a mixture of platforms including Prolific, social media outlets, and crowd-sourcing. Only one participant was excluded according to pre-registered exclusion criterion of indicating they did not take the survey seriously. A simulation-based sensitivity power analysis indicated that our sample size ($N = 260$) was large enough to detect small to medium sized simple slopes (minimum detectable $\beta = 0.18$) between any given social-cognitive process and endorsement of any given pronoun type in a multilevel model structured the same as in Study 2a (number of simulations = 1000; power = 77%; $\alpha = .05$; Cohen, 1992; Green & MacLeod, 2016).

Procedure and Materials

The procedure and materials were similar to Study 2a, except the study was conducted entirely in Turkish. Materials were forward-translated from English into Turkish by the fourth author; then, another fluent Turkish speaker back-translated them into English. English back-translations were verified by the first and fourth author and minor discrepancies between the English originals and back-translations were resolved by consulting additional Turkish speakers. To adapt to a Turkish context, we used the term "race/ethnicity" throughout instead of "race" and changed a CRT item to be about yogurt instead of sun tea, an unfamiliar food item in Turkey. Critically, we also changed the introductory information about pronouns to match the new linguistic context. Recall that in the studies with English speakers, participants were first remined that English has binary gender pronouns and were then introduced to the possibility that languages could have identity-neutral pronouns or pronouns distinguishing other identities,

including race. Here, Turkish-speaking participants were reminded that Turkish has identity-neutral pronouns and introduced to the possibility that languages could have pronouns distinguishing various social identities, including binary gender and race/ethnicity. As in Study 2a, participants needed to answer a comprehension check question correctly to proceed. (Any respondents who failed this initial screening were redirected to the end of the survey.)

Next, participants reported their endorsement of binary gender (α = .86), race/ethnicity (α = .81), and identity-neutral (α = .82) pronouns. Participants then completed the preregistered measures of social ideologies—specifically, gender essentialism, race/ethnicity essentialism, and social dominance orientation. We included the CRT as our only measure of cognitive heuristics because it was the only measure from Study 2a that was predictive in the simultaneous model. Median completion time was 12 minutes, and participants were compensated between $0-2.25 depending on recruitment method.

Results and Discussion

Using the same preregistered analytic approach described in Study 2a, we first found an overall effect of pronoun type, $\chi^2(2) = 613.56$, $p < .001$ (see Figure 1 and Table 2). Specifically, identity-neutral pronouns were endorsed the most and more than gender pronouns, $B = -2.35$, $SE = 0.12$, $p < .001$, $\beta = -1.16$. In turn, gender pronouns were endorsed more than race/ethnicity pronouns, $B = -1.27$, $SE = 0.12$, $p < .001$, $\beta = -0.62$. This pattern is strikingly different than that found for English speakers but consistent with the pronoun status quo in Turkish.

Next, we used the same preregistered analytic approach described in Study 2a to examine whether social ideologies might predict pronoun endorsement. The results were almost identical to the results for English-speaking participants: The more Turkish-speaking participants endorsed ideological beliefs portraying gender groups and racial/ethnic groups as inherently

discrete (essentialism) and hierarchal (social dominance orientation), the more they endorsed binary gender pronouns and race/ethnicity pronouns. Also similar to Study 2a, participants with higher CRT scores were more likely to endorse identity-neutral pronouns. These results emerged both in the models in which each social-cognitive process was entered by itself and in the simultaneous model that adjusted for the other processes. Note that, again, despite the large number of intercorrelated predictors, variance inflation factors were acceptable (range = 1.01 to 2.20), indicating multicollinearity was not a concern in the simultaneous model. Finally, in exploratory analyses reported in supplementary materials, we did not find any evidence that participant gender moderated these results (see SI.2b.1).

In summary, in contrast to English-speaking participants in Studies 1 and 2a, Turkish-speaking participants showed a different pattern of pronoun preferences: Turkish speakers endorsed identity-neutral pronouns the most, consistent with the status quo in their language. However, this language-specific preference was accompanied by a language-general process: Social ideologies predicted Turkish-speaking participants' beliefs about pronouns in the same way they did for English-speaking participants (see Table 5). Even though Turkish-speaking participants were below the midpoint in their overall endorsement of binary gender pronouns and race/ethnicity pronouns, the more they thought that gender and racial/ethnic groups are distinct (essentialized) and that social groups should be hierarchal (social dominance orientation), the more they endorsed binary gender pronouns and race/ethnicity pronouns. It is noteworthy that these effects were not trivial in magnitude. For example, in the simultaneous model controlling for other factors (see Table 5, rightmost column), for every 1 *SD* increase in gender essentialism there was a .27 *SD* increase in endorsement of binary gender pronouns, which translates into a .54 *SD* difference between participants who are 1 *SD* above vs. below the mean or approximately

0.84 scale points on the 1–7 scale. The analogous statistic for English-speaking participants in Study 2a is 0.85 scale points.

Table 5

Relationships Between Four Social-Cognitive Processes and Endorsement of Three Pronoun Types in Study 2b, in Separate and Simultaneous Models

Measure	Term	Separate Models for Each Process			Simultaneous Model	
		χ^2 (*df*)	*B* (*SE*)	β	*B* (*SE*)	β
CRT	CRT × Pronoun Type	10.37 (2) **				
	CRT × Neutral (= 1) vs. Gender (= 0) Pronouns		1.22 (0.40) **	0.19	1.25 (0.37) ***	0.19
	CRT × Race (= 1) vs. Gender (= 0) Pronouns		0.21 (0.40)	0.03	0.31 (0.37)	0.04
	CRT × Neutral (= 1) vs. Race (= 0) Pronouns		1.00 (0.40) *	0.16	0.94 (0.37) *	0.15
	Simple slope of CRT for Gender Pronouns		−0.50 (0.29)	−0.08	−0.58 (0.27) *	−0.09
	Simple slope of CRT for Race Pronouns		−0.28 (0.29)	−0.04	−0.27 (0.27)	−0.04
	Simple slope of CRT for Neutral Pronouns		0.72 (0.29) *	0.11	0.67 (0.27) *	0.10
GE	GE × Pronoun Type	45.91 (2) ***				
	GE × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.49 (0.07) ***	−0.40	−0.40 (0.09) ***	−0.33
	GE × Race (= 1) vs. Gender (= 0) Pronouns		−0.14 (0.07)	−0.12	−0.26 (0.09) **	−0.21
	GE × Neutral (= 1) vs. Race (= 0) Pronouns		−0.34 (0.07) ***	−0.29	−0.14 (0.09)	−0.11
	Simple slope of GE for Gender Pronouns		0.39 (0.05) ***	0.33	0.33 (0.06) ***	0.27
	Simple slope of GE for Race Pronouns		0.25 (0.05) ***	0.21	0.07 (0.06)	0.06
	Simple slope of GE for Neutral Pronouns		−0.09 (0.05)	−0.07	−0.06 (0.06)	−0.05
RE	RE × Pronoun Type	31.24 (2) ***				
	RE × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.32 (0.07) ***	−0.28	−0.05 (0.08)	−0.04
	RE × Race (= 1) vs. Gender (= 0) Pronouns		0.03 (0.07)	0.03	0.17 (0.08) *	0.14
	RE × Neutral (= 1) vs. Race (= 0) Pronouns		−0.36 (0.07) ***	−0.31	−0.22 (0.00) **	−0.19
	Simple slope of RE for Gender Pronouns		0.25 (0.05) ***	0.21	0.02 (0.06)	0.02
	Simple slope of RE for Race Pronouns		0.28 (0.05) ***	0.24	0.19 (0.06) **	0.16
	Simple slope of RE for Neutral Pronouns		−0.08 (0.05)	−0.07	−0.03 (0.06)	−0.03
SDO	SDO × Pronoun Type	29.34 (2) ***				
	SDO × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.59 (0.12) ***	−0.31	−0.43 (0.12) ***	−0.22
	SDO × Race (= 1) vs. Gender (= 0) Pronouns		−0.12 (0.12)	−0.06	−0.08 (0.12)	−0.04
	SDO × Neutral (= 1) vs. Race (= 0) Pronouns		−0.47 (0.12) ***	−0.25	−0.35 (0.12) **	−0.18
	Simple slope of SDO for Gender Pronouns		0.43 (0.08) ***	0.22	0.28 (0.08) ***	0.15
	Simple slope of SDO for Race Pronouns		0.31 (0.08) ***	0.16	0.29 (0.08) *	0.10
	Simple slope of SDO for Neutral Pronouns		−0.17 (0.08) *	−0.09	−0.14 (0.08)	−0.07

Note. CRT = Cognitive Reflection Test. GE = Gender essentialism. RE = Race essentialism. SDO = Social dominance orientation. The "Separate Models" column shows four different models, each with a single process and its interaction with pronoun type predicting pronoun endorsement, with a random intercept for participant. The rightmost column displays the results of a fifth model, in which each process and its interactions with pronoun type were entered simultaneously. The beta coefficient (β) is a measure of effect size derived from a model with standardized process(es) and standardized pronoun endorsement. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Study 3

The language-specific status quo evidently shapes pronoun preferences, but social ideologies also appear to matter, regardless of the language participants speak (English vs. Turkish). The present preregistered study built on the correlational findings from Studies 2a and 2b to test whether experimentally manipulating social ideologies would cause corresponding changes in pronoun endorsement.

Specifically, we manipulated gender essentialism in an English-speaking sample. We focused on gender essentialism both because of its conceptual relevance to current debates about the use of gender-inclusive pronouns (and language more generally; Saguy et al., 2021) and because it showed the strongest empirical relationship with pronoun endorsement among all constructs investigated with English-speaking participants (Study 2a, $\beta = 0.37$ with gender pronouns) and Turkish-speaking participants (Study 2b, $\beta = 0.27$ with gender pronouns). On a more practical level, prior research also provides validated means of manipulating gender essentialism (e.g., Christy et al., 2019), making an experimental approach to our questions feasible. We expected that manipulating gender essentialism would lead to corresponding changes in participants' endorsement of binary gender pronouns. We did not have strong predictions about the effect of the manipulation on endorsement of identity-neutral or race pronouns, which were also included for comparison in the present study.

Method

Prior to this study, we conducted an initial pilot study with near-identical methods. The pilot study is reported in supplementary materials (SI.3.3) and meta-analyzed with the present study at the end of this section.

Participants

The sample consisted of 624 English-speaking participants: all U.S. adults; M_{age} = 43.62; 51% women, 48% men, 1% gender non-binary; and 62% self-identified as "White" or "Caucasian" and the remaining 38% self-identified as a range of other race and ethnicities (e.g., "Hispanic," "Indigenous American," and "mixed"). Following the pre-registered recruitment plan, we recruited 775 participants to arrive at a sample of at least 600 participants after exclusions. We only needed to exclude 151 participants according to preregistered criteria (detailed below), leaving a final sample of 624 participants. A sensitivity power analysis indicated that this sample was able to detect a small effect of condition on pronoun endorsement (d = 0.22 on a two-tailed independent-samples t test, power = 80%, α = .05; Cohen, 1992; Faul et al., 2007).

Procedure and Materials

Participants were told that they would be completing two short, unrelated surveys. In reality, the first survey administered the essentialism manipulation and the second survey measured pronoun endorsement. We removed respondents as preregistered who only completed the first survey (n = 16) or who indicated that they were two different ages when asked twice, suggesting low attentiveness (n = 8).

The essentialism manipulation was taken exactly from prior work (Christy et al., 2019). Participants read an article about gender that emphasized gender as either binary, biologically-determined, and fixed in the "high essentialism" condition or as complex, culturally-determined, and malleable in the "low essentialism" condition. This article was presented to participants as part of a brief survey on scientific literacy and included two comprehension check questions. Participants who failed both of these questions were excluded (n = 7). To bolster the cover story, the survey also included an article about birds and questions about reading habits. As stipulated

in our preregistration, these materials were only included as part of the cover story. Finally, participants answered questions about their "personal opinions on some of the topics" that they had read about. This section included the gender essentialism items from Study 2a (α = .92).

Participants were then directed to an ostensibly unrelated survey on language. As in Study 1, participants read information about personal pronouns and completed a comprehension check question. Participants who failed this check were excluded from analyses as preregistered and as in Studies 1, 2a, and 2b (n = 114). Next, using questions that were nearly identical to those from the preceding studies, we elicited participants' endorsement of gender (α = .86), race (α = .87), and identity-neutral (α = 0.88) pronouns. We modified the wording slightly relative to the preceding studies to avoid any possible misreading and to further clarify that our questions about gender pronouns were about *binary* gender pronouns. Whereas in previous studies these pronouns were described as "gender-based pronouns (e.g., males vs. females)," in the present study they were described as "gender-based pronouns (one for men, one for women)." (However, note that in the pilot study we used the exact same wording as in Studies 1, 2a, and 2b with similar results, as described in SI.3.3.)

At the end of each "study," participants were asked what they thought its purpose had been. Unaware of participants' condition, the first author reviewed these open-ended responses and excluded participants who expressed suspicion about the cover story (n = 6). Finally, participants were thanked for their participation, debriefed about the full purpose of the study (including that the two ostensibly unrelated surveys were part of the same study), and compensated $2.00. Median completion time was 16 minutes.

Results and Discussion

We preregistered five sets of analyses. First, we assessed that the manipulation worked as

intended. Second, similar to Studies 2a and 2b, we tested the relationships between gender essentialism and endorsement of different pronoun types. Third, we tested the causal effect of the manipulation on pronoun endorsement. Fourth, we conducted a mediation analysis to test the indirect effect of the manipulation on pronoun endorsement through gender essentialism. Fifth, we meta-analyzed key findings from the present study together with the pilot study.

Manipulation Check

The essentialism manipulation worked as intended. In a preregistered linear regression, gender essentialism was higher in the high-essentialism condition (M = 6.60, SD = 1.57) compared to the low-essentialism condition (M = 5.30, SD = 1.97), B = 1.30, SE = 0.14, $p <$ .001, $\beta = 0.68$.

Relationships Between Gender Essentialism and Pronoun Endorsement

As preregistered, we computed a linear mixed-effects model with gender essentialism, pronoun type, and their two interaction terms predicting pronoun endorsement while controlling for condition. The model included a random intercept for participant. We compared this model to an identical model without the two interaction terms and found a significant difference in fit, indicating an overall gender essentialism × pronoun type interaction (for statistics, see Table 6). Consistent with Studies 2a and 2b (specifically, the single-variable model results in Table 4 and 5), individuals who were higher in gender essentialism showed stronger endorsement of binary gender pronouns and race pronouns and weaker endorsement of identity-neutral pronouns.

Table 6

Relationship Between Essentialism and Endorsement of Three Pronoun Types in Study 3

Term			
	χ^2 (*df*)	*B* (*SE*)	β
GE × Pronoun Type	222.42 (2) ***		
GE × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.63 (0.04) ***	−0.64
GE × Race (= 1) vs. Gender (= 0) Pronouns		−0.18 (0.04) ***	−0.19
GE × Neutral (= 1) vs. Race (= 0) Pronouns		−0.44 (0.04) ***	−0.46
Simple slope of GE for Gender Pronouns		0.33 (0.03) ***	0.33
Simple slope of GE for Race Pronouns		0.14 (0.03) ***	0.15
Simple slope of GE for Neutral Pronouns		−0.30 (0.03) ***	−0.31

Note. GE = Gender essentialism. The beta coefficient (β) is a measure of effect size derived from a model with standardized GE and standardized pronoun endorsement. This model also controlled for experimental condition, although this term is not listed in the table for brevity.

*p < .05. **p < .01. ***p < .001.

Experimental Effect of Condition on Pronoun Endorsement

Next, we tested whether the essentialism manipulation affected participants' endorsement of binary gender pronouns. As preregistered, we computed a linear mixed-effects model with condition, pronoun type, and their interaction terms predicting pronoun endorsement, with a random intercept for participant. We compared this model to an identical model without the two interaction terms to compute the overall condition × pronoun type interaction. The model with the interaction terms showed better fit, χ^2 (2) = 10.16, p = .006, indicating that the condition × pronoun type interaction was significant.[6] Examining the simple slopes, participants in the high-essentialism condition (M = 5.89, SD = 1.30) showed greater endorsement of binary gender pronouns compared to those in the low-essentialism condition (M = 5.43, SD = 1.31), B = 0.46, SE = 0.12, p < .001, β = 0.25.[7] There was no evidence that the gender essentialism manipulation impacted endorsement of race pronouns, B = 0.07, SE = 0.12, p = .559, β = 0.04, or identity-neutral pronouns, B = −0.06, SE = 0.12, p = .607, β = −0.03.

These results suggest that, as hypothesized, manipulating participants' endorsement of gender essentialism produced corresponding changes in their endorsement of binary gender pronouns. Notably, the effect of the manipulation seems to have been driven by exposure to *anti*-essentialist messages: English speakers' endorsement of binary gender pronouns in the high-essentialism condition (M = 5.89) was almost identical to that in the preceding studies, where

---

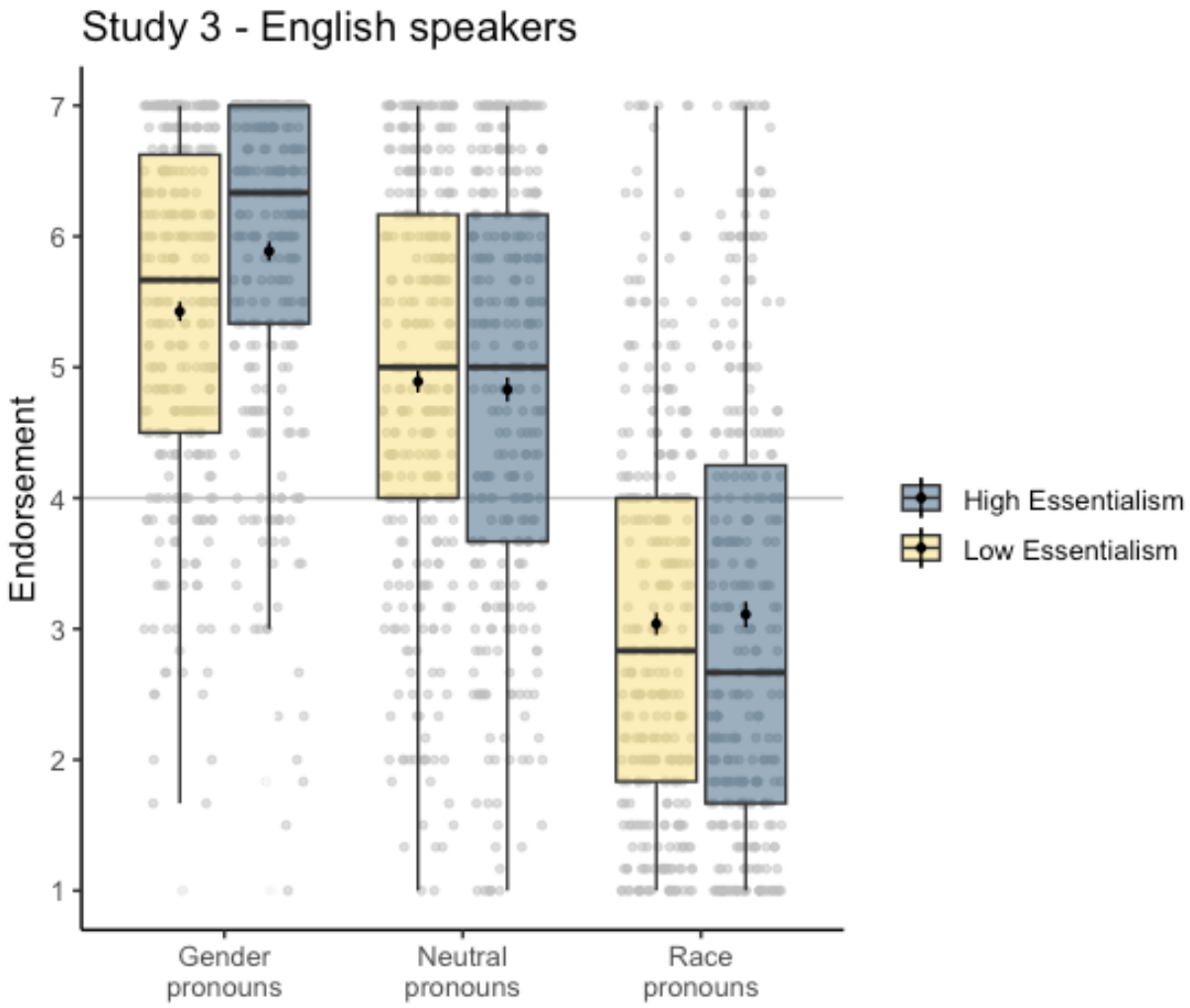
[6] This model exhibited singular fit, likely due to the random intercept explaining little variance. Fixed effect estimates were identical when omitting the random intercept and running an ordinary least-squares (OLS) regression (Muradoglu et al., 2023). The mixed-effects model computed in another software package (Stata 18.0 [StataCorp, 2023]) was able to produce identical estimates without indicating a singularity error. Finally, conclusions were the same from a mixed analysis of variance (ANOVA). Scripts for reproducing these OLS, Stata, and mixed ANOVA analyses are available on the OSF page.

[7] The Cohen's d for this pairwise comparison was 0.35, which is above our minimum detectable effect of d = 0.22. Thus, the study was adequately powered to detect this difference.

there was no experimental manipulation (*M*s = 5.92 and 5.84 in Studies 1 and 2a, respectively), whereas endorsement in the low-essentialism condition was lower (*M* = 5.43). In exploratory analyses, we also again considered differences based on participant gender. There was no evidence that participant gender moderated these results (see SI.3.1).

Figure 2

Experimental Effect of Condition on Pronoun Endorsement



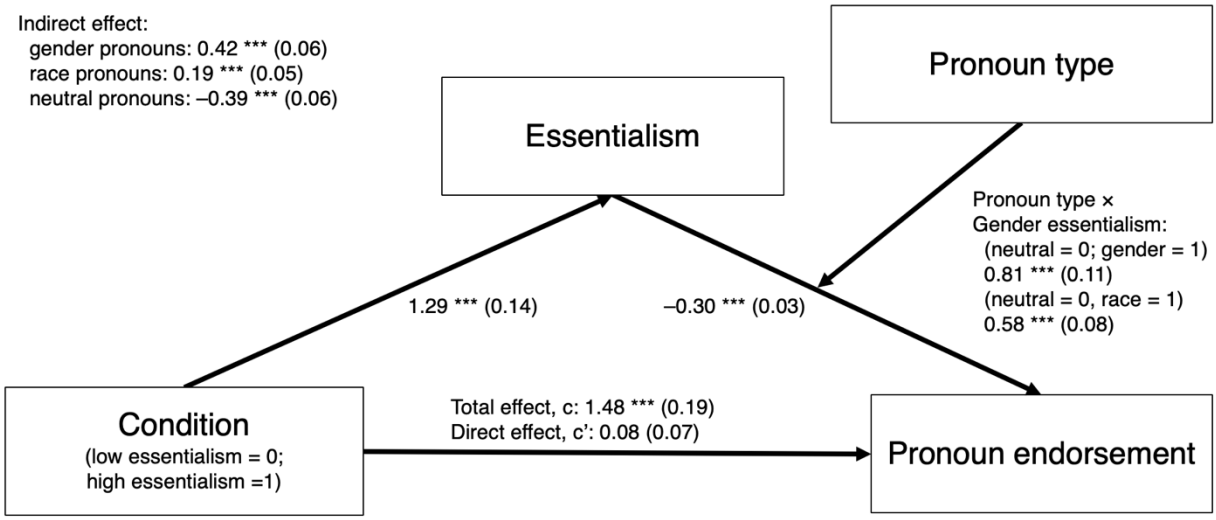
Note. Boxplots with means (black dots), medians (horizontal lines in each of the boxplots), and standard error bars. Raw data are underlaid (gray dots).

Mediation of Condition Effect on Pronoun Endorsement

We expected that the manipulation would impact pronoun endorsement specifically by shifting gender essentialism. We tested this prediction in a preregistered multilevel moderated mediation model (Figure 3). The moderator was pronoun type: We expected that the indirect effect of the experimental manipulation on pronoun endorsement via gender essentialism would differ by pronoun type. Consistent with this prediction, the indirect effect differed for binary gender pronouns compared to identity-neutral pronouns, $B = 0.81$, $SE = 0.11$, $p < .001$, and for race pronouns compared to identity-neutral pronouns, $B = 0.58$, $SE = 0.08$, $p < .001$. For both binary gender and race pronouns, the indirect effects of the manipulation were positive: Participants in the high (vs. low) essentialism condition showed greater gender essentialism, which in turn related to greater endorsement of binary gender pronouns ($ab = 0.42$, $p < .001$) and race pronouns ($ab = 0.19$, $p < .001$). In contrast, the indirect effect of the manipulation was negative for identity-neutral pronouns: Participants in the high (vs. low) essentialism condition showed greater gender essentialism, which in turn related to lower endorsement of identity-neutral pronouns ($ab = -0.39$, $p < .001$).

Figure 3

Multilevel Moderated Mediation of the Effect of Condition on Pronoun Endorsement Through Gender Essentialism



Note. Unstandardized coefficients are reported with *SE*s in parentheses. Gender essentialism is mean centered, and the essentialism condition and pronoun type are dummy coded as specified. *** $p < .001$.

Meta-analysis of the Main Study and the Pilot Study

We previously conducted a pilot study ($N = 600$) similar to the one reported here (see SI.3.3 for details). As preregistered, we meta-analyzed the standardized estimates (βs) from this pilot and the present study. Specifically, we conducted a series of random-effects meta-analyses using the restricted maximum likelihood estimator in JASP 0.18.1 (JASP Team, 2023). Below, we report the relevant meta-analytic estimates, followed by their 95% confidence intervals. The full output (including heterogeneity statistics) can be found on the OSF page.

We first meta-analyzed the correlational relationships between gender essentialism and pronoun endorsement. Gender essentialism was positively related to endorsement of binary gender pronouns, $\beta_+ = 0.30$ [0.24, 0.37], and race pronouns, $\beta_+ = 0.20$ [0.09, 0.31], but negatively related to endorsement of identity-neutral pronouns, $\beta_+ = -0.29$ [−0.34, −0.25]. Next, we meta-analyzed the effects of the manipulation of gender essentialism on pronoun endorsement. The experimental manipulation shifted endorsement of binary gender pronouns, $\beta_+ = 0.18$ [0.03, 0.32], but did not affect endorsement of race pronouns, $\beta_+ = 0.13$ [−0.05, 0.31], or identity-neutral pronouns, $\beta_+ = -0.04$ [−0.13, 0.05].

Summary

Activating anti-essentialist ideas about gender as complex, culturally-determined, and malleable (vs. binary, biological, and fixed) lowered participants' endorsement of binary gender pronouns. Combined with the correlational evidence from English (Study 2a) and Turkish (Study 2b) speakers, the present study reinforces the conclusion that—above and beyond language-specific familiarity effects—social ideologies also shape people's beliefs about pronouns.

General Discussion

The present studies investigated what lay speakers think about pronouns and why. We focused on pronouns both because, as function words, they are extremely common and because they feature prominently in debates about how to encode and convey identity as part of broader controversies about the nature of gender and gender identity (Davis, 2008; Parker et al., 2019). Our first goal was to understand lay people's beliefs about whether—and which—social identities should be encoded in pronouns. We compared endorsement of binary gender pronouns to identity-neutral pronouns as well as to other plausible types of identity-based pronouns (e.g., race pronouns). We found that English-speaking participants endorsed binary gender pronouns the most, consistent with the status quo in English (Studies 1, 2a, and 3). We also found that Turkish-speaking participants endorsed identity-neutral pronouns the most, consistent with status quo pronouns in Turkish (Study 2b). Thus, English and Turkish speakers generally endorsed the specific types of pronouns that were also common in each respective language context. However, we also observed considerable variability among participants, suggesting that beliefs about pronouns were not only a matter of familiarity with the language-specific status quo.

Our second, and primary, goal was to investigate whether some of this variability might be explained by social-cognitive processes that shape individuals' reasoning about pronouns in similar ways across languages. Drawing on theory and previous investigations (e.g., Bradley, 2020; Sarrasin et al., 2012; Sczesny et al., 2015; Vergoossen et al., 2020), we narrowed our focus to three types of processes: cognitive heuristics, social ideologies, and meta-linguistic beliefs. Our findings identified social ideologies that emphasize social distinctions (essentialism) and hierarchies (social dominance orientation) as key aspects of people's reasoning about pronouns, regardless of the language they speak. For instance, the more English- and Turkish-speaking

participants thought of gender as essentialized—reflecting an underlying "true" reality that is binary, biological, and fixed—the more they endorsed binary gender pronouns (Studies 2a and 2b), with small but robust experimental effects in English (Study 3). Taken together, these results indicate that English- and Turkish-speaking participants' reasoning about pronouns is impacted both by the status quo in their language and by individuals' ideological beliefs about the relevant social groups—and in particular by ideologies that reinforce social distinctions (essentialism) and hierarchies (social dominance orientation).

Another noteworthy finding was that reflective, analytic thinking (as measured by the Cognitive Reflection Test) related to greater endorsement of identity-neutral pronouns in both the English- and Turkish-speaking samples (Studies 2a and 2b), even after accounting for participants' social ideologies. This finding is consistent with previous work suggesting that the Cognitive Reflection Test captures analytic thinking independent of ideology. For instance, people who score higher on the Cognitive Reflection Test are more skeptical of inaccurate news headlines ("fake news") regardless of whether the headlines align with or run counter to their own political ideology (Pennycook & Rand, 2019). Thus, individual support for identity-neutral pronouns may be, at least in part, a function of careful reasoning and not only a function of social ideology.

Essentialism About Social Groups and Identities

Essentialism about social groups and identities relates to a range of downstream beliefs, attitudes, and behaviors. For instance, essentialism has been connected to greater stereotyping (Brescoll & LaFrance, 2004), more negative attitudes toward certain groups (e.g., Black people; Chen & Ratliff, 2018), more positive attitudes toward other groups (e.g., gay people; Haslam & Levy, 2006), and less blame for moral wrongdoing (e.g., Bailey et al., 2021). These relationships

can be quite complex, as highlighted in recent work by Bailey and Knobe (2023), Mandalaywala (2020), and Peretz-Lange (2021). But most relevant for our purposes, essentialism has been shown to shape language use and vice versa (e.g., Ritchie, 2021). For example, the more people essentialize a group, the more willing they are to produce generic statements about the group (e.g., "men are good at math") compared to quantified statements (e.g., "some men are good at math"; Rhodes et al., 2012; Wodak et al., 2015). Here, we add a novel finding: the more people essentialize a group, the more they endorse the use of identity-based terms for that group. We found this relationship for both binary gender pronouns and race pronouns and among speakers of two languages with very different pronoun systems, English and Turkish.

Our finding that essentialism about a social identity, whether binary gender or race, predicts endorsement of pronouns that encode that distinction may suggest a broader tendency to endorse obligatory marking in language for essentialized categories. It could be that people generally prefer language to include more specific terms for categories that they essentialize—that they view as "deep" and real. This possibility is consistent with participants' absolute agreement with items on our language–metaphysics link measure (mean agreement was 5.14 in Study 2 on a 1–9 scale), all of which convey the idea that language marks important ontological distinctions. Given that social (and non-social) categories vary in terms of how strongly they are essentialized (Gelman, 2003; e.g., gender tends to be essentialized more than race; Rhodes & Gelman, 2009), in future work it would be useful to systematically explore whether the degree to which a category is essentialized corresponds to the degree to which people believe it should be marked in language.

The link between gender essentialism and endorsement of binary gender pronouns is also reminiscent of a previously documented link between gender essentialism and attitudes toward

non-binary and transgender individuals (Axt et al., 2021; Broussard et al., 2018; Gallagher & Bodenhausen, 2021; Norton & Herek, 2013; Prusaczyk & Hodson, 2020; Rad et al., 2019; Saguy et al., 2021; Schudson & van Anders, 2021; Tee & Hegarty, 2006; Wilton et al., 2019). For instance, Ching and Xu (2018) found that experimentally increasing gender essentialism by exposing participants to gender essentialist arguments (similar to the present Study 3) in turn increased negative attitudes toward transgender people. Thus, support for binary gender pronouns may provide an indirect means for people high in (binary) gender essentialism to express negative attitudes toward transgender and non-binary individuals. This possibility awaits future research.

Limitations and Future Directions

In terms of limitations, it is worth noting that the experimental effect in Study 3 was small in magnitude ($\beta = 0.25$). In part, the modest magnitude may be because English-speaking participants' endorsement of binary gender pronouns was close to ceiling across studies, meaning that the manipulation could only feasibly *decrease* endorsement. A similar point likely applies to participants' gender essentialism which was also high across studies. Furthermore, gender essentialism might be too entrenched to be dramatically altered by brief exposure to a short article. As such, we consider Study 3 to be a "proof of concept" that essentialism has a causal effect on lay beliefs about pronouns. Future research designed to intervene on reasoning about pronouns could use the present work as a starting point and target essentialism in a more comprehensive fashion. For instance, teaching people about the complexity of biological sex and gender has been shown to reduce binary gender essentialism (Donovan et al., 2019; Fausto-Sterling, 2000). Participants learning such information over a longer period may show more substantial changes in their essentialism than we achieved here using brief articles and thus may

show more substantial corresponding changes in their beliefs about pronouns.

In addition, manipulating individuals' essentialism may have had a weak effect because lay beliefs about pronouns may be particularly complex and nuanced, consistent with the complexity in the academic and public debates on the benefits and drawbacks of gendered language. For instance, there are arguments for and against binary gender pronouns at both ends of the ideological spectrum: Some advocate for binary gender pronouns while marginalizing non-binary people—for example, arguing that only binary gender pronouns are necessary even for people who do not have binary gender identities (Murphy, 2016). In contrast, some advocate for binary gender pronouns by centering transgender people—for example, arguing that pronouns are an important way that people, and especially transgender people, convey their gender identity to others and have that identity affirmed by others (Hanna et al., 2019). Thus, in our studies, some participants high in gender essentialism may have endorsed binary gender pronouns believing that they reinforce a fixed gender binary, while some participants low in gender essentialism may have *also* endorsed binary gender pronouns but for a very different reason: because they believe that gendered pronouns allow all people, including transgender people, to express and affirm their gender identity. The complexity of these ideological debates (e.g., binary gender pronouns have even been raised with respect to animal rights; Simon, 2021) highlights the importance of conducting more research on lay beliefs to understand how lay people reason about identity-relevant language.

A further limitation of our work is that we only included speakers of two languages. Although English and Turkish provide two valuable language contexts given the theory-relevant differences in their pronoun systems, our claims about the language-generality of some of the mechanisms investigated here (particular, social ideologies) would be strengthened by recruiting

speakers of additional languages worldwide. Given the differences between English and Turkish pronoun systems and the differences between our English- and Turkish-speaking samples (e.g., in recruitment procedures), it is particularly remarkable that we find such consistent evidence that social ideologies predict relative endorsement of identity-relevant pronouns.

Conclusion

In academic debates, psychologists, linguists, and philosophers point out several benefits and drawbacks of gendered language including binary gender pronouns (for reviews, see Atir, 2022; Saguy & Williams, 2019). To name a few proposed benefits, binary gender pronouns facilitate effective communication (Sanford & Filik, 2007), convey gender identity in shorthand (e.g., "My pronouns are he/him"), and make women visible in male-dominated contexts (e.g., "The surgeon operated on the patient; she was very skilled"; Morehouse et al., 2022). To name a few proposed drawbacks, binary gender pronouns reinforce the idea that gender is binary, exclude gender non-binary individuals, and may promote gender hierarchy that disadvantages women by making gender "omnirelevant" (Bigler & Leaper, 2015). In these debates, some have pointed out that other languages function without binary gender pronouns and have advocated for eliminating gender pronouns entirely in English (Dembroff & Wodak, 2018; 2021). The present work informs these debates by investigating lay beliefs about pronouns in English and Turkish. People's reasoning about pronouns appears to be influenced both by familiarity with the specific pronoun types in their language (i.e., the language-specific status quo) and by people's ideological commitments—in particular, believing social groups are inherently distinct (essentialism) and ought to be unequal (social dominance orientation). Our findings belie the possibility that English speakers' commitments to binary gender pronouns are merely a matter of familiarity or ease and instead show that ideology is also playing a role.

References

Arnold, J. E., Mayo, H. C., & Dong, L. (2021). My pronouns are they/them: Talking about pronouns changes how pronouns are understood. *Psychonomic Bulletin & Review, 28*(5), 1688-1697.

Atir, S. (2022). Girlboss? Highlighting versus downplaying gender through language. *Trends in Cognitive Sciences*, 26(8), P623-625.

Axt, J. R., Conway, M. A., Westgate, E. C., & Buttrick, N. R. (2021). Implicit transgender attitudes independently predict beliefs about gender and transgender people. *Personality and Social Psychology Bulletin*, *47*(2), 257-274.

Bäck, E. A., Lindqvist, A., & Gustafsson Sendén, M. (2015). Hen can do it!: Effects of using a gender-neutral pronoun in recruitment. In J. Magnusson, K. Milles, & Z. Nikolaidou (Eds.), *Gender constructions and language change: Report from the 8th Nordic Conference on Language and Gender* (pp. 71–90). Flemingsberg, Sweden: Södertörn University.

Bailey, A. H., & Knobe, J. (2023). Biological essentialism correlates with (but doesn't cause?) intergroup bias. *Personality and Social Psychology Bulletin.*

Bailey, A. H., Knobe, J., & Newman, G. E. (2021). Value-based essentialism: Essentialist beliefs about social groups with shared values. *Journal of Experimental Psychology: General, 150*(10), 1994–2014.

Bailey, A. H., Williams, A., & Cimpian, A. (2022). Based on billions of words on the internet, PEOPLE = MEN. *Science Advances*, *8*(13), eabm2463.

Baron, D. (2010, August 26). The gender-neutral pronoun: 150 years later, still an epic fail. *OUPblog*. https://blog.oup.com/2010/08/gender-neutral-pronoun/

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, *4*(3), 265-284.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bigler, R. S., & Leaper, C. (2015). Gendered language: Psychological principles, evolving practices, and inclusive policies. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), 187-194.

Blanchar, J. C., & Eidelman, S. (2019). Everything in its right place: Tradition, order, and the legitimation of long-standing inequality. In *The Social Psychology of Inequality* (pp. 349-363). Springer.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin, 106*(2), 265–289.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*(1), 1-22.

Bradley, E. D. (2020). The influence of linguistic and social attitudes on grammaticality judgments of singular 'they'. *Language Sciences, 78*, 101272.

Bradley, E. D., Salkind, J., Moore, A., & Teitsort, S. (2019). Singular 'they' and novel pronouns: Gender-neutral, nonbinary, or both? *Proceedings of the Linguistic Society of America, 4*(36), 1-7.

Brescoll, V., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on sex differences. *Psychological Science, 15*(8), 515-520.

Broussard, K. A., Warner, R. H., & Pope, A. R. (2018). Too many boxes, or not enough?

Preferences for how we ask about gender in cisgender, LGB, and gender-diverse samples. *Sex Roles*, *78*(9), 606-624.

Burgess, A., & Plunkett, D. (2013). Conceptual ethics I. *Philosophy Compass*, *8*(12), 1091-1101.

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306-307.

Chen, J. M., & Ratliff, K. A. (2018). Psychological essentialism predicts intergroup bias. *Social Cognition, 36*(3), 301-323.

Ching, B. H. H., & Xu, J. T. (2018). The effects of gender neuroessentialism on transprejudice: An experimental study. *Sex Roles, 78*(3), 228-241.

Christy, A. G., Schlegel, R. J., & Cimpian, A. (2019). Why do people believe in a "true self"? The role of essentialist reasoning about personal identity and the self. *Journal of Personality and Social Psychology, 117*(2), 386-416.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, 7(4), 173-179.

Ćwiek, A., Fuchs, S., Draxler, C., Asu, E. L., Dediu, D., Hiovain, K., ... & Winter, B. (2022). The bouba/kiki effect is robust across cultures and writing systems. *Philosophical Transactions of the Royal Society B*, *377*(1841), 20200390.

Davies, M. (2008-2023). *The corpus of contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/

DeFranza, D., Mishra, H., & Mishra, A. (2020). How language shapes prejudice against women: An examination across 45 world languages. *Journal of Personality and Social Psychology, 119*(1), 7-22.

Dembroff, R., & Wodak, D. (2018). He/she/they/ze. *Ergo: An Open Access Journal of Philosophy*.

Dembroff, R., & Wodak, D. (2021). How much gender is too much gender?. In *The Routledge Handbook of Social and Political Philosophy of Language* (pp. 362-376). Routledge.

Donovan, B. M. (2022). Ending genetic essentialism through genetics education. *Human Genetics and Genomics Advances*, *3*(1), 100058.

Donovan, B. M., Stuhlsatz, M. A., Edelson, D. C., & Buck Bracey, Z. E. (2019). Gendered genetics: How reading about the genetic basis of sex differences in biology textbooks could affect beliefs associated with science gender disparities. *Science Education, 103*(4), 719-749.

Eagly, A. H., & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *European Review of Social Psychology*, *5*(1), 1-35.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Fausto-Sterling, A. (2000). *Sexing the body: Gender politics and the construction of sexuality*. Basic Books.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902.

Foertsch, J., & Gernsbacher, M. A. (1997). In search of gender neutrality: Is singular they a cognitively efficient substitute for generic he?. *Psychological Science*, *8*(2), 106-111.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25-42.

Gabriel, U., Gygax, P. M., & Kuhn, E. A. (2018). Neutralising linguistic sexism: Promising but cumbersome? *Group Processes & Intergroup Relations, 21*(5), 844–858. https://doi.org/10.1177/1368430218771742

Gallagher, N. M., & Bodenhausen, G. V. (2021). Gender essentialism and the mental representation of transgender women and men: A multimethod investigation of stereotype content. *Cognition, 217*, 104887.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Pres.

Gelman, S. A., & Taylor, M. G. (2000). Gender essentialism in cognitive development. In P. H. Miller & E. Kofsky Scholnick (Eds.), *Toward a feminist developmental psychology* (pp. 169–190). Taylor & Frances/Routledge.

Glick, P., & Fiske, S. T. (1997). Hostile and benevolent sexism: Measuring ambivalent sexist attitudes toward women. *Psychology of Women Quarterly*, *21*(1), 119-135.

Green P., & MacLeod, C. J. (2016). simr: an R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, **7**(4), 493–498.

Guiora, A. Z., Beit-Hallahmi, B., Fried, R., & Yoder, C. (1982). Language environment and gender identity attainment. *Language Learning*, *32*(2), 289-304.

Gustafsson Sendén, M., Bäck, E. A., & Lindqvist, A. (2015). Introducing a gender-neutral pronoun in a natural gender language: The influence of time on attitudes and behavior. *Frontiers in Psychology*, *6*, 893.

Gustafsson Sendén, M., Renström, E., & Lindqvist, A. (2021). Pronouns beyond the binary: The

change of attitudes and use over time. *Gender & Society*, *35*(4), 588-615.

Haslam, N., & Levy, S. R. (2006). Essentialist beliefs about homosexuality: Structure and implications for prejudice. *Personality and Social Psychology Bulletin*, *32*(4), 471-485.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*(1), 113-127.

Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to language. *Nature*, *430*(6998), 453-456.

Hilliard, L. J., & Liben, L. S. (2010). Differing levels of gender salience in preschool classrooms: Effects on children's gender attitudes and intergroup bias. *Child Development*, *81*(6), 1787-1798.

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., ... & Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO_7 scale. *Journal of Personality and Social Psychology, 109*(6), 1003.

Hofstadter, D. R. (1985). A person paper on purity in language. In *Metamagical themas; Questing for the essence of mind and pattern* (pp. 159-172. Basic Books.

Hussak, L. J., & Cimpian, A. (2015). An early-emerging explanatory heuristic promotes support for the status quo. *Journal of Personality and Social Psychology, 109*(5), 739–752.

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist, 74*(2), 171-193.

JASP Team. (2023). *JASP* (Version 0.18.1) [Computer software].

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory:

Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*(6), 881-919.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kapusta, S. J. (2016). Misgendering and its moral contestability. *Hypatia*, *31*(3), 502-519.

Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. Liveright.

Kreft, I. G., & De Leeuw, J. (1998). *Introducing multilevel modeling*. Sage.

Lakoff, R. T. (2004). *Language and women's place* (M. Bucholtz, Ed., Revised Edition). Oxford University Press.

Leventhal, O., Camilliere, S., Chen, Peiyao, & Grodner, D. (2020). *Using ERPs to investigate the processing of singular* they. Poster presented at the CUNY conference on human sentence processing, Amherst, MA.

Lewis, M., & Lupyan, G. (2020). Gender stereotypes are reflected in the distributional structure of 25 languages. *Nature Human Behaviour*, *4*(10), 1021-1028.

Mandalaywala, T. M. (2020). Does essentialism lead to racial prejudice? It is not so Black and White. *Advances in Child Development and Behavior, 59*, 195-245.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). Cambridge University Press.

Meyer, A., Zhou, E., & Shane, F. (2018). The non-effects of repeated exposure to the Cognitive Reflection Test. *Judgment and Decision Making*, *13*(3), 246-259.

Morehouse, K. N., Kurdi, B., Hakim, E., & Banaji, M. R. (2022). When a stereotype dumbfounds: Probing the nature of the surgeon= male belief. *Current Research in Ecological and Social Psychology*, *3*, 100044.

Morgenroth, T., & Ryan, M. K. (2021). The effects of gender trouble: An integrative theoretical framework of the perpetuation and disruption of the gender/sex binary. *Perspectives on Psychological Science*, *16*(6), 1113-1142.

Muradoglu, M., Cimpian, J. R., & Cimpian, A. (2023). Mixed-effects models for cognitive development researchers. *Journal of Cognition and Development*, 1-34.

Murphy, J. (2016, November 4). Toronto professor Jordan Peterson takes on gender-neutral pronouns. *British Broadcasting Corporation.* https://www.bbc.com/news/world-us-canada-37875695

Norton, A. T., & Herek, G. M. (2013). Heterosexuals' attitudes toward transgender people: Findings from a national probability sample of US adults. *Sex roles*, *68*(11), 738-753.

Parker, K., Graf, N., & Igielnik, R. (2019). *Generation Z looks a lot like Millennials on key social and political issues.* Pew Research Center, from https://www.pewresearch.org/social-trends/2019/01/17/generation-z-looks-a-lot-like-millennials-on-key-social-and-political-issues/

Pennycook, G. & Rand, D. G. (2019). Lazy, no biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition, 188*, 39-50.

Peretz-Lange, R. (2021). Why does social essentialism sometimes promote, and other times mitigate, prejudice development? A causal discounting perspective. *Cognitive Development*, *59*, 101085.

Perez-Arche, H., & Miller, D. J. (2021). What predicts attitudes toward transgender and nonbinary people? An exploration of gender, authoritarianism, social dominance, and gender ideology. *Sex Roles*, *85*(3), 172-189.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A

personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67*(4), 741–763.

Prewitt-Freilino, J. L., Caswell, T. A., & Laakso, E. K. (2012). The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and genderless languages. *Sex Roles*, *66*(3), 268-281.

Prusaczyk, E., & Hodson, G. (2020). The roles of political conservatism and binary gender beliefs in predicting prejudices toward gay men and people who are transgender. *Sex Roles*, *82*(7), 438-446.

Rad, M. S., Shackleford, C., Lee, K. A., Jassin, K., & Ginges, J. (2019). Folk theories of gender and anti-transgender attitudes: Gender differences and policy preferences. *PloS One*, *14*(12), e0226967.

Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences*, *13*(10), 439-446.

Renström, E. A., Lindqvist, A., & Sendén, M. G. (2022). The multiple meanings of the gender-inclusive pronoun hen: Predicting attitudes and use. *European Journal of Social Psychology*, *52*(1), 71-90.

Renström, E. A., Lindqvist, A., Akbas, G., Hekanaho, L., & Sendén, M. G. (2022). Are gender-neutral pronouns really neutral? Testing a male bias in the grammatical genderless languages Turkish and Finnish. *Journal of Language and Social Psychology*.

Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, *59*(3), 244-274.

Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social

essentialism. *Proceedings of the National Academy of Sciences*, *109*(34), 13526-13531.

Ritchie, K. (2021). Essentializing inferences. *Mind & Language*, *36*(4), 570-591.

Saguy, A. C., & Williams, J. A. (2019). Why we should all use they/them pronouns. *Scientific American.*

Saguy, T., Reifen-Tagar, M., & Joel, D. (2021). The gender-binary cycle: the perpetual relations between a biological-essentialist view of gender, gender ideology, and gender-labelling and sorting. *Philosophical Transactions of the Royal Society B*, *376*(1822), 20200141.

Salomon, E., & Cimpian, A. (2014). The inherence heuristic as a source of essentialist thought. *Personality and Social Psychology Bulletin, 40*(10), 1297-1315.

Sanford, A. J., & Filik, R. (2007). "They" as a gender-unspecified singular pronoun: Eye tracking reveals a processing cost. *Quarterly Journal of Experimental Psychology, 60*(2), 171–178.

Sarrasin, O., Gabriel, U., & Gygax, P. (2012). Sexism and attitudes toward gender-neutral language: The case of English, French, and German. *Swiss Journal of Psychology, 71*(3), 113–124. https://doi.org/10.1024/1421-0185/a000078

Schudson, Z. C., & van Anders, S. M. (2021). Gender/sex diversity beliefs: Scale construction, validation, and links to prejudice. *Group Processes & Intergroup Relations, 25*(4), 1011-1036.

Sczesny, S., Formanowicz, M., & Moser, F. (2016). Can gender-fair language reduce gender stereotyping and discrimination? *Frontiers in Psychology, 7*, 25.

Sczesny, S., Moser, F., & Wood, W. (2015). Beyond sexist beliefs: How do people decide to use gender-inclusive language? *Personality and Social Psychology Bulletin, 41*(7), 943-954. https://doi.org/10.1177/0146167215585727

Sibley, C. G., Wilson, M. S., & Duckitt, J. (2007). Antecedents of men's hostile and benevolent sexism: The dual roles of social dominance orientation and right-wing authoritarianism. *Personality and Social Psychology Bulletin*, *33*(2), 160–172.

Simon, S. (2021). Animals deserve gender pronouns, too. *National Public Radio.* https://www.npr.org/2021/04/03/984008349/opinion-animals-deserve-gender-pronouns-too

Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representation of the sexes in language. *Social Communication*, 163-187.

StataCorp. (2023). *Stata Statistical Software: Release 18* [Computer software]. StataCorp LLC.

Sutherland, S. L., & Cimpian, A. (2015). An explanatory heuristic gives rise to the belief that words are well suited for their referents. *Cognition*, *143*, 228-240.

Tavits, M., & Pérez, E. O. (2019). Language influences mass opinion toward gender and LGBT equality. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(34), 16781–16786.

Tee, N., & Hegarty, P. (2006). Predicting opposition to the civil rights of trans persons in the United Kingdom. *Journal of Community & Applied Social Psychology*, *16*(1), 70-80.

Tworek, C. M., & Cimpian, A. (2016). Why do people tend to infer "ought" from "is"? The role of biases in explanation. *Psychological Science, 27*(8), 1109–1122.

Vergoossen, H. P., Pärnamets, P., Renström, E. A., & Gustafsson Sendén, M. (2020). Are new gender-neutral pronouns difficult to process in reading? The case of *hen* in Swedish. *Frontiers in Psychology*, *11*, 2967.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (J. B. Carroll, Ed.). MIT Press.

Wilton, L. S., Bell, A. N., Carpinella, C. M., Young, D. M., Meyers, C., & Clapham, R. (2019). Lay theories of gender influence support for women and transgender people's legal rights. *Social Psychological and Personality Science*, *10*(7), 883-894.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, *104*(19), 7780-7785.

Wodak, D., Leslie, S. J., & Rhodes, M. (2015). What a loaded generalization: Generics and social cognition. *Philosophy Compass*, *10*(9), 625-635.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2), 1–27.

People's Beliefs About Pronouns Reflect Both the Language They Speak and Their Ideologies

SUPPLEMENTARY INFORMATION

Study 1

SI.1.1 Analyses of the Inherence Bias and Need for Cognition Measures

Items measuring the inherence bias ($\alpha = 0.88$) and need for cognition ($\alpha = 0.97$) were combined into two respective indices. As described in the main text, the inherence bias scale also included items designed as exclusion criteria (Salomon & Cimpian, 2014); these were not combined into the index.

To assess the relationship between pronoun endorsement and, respectively, the inherence bias and need for cognition, we conducted two linear mixed-effects models with the relevant process (e.g., inherence bias; a continuous variable), pronoun type (a six-level categorical variable), and their interaction terms predicting pronoun endorsement. Each model included a random intercept for participant. Because we had six types of pronouns, the interaction between each social-cognitive process and pronoun type was represented in the model by five terms. Thus, to compute the interaction between the process and pronoun type, we used a likelihood ratio test to compare this model to an identical model without the five interaction terms. This model comparison indicated whether or not there was a significant process × pronoun type interaction. To interpret this interaction, we focused especially on the simple slopes for each process and each pronoun type (Table S1).

Inherence bias and need for cognition were inversely correlated ($r = -.22$, $p = .001$) and tended to have corresponding inverse relationships with pronoun endorsement. For instance, in the simple slopes, inherence bias was negatively correlated with endorsement of identity-neutral pronouns and positively correlated with endorsement of binary gender pronouns, but need for cognition was positively correlated with endorsement identity-neutral pronouns. This pattern suggests that, as expected, inherence bias and need for cognition both capture some overlapping tendencies toward heuristic versus analytic thinking about pronouns (see Table S1 for statistics).

Table S1

Relationships Between Social-Cognitive Processes and Endorsement of Six Pronoun Types in Study 1

Construct	Term	Separate Models for Each Process		
		χ^2 (*df*)	*B* (*SE*)	β
IB	IB × Pronoun Type	23.87 (5) ***		
	IB × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.35 (0.09) ***	−0.26
	IB × Race (= 1) vs. Gender (= 0) Pronouns		−0.06 (0.09)	−0.04
	IB × Wealth (= 1) vs. Gender (= 0) Pronouns		< 0.01 (0.09)	< 0.01
	IB × Weight (= 1) vs. Gender (= 0) Pronouns		−0.09 (0.09)	−0.07
	IB × Age (= 1) vs. Gender (= 0) Pronouns		−0.02 (0.09)	−0.01
	Simple slope of IB for Gender Pronouns		0.16 (0.09) *	0.12
	Simple slope of IB for Race Pronouns		0.10 (0.09)	0.08
	Simple slope of IB for Neutral Pronouns		−0.19 (0.09) *	−0.14
	Simple slope of IB for Wealth Pronouns		0.16 (0.09) *	0.12
	Simple slope of IB for Weight Pronouns		0.07 (0.09)	0.05
	Simple slope of IB for Age Pronouns		0.14 (0.09)	0.11
NFC	NFC × Pronoun Type	21.74 (5) ***		
	NFC × Neutral (= 1) vs. Gender (= 0) Pronouns		0.12 (0.06) *	0.13
	NFC × Race (= 1) vs. Gender (= 0) Pronouns		−0.13 (0.06) *	−0.15
	NFC × Wealth (= 1) vs. Gender (= 0) Pronouns		−0.07 (0.06)	−0.07
	NFC × Weight (= 1) vs. Gender (= 0) Pronouns		−0.08 (0.06)	−0.09
	NFC × Age (= 1) vs. Gender (= 0) Pronouns		−0.04 (0.06)	−0.04
	Simple slope of NFC for Gender Pronouns		−0.01 (0.05)	−0.01
	Simple slope of NFC for Race Pronouns		−0.14 (0.05) **	−0.16
	Simple slope of NFC for Neutral Pronouns		0.11 (0.05) *	0.12
	Simple slope of NFC for Wealth Pronouns		−0.08 (0.05)	−0.09
	Simple slope of NFC for Weight Pronouns		−0.09 (0.05)	−0.10
	Simple slope of NFC for Age Pronouns		−0.05 (0.05)	−0.06

Note. IB = Inherence bias. NFC = Need for cognition. This table shows two different models, each with a single process and its interaction with pronoun type predicting pronoun endorsement, with a random intercept for participant. The beta coefficient (β) is a measure of effect size derived from a model with standardized process and standardized pronoun endorsement.

*$p < .05$. **$p < .01$. ***$p < .001$.

Study 2a

SI.2a.1 Additional Details About the Language-metaphysics Link Measure

Because the LML measure was created for this study, we conducted further preregistered analyses to explore its factor structure in addition to computing Cronbach's alpha ($\alpha = 0.88$). Based on examination of the scree plot (Figure S1), we retained a single factor (eigenvalue = 4.29). This factor explained 43% of the variance, and all items loaded onto it at ≥ 0.54. See Table S2 for a complete list of items.

Figure S1

Scree Plot Concerning the LML Measure in Study 2a

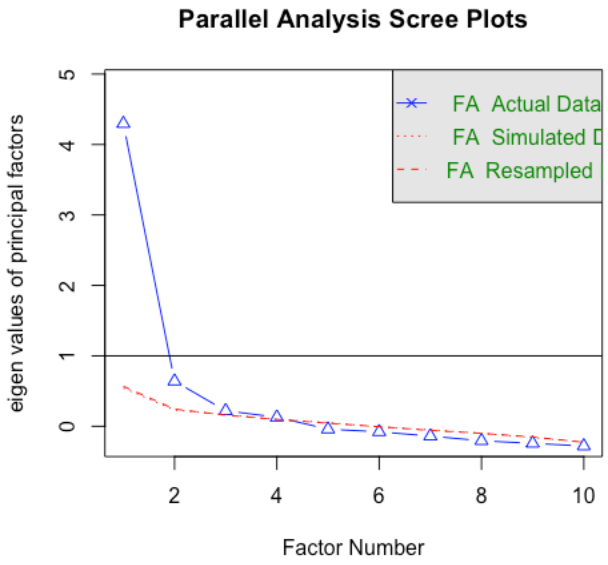


Table S2

Language-metaphysics Link Items in Study 2a

Item	Scale
1. If a language uses different words for groups that seem similar, there MUST be some deeper difference between those groups.	*1=strongly disagree, 9=strongly agree*
2. If a language uses a single word for groups that seem different from each other, there MUST be a deeper similarity at play.	*1=strongly disagree, 9=strongly agree*
3. Two groups MUST be the same if a language only has a single word for them, even if it is not obvious how they are the same.	*1=strongly disagree, 9=strongly agree*
4. Languages have different words for groups that seem similar BECAUSE those groups really are different.	*1=strongly disagree, 9=strongly agree*
5. If a language uses different words for groups that seem similar, people MUST have found it useful to have separate words.	*1=strongly disagree, 9=strongly agree*
6. Speakers MUST find it useful if a language only has a single word for two groups, even if it is not obvious how it is useful.	*1=strongly disagree, 9=strongly agree*
7. Languages have different words for groups that seem similar BECAUSE it is useful to do so.	*1=strongly disagree, 9=strongly agree*
8. If a language has different words for groups that seem similar, other languages will also have multiple words for those same groups.	*1=strongly disagree, 9=strongly agree*
9. If a language uses a single word for groups that seem different, other languages will also have a single word for those same groups.	*1=strongly disagree, 9=strongly agree*
10. If a language uses a single word for groups that seem different from each other, people MUST have found it useful to have just one word.	*1=strongly disagree, 9=strongly agree*

SI.2a.2 Simultaneous Model with Composite Scores for Cognitive Heuristics and Social Ideologies

Study 2a included multiple measures of cognitive heuristics (CRT, longevity bias, and inherence bias) and social ideologies (gender essentialism, race essentialism, and social dominance orientation), which raises the possibility that conceptually similar and statistically interrelated measures "cancelled each other out" when entered into the same model simultaneously. To address this concern, we conducted additional, non-preregistered exploratory analyses.

Factor analysis. We conducted a factor analysis on the six scale-specific composite scores for each participant (three cognitive heuristics, three social ideologies). Examination of the scree plot suggested one factor. However, given that there were two a priori conceptual groupings (i.e., cognitive heuristics and social ideologies), we nevertheless requested that *two* factors be extracted in an exploratory varimax-rotated factor analysis. Gender essentialism, race essentialism, and social dominance orientation loaded most strongly on the first factor, which had an eigenvalue of 2.16 and explained 36% of the variance (Table S3). CRT, inherence bias, and longevity bias loaded most strongly on the second factor, which had an eigenvalue of 1.18 and explained 20% of the variance. Note that the CRT had a negative loading, which is sensible because higher scores on this measure indicate more reflective and *less* heuristic thinking.

In light of this factor analysis, we created a cognitive heuristic composite (averaging inherence bias, longevity bias, and reverse-scored CRT) and a social ideology composite (averaging gender essentialism, race essentialism, and social dominance orientation). The individual variables were standardized before averaging into the relevant composite because they were on different scales. Standardizing brought them on the same scale ($M = 0$, $SD = 1$).

Simultaneous model with composite scores. Similar to the simultaneous analysis reported in the main text, we entered the cognitive heuristic composite, the social ideology composite, and the language–metaphysics link measure as predictors of pronoun endorsement, along with pronoun type (a three-level categorical variable) and their six interaction terms (Table S4). We included a random intercept for participant.

In this analysis, the social ideology composite stood out as a key predictor compared to the cognitive heuristics composite and the language–metaphysics link measure. Higher scores on the social ideology composite related to higher endorsement of binary gender pronouns, higher endorsement of race pronouns, and lower endorsement of identity-neutral pronouns.

Overall, this supplementary analysis points to a similar conclusion as the preregistered analysis reported in the main text. Even accounting for cognitive heuristics, participants who endorsed hierarchy-enhancing social ideologies more strongly also showed stronger endorsement of binary gender pronouns and even relative openness to race pronouns.

Table S3
Factor Loadings

Measure	Factor 1	Factor 2
Longevity bias	0.60	**0.61**
Inherence bias	0.50	**0.67**
Cognitive Reflection Test	−0.08	**−0.33**
Gender essentialism	**0.98**	0.19
Race essentialism	**0.64**	0.40
Social dominance orientation	**0.41**	0.22

Table S4
Relationships Among Social-Cognitive Process Composites and Pronoun Types in Study 2a in a Simultaneous Model

Construct	Term		
		B (*SE*)	β
Cognitive heuristic composite	Composite × Neutral (= 1) vs. Gender (= 0) Pronouns	−0.45 (0.22) *	−0.34
	Composite × Race (= 1) vs. Gender (= 0) Pronouns	−0.02 (0.22)	−0.02
	Composite × Neutral (= 1) vs. Race (= 0) Pronouns	−0.42 (0.22)	−0.33
	Simple slope of Composite for Gender Pronouns	0.15 (0.16)	0.12
	Simple slope of Composite for Race Pronouns	0.13 (0.16)	0.10
	Simple slope of Composite for Neutral Pronouns	−0.29 (0.16)	−0.22
Social ideology composite	Composite × Neutral (= 1) vs. Gender (= 0) Pronouns	−0.75 (0.19) ***	−0.61
	Composite × Race (= 1) vs. Gender (= 0) Pronouns	−0.09 (0.19)	−0.07
	Composite × Neutral (= 1) vs. Race (= 0) Pronouns	−0.66 (0.19) ***	−0.54
	Simple slope of Composite for Gender Pronouns	0.46 (0.14) ***	0.37
	Simple slope of Composite for Race Pronouns	0.37 (0.14) **	0.30
	Simple slope of Composite for Neutral Pronouns	−0.29 (0.14) *	−0.24
Language-metaphysics link	LML × Neutral (= 1) vs. Gender (= 0) Pronouns	0.01 (0.10)	0.02
	LML × Race (= 1) vs. Gender (= 0) Pronouns	0.16 (0.10)	0.24
	LML × Neutral (= 1) vs. Race (= 0) Pronouns	−0.15 (0.10)	−0.22
	Simple slope of LML for Gender Pronouns	0.01 (0.07)	0.01
	Simple slope of LML for Race Pronouns	0.17 (0.07) *	0.26
	Simple slope of LML for Neutral Pronouns	0.02 (0.07)	0.03

Note. LML = language–metaphysics link measure. This table shows a single model with the cognitive heuristic composite, the social ideology composite, and the language–metaphysics link composite and their interactions with pronoun type entered simultaneously to predict pronoun endorsement, with a random intercept for participant. The beta coefficient (β) is a measure of effect size derived from a model with standardized process and standardized pronoun endorsement.
*$p < .05$. **$p < .01$. ***$p < .001$.

SI.2a.3 Analyses With the Full Sample

Because the preregistered exclusion criteria resulted in unusually high exclusion rates in the Mechanical Turk sample, we report results from the full sample here. This analysis was not preregistered and is included for exploratory purposes. None of the overall conclusions reported in the main text change in these analyses with the full sample.

We compared a linear mixed-effects model with pronoun type predicting pronoun endorsement and a random intercept for participant to an intercept-only model. We found a significant difference in the fit of these models, $\chi^2(2) = 436.37$, $p < .001$, indicating a significant main effect of pronoun type. In follow-up analyses, we found that binary gender pronouns were endorsed the most and more than identity-neutral pronouns, $B = -0.27$, $SE = 0.08$, $p < .001$, $\beta = -0.18$. In turn, identity-neutral pronouns were endorsed more than race pronouns, $B = -1.35$, $SE = 0.08$, $p < .001$, $\beta = -0.94$. These findings were the same in direction and significance as those reported in the main text. Note that this model again exhibited singular fit.

To assess the relationship between each social-cognitive process and pronoun endorsement, we conducted a series of linear mixed-effects models with the relevant process (e.g., gender essentialism; a continuous variable), pronoun type (a three-level categorical variable), and their interaction terms predicting pronoun endorsement. We also computed a model with all processes and their interactions with pronoun type entered simultaneously. In this simultaneous model, which is the most informative analysis because it estimates the relationship of each construct controlling for the others, all of the results were the same in direction and significance as those reported in the main text—the only exception to this was that the simple slope for race pronouns for the language–metaphysics link measure was now non-significant at $p = .056$ (see Table S5).

Table S5

Relationships Between Seven Social-Cognitive Processes and Endorsement of Three Pronoun Types in Study 2a, in Separate and Simultaneous Models with the Full Sample

Construct	Term	Separate Models for Each Process			Simultaneous Model	
		χ^2 (*df*)	*B* (*SE*)	β	*B* (*SE*)	β
LB	LB × Pronoun Type	175.98 (2) ***				
	LB × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.31 (0.06) ***	−0.26	−0.16 (0.15)	−0.14
	LB × Race (= 1) vs. Gender (= 0) Pronouns		0.53 (0.06) *** †	0.44	−0.05 (0.15)	−0.04
	LB × Neutral (= 1) vs. Race (= 0) Pronouns		−0.85 (0.06) ***	−0.70	−0.11 (0.15)	−0.09
	Simple slope of LB for Gender Pronouns		0.03 (0.04) †	0.02	0.04 (0.11)	0.03
	Simple slope of LB for Race Pronouns		0.56 (0.04) ***	0.47	−0.02 (0.11)	0.01
	Simple slope of LB for Neutral Pronouns		−0.29 (0.04) ***	−0.24	−0.13 (0.11)	−0.11
IB	IB × Pronoun Type	42.72 (2) ***				
	IB × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.44 (0.07) ***	−0.46	−0.03 (0.11)	−0.04
	IB × Race (= 1) vs. Gender (= 0) Pronouns		−0.03 (0.07)	−0.03	−0.04 (0.11)	−0.04
	IB × Neutral (= 1) vs. Race (= 0) Pronouns		−0.41 (0.07) ***	−0.33	<0.01 (0.11)	<0.01
	Simple slope of IB for Gender Pronouns		0.27 (0.05) ***	0.29	0.11 (0.08)	0.11
	Simple slope of IB for Race Pronouns		0.24 (0.05) ***	0.20	0.07 (0.08)	0.07
	Simple slope of IB for Neutral Pronouns		−0.16 (0.05) **	−0.17	0.07 (0.08)	0.08
CRT	CRT × Pronoun Type	200.31 (2) ***				
	CRT × Neutral (= 1) vs. Gender (= 0) Pronouns		0.14 (0.20) †	0.03	0.39 (0.31)	0.10
	CRT × Race (= 1) vs. Gender (= 0) Pronouns		−2.46 (0.20) ***	−0.62	−0.58 (0.31)	−0.15
	CRT × Neutral (= 1) vs. Race (= 0) Pronouns		2.59 (0.20) ***	0.65	0.97 (0.31) **	0.24
	Simple slope of CRT for Gender Pronouns		0.97 (0.14) *** †	0.24	0.36 (0.22)	0.09
	Simple slope of CRT for Race Pronouns		−1.49 (0.14) ***	−0.37	−0.22 (0.22)	−0.06
	Simple slope of CRT for Neutral Pronouns		1.11 (0.14) ***	0.28	0.75 (0.22) ***	0.19
GE	GE × Pronoun Type	138.99 (2) ***				
	GE × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.30 (0.04) ***	−0.36	−0.44 (0.09) ***	−0.54
	GE × Race (= 1) vs. Gender (= 0) Pronouns		0.21 (0.04) *** †	0.26	−0.55 (0.10) ***	−0.67
	GE × Neutral (= 1) vs. Race (= 0) Pronouns		−0.51 (0.04) ***	−0.62	0.11 (0.10)	0.13
	Simple slope of GE for Gender Pronouns		0.11 (0.03) ***	0.14	0.37 (0.07) ***	0.45
	Simple slope of GE for Race Pronouns		0.32 (0.03) ***	0.39	−0.18 (0.07) **	−0.22
	Simple slope of GE for Neutral Pronouns		−0.18 (0.03) ***	−0.22	−0.08 (0.07)	−0.09
RE	RE × Pronoun Type	297.77 (2) ***				

	RE × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.11 (0.04) **	−0.15	0.16 (0.09)	0.22
	RE × Race (= 1) vs. Gender (= 0) Pronouns		0.50 (0.04) ***	0.68	0.45 (0.09) ***	0.61
	RE × Neutral (= 1) vs. Race (= 0) Pronouns		−0.61 (0.04) ***	−0.83	−0.28 (0.09) **	−0.38
	Simple slope of RE for Gender Pronouns		−0.07 (0.03) ** †	−0.09	−0.14 (0.07) *	−0.19
	Simple slope of RE for Race Pronouns		0.43 (0.03) ***	0.58	0.31 (0.07) ***	0.42
	Simple slope of RE for Neutral Pronouns		−0.18 (0.03) ***	−0.24	0.03 (0.07)	0.03
SDO	SDO × Pronoun Type	284.76 (2) ***				
	SDO × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.16 (0.06) **	−0.14	−0.13 (0.09)	−0.12
	SDO × Race (= 1) vs. Gender (= 0) Pronouns		0.76 (0.06) *** †	0.67	0.26 (0.09) **	0.23
	SDO × Neutral (= 1) vs. Race (= 0) Pronouns		−0.92 (0.06) ***	−0.81	−0.39 (0.09) ***	−0.34
	Simple slope of SDO for Gender Pronouns		−0.23 (0.04) *** †	−0.21	−0.08 (0.07)	−0.07
	Simple slope of SDO for Race Pronouns		0.53 (0.04) ***	0.47	0.18 (0.07) **	0.16
	Simple slope of SDO for Neutral Pronouns		−0.39 (0.04) ***	−0.35	−0.21 (0.07) **	−0.18
LML	LML × Pronoun Type	247.01 (2) ***				
	LML × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.12 (0.04) **	−0.14	−0.02 (0.10)	−0.02
	LML × Race (= 1) vs. Gender (= 0) Pronouns		0.54 (0.04)	0.63	0.11 (0.10)	0.12
	LML × Neutral (= 1) vs. Race (= 0) Pronouns		−0.66 (0.04) ***	−0.77	−0.13 (0.10)	−0.15
	Simple slope of LML for Gender Pronouns		−0.08 (0.03) ** †	0.09	0.02 (0.07)	0.03
	Simple slope of LML for Race Pronouns		0.46 (0.03) ***	0.54	0.13 (0.07) †	0.15
	Simple slope of LML for Neutral Pronouns		−0.20 (0.03) ***	−0.23	<0.01 (0.07)	<0.01

Note. LB = Longevity bias. IB = Inherence bias. CRT = Cognitive Reflection Test. GE = Gender essentialism. RE = Race essentialism. SDO = Social dominance orientation. LML = Language–metaphysics link. The "Separate Models" column shows seven different models, each with a single process and its interaction with pronoun type predicting pronoun endorsement, with a random intercept for participant. Note that these seven models exhibited singular fit. The rightmost column displays the results of an eighth model, in which each process and its interactions with pronoun type were entered simultaneously. The beta coefficient (β) is a measure of effect size derived from a model with standardized process(es) and standardized pronoun endorsement.

† Result changed in significance or in significance and direction in this analysis compared to that reported in the main text with the more restricted sample.

*p < .05. **p < .01. ***p < .001.

SI.2a.4 Analyses of Potential Moderation by Participant Gender

We tested whether participants' gender—comparing only women and men because non-binary individuals were a relatively small portion of the sample (2%)—might moderate the results reported in the main text. These analyses were not preregistered and were conducted for exploratory purposes. There was no evidence of moderation by participant gender, all *p*s > .26.

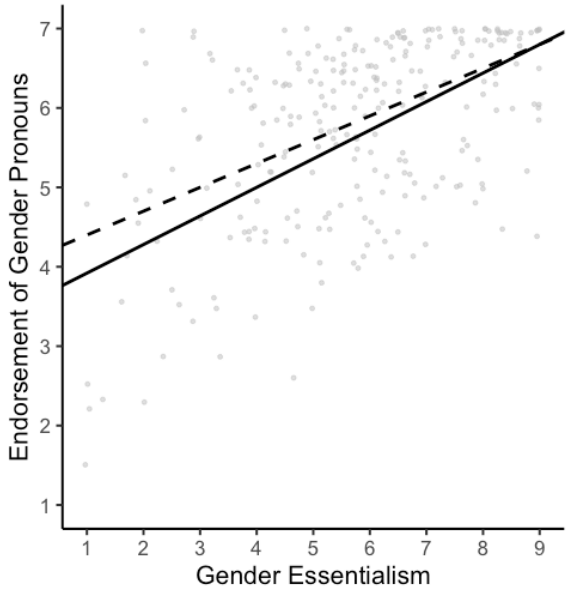
SI.2a.5 Additional Details About the Gender Essentialism Measure

One of the key findings reported in the main text was that gender essentialism was positively related to endorsement of binary gender pronouns. Figure S2 depicts this relation (i.e., endorsement of binary gender pronouns predicted by gender essentialism) with and without the other social-cognitive constructs entered as covariates.

Here, we report an additional analysis that followed up on this finding. We analyzed a single item from the eight-item measure of gender essentialism (for the full list, see Table S6 and the OSF repository for this project). Specifically, we analyzed the item "Gender is an all-or-none category; people are either male OR female, there is nothing in-between" because it expresses a specific commitment to the gender binary. Results from this analysis were largely similar to that of the complete gender essentialism scale reported in the main text (see Table S7 for details).

Figure S2

Relationship Between Gender Essentialism and Endorsement of Binary Gender Pronouns Without (Dashed Line) and With Covariates (Solid Line)



Note. Covariates = longevity bias, inherence bias, CRT score, race essentialism, social dominance orientation, and the language–metaphysics link measure. Raw data are underlaid (gray dots).

Table S6

Gender Essentialism Items in Studies 2 and 3

Item	Scale
1. Gender is a very important part of what makes people who they are.	*1=strongly disagree, 9=strongly agree*
2. People that are the same gender have many things in common.	*1=strongly disagree, 9=strongly agree*
3. Knowing someone's gender tells you a lot about a person.	*1=strongly disagree, 9=strongly agree*
4. Gender is an all-or-none category; people are either male OR female, there is nothing in between.	*1=strongly disagree, 9=strongly agree*
5. Gender is a natural category.	*1=strongly disagree, 9=strongly agree*
6. Gender categories are important in all cultures around the world.	*1=strongly disagree, 9=strongly agree*
7. Males share an underlying property that causes them to have many similarities.	*1=strongly disagree, 9=strongly agree*
8. Females share an underlying property that causes them to have many similarities.	*1=strongly disagree, 9=strongly agree*

Table S7

Interaction Between a Single Gender Essentialism Item (#4) and Pronoun Type in Study 2a

Term			
	χ^2 (*df*)	*B* (*SE*)	β
GE Item × Pronoun Type	45.60 (2) ***		
GE Item × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.23 (0.04) ***	−0.37
GE Item × Race (= 1) vs. Gender (= 0) Pronouns		−0.02 (0.04)	−0.04
GE Item × Neutral (= 1) vs. Race (= 0) Pronouns		−0.20 (0.04) ***	−0.34
Simple slope of GE Item for Gender Pronouns		0.11 (0.03) ***	0.19
Simple slope of GE Item for Race Pronouns		0.09 (0.03) ***	0.15
Simple slope of GE Item for Neutral Pronouns		−0.11 (0.03) ***	−0.19

Note. GE Item = Gender essentialism item: "Gender is an all-or-none category; people are either male OR female, there is nothing in between." The beta coefficient (β) is a measure of effect size derived from a model with standardized GE Item and standardized pronoun endorsement. *$p < .05$. **$p < .01$. ***$p < .001$.

SI.2a.6 Additional Analyses of Pronoun Endorsement

In supplementary analyses, we conducted more fine-grained analyses of the pronoun endorsement measure. That is, although all six pronoun endorsement items (see Table 1 in the main text) were correlated and averaged into a single index, these six items capture conceptually distinct beliefs about pronouns. In this supplementary analysis, we created two separate indices: The *pronoun endorsement—evaluative* index was an average of the "useful," "natural," "offensive," and "difficult" items ($\alpha_{gender} = .81$, $\alpha_{neutral} = .74$, $\alpha_{race} = .76$), and the *pronoun endorsement—prescriptive* index was an average of the "build a language" and "should" items ($\alpha_{gender} = .93$, $\alpha_{neutral} = .80$, $\alpha_{race} = .88$). These analyses did not change the conclusions reported in the main text. For simplicity, we report just the analysis comparing endorsement across pronoun types below, but the data file and analysis code for all analyses is available in the OSF repository for this project.

Separately for the evaluative and prescriptive sub-indices of pronoun endorsement, we first compared a linear mixed-effects model with pronoun type predicting pronoun endorsement and a random intercept for participant to an intercept-only model. For both sub-indices, we found a significant difference in the fit of these models, indicating a significant main effect of pronoun type (evaluative: $\chi^2(2) = 454.58$, $p < .001$; prescriptive: $\chi^2(2) = 458.78$, $p < .001$). In follow-up analyses for the evaluative sub-index, we found that binary gender pronouns were endorsed the most and more than identity-neutral pronouns, $B = -0.58$, $SE = 0.11$, $p < .001$, $\beta = -0.32$. In turn, identity-neutral pronouns were endorsed more than race pronouns, $B = -2.09$, $SE = 0.11$, $p < .001$, $\beta = -1.17$. Similarly, in follow-up analyses for the prescriptive sub-index, we found that binary gender pronouns were endorsed the most and directionally but non-significantly more than identity-neutral pronouns, $B = -0.08$, $SE = 0.15$, $p = .610$, $\beta = -0.03$. In turn, identity-neutral pronouns were endorsed more than race pronouns, $B = -3.12$, $SE = 0.15$, $p < .001$, $\beta = -1.34$. Overall, the pattern was fairly consistent across the two sets of sub-indices.

Study 2b

SI.2b.1 Analyses of Potential Moderation by Participant Gender

We tested whether participants' gender—comparing only women and men because non-binary individuals were a relatively small portion of the sample (1%)—might moderate the key results reported in the main text. These analyses were not preregistered and were conducted for exploratory purposes. There was no evidence of moderation by participant gender, all *p*s > .36.

Study 3

SI.3.1 Analyses of Potential Moderation by Participant Gender

We tested whether participants' gender—comparing only women and men because non-binary individuals were a relatively small portion of the sample (2%)—might moderate the results reported in the main text. These analyses were not preregistered and were conducted for exploratory purposes. There was no evidence of moderation by participant gender, all *p*s > .16.

SI.3.2 Article Manipulation Materials

Materials for the essentialism manipulation were adapted from Christy et al. (2019). The two key articles from the high- and low-essentialism conditions are replicated below for reference. Complete materials for all studies are also available in the OSF repository for this research.

High essentialism article:

THE TRUTH ABOUT THE GENDER BINARY: IT'S ALWAYS BEEN THERE AND ALWAYS WILL BE

OXFORD, UK – As people increasingly demand recognition for transgender and other non-binary gender identities (e.g., genderqueer, agender, gender nonconforming, etc.), we are confronting many important new questions, both legal and philosophical. Perhaps the most basic question that has been raised is whether the traditional man-woman gender binary has any firm basis in reality, or whether it is instead a mere social convention to categorize people in this manner.

This, like all fundamental questions of human nature, is a difficult question to answer. However, this is exactly the kind of question that anthropologists are trained to answer. Dr. Hazel Pretchik of Oxford University has devoted her entire career to studying the anthropology of gender. Dr. Pretchik has traveled the world observing and documenting how hundreds of different cultures approach gender, from small-scale hunter-gatherer communities in Africa and the Pacific Islands to fully industrialized societies in North America, Asia, and Europe. Based on her extensive experience, Dr. Pretchik felt that she could give a confident answer to the fundamental question: Is the gender binary real?

"Of course it's real!" says Dr. Pretchik. "In every society I have studied, they make a basic distinction between women on the one hand and men on the other. I have never seen or heard about a culture that doesn't recognize men and women as distinct groups." This means that a gender binary is what anthropologists call a cultural universal, a feature that is present in all known human societies. If it can be demonstrated that a certain behavior or practice is a cultural universal, this is strong evidence that it arises from our basic human nature. For example, language and music are found in every known human society, and we conclude from this that being speakers and music-makers are inherent parts of what it means to be human. Dr. Pretchik argues the same can be said of the man-woman binary.

"Not only do male and female gender concepts exist in every society I know of," Dr. Pretchik continued, "There is also a remarkable degree of consistency in the specific content of what it means to be male or female across cultures." According to Dr. Pretchik's research, there are cross-culturally consistent patterns of gender differences in appearance and dress, occupation, personal preferences and tastes, personality traits, and styles of communication and interaction with others. For example, in virtually all known cultures women wear their hair longer than men, men tend to pursue more competitive occupations and pastimes while women tend to pursue more cooperative activities, and women tend to do more listening than men in interpersonal communication. Thus, the typical differences between men and women that Americans recognize are also well-represented in other cultures.

Dr. Pretchik also notes that some societies do recognize genders beyond the man-woman binary, saying "Without question there are societies that have concepts of a 'third gender.' However, the man-woman distinction is ubiquitous, and there is no denying that the overwhelming majority of people who have existed on this Earth have lived and died as either men or women. That is the anthropological and historical fact, and I don't see that changing anytime soon."

Low essentialism article:

THE TRUTH ABOUT THE GENDER BINARY: VARIATION IS THE ONLY CONSTANT

OXFORD, UK – As people increasingly demand recognition for transgender and other non-binary gender identities (e.g., genderqueer, agender, gender nonconforming, etc.), we are confronting many important new questions, both legal and philosophical. Perhaps the most basic question that has been raised is whether the traditional man-woman gender binary has any firm basis in reality, or whether it is instead a mere social convention to categorize people in this manner.

This, like all fundamental questions of human nature, is a difficult question to answer. However, this is exactly the kind of question that anthropologists are trained to answer. Dr. Hazel Pretchik of Oxford University has devoted her entire career to studying the anthropology of gender. Dr. Pretchik has traveled the world observing and documenting how hundreds of different cultures approach gender, from small-scale hunter-gatherer communities in Africa and the Pacific Islands to fully industrialized societies in North America, Asia, and Europe. Based on her extensive experience, Dr. Pretchik felt that she could give a confident answer to the fundamental question: Is the gender binary real?

"Of course it's not real!" says Dr. Pretchik. "In every society I have studied, they define gender in a unique and highly idiosyncratic way. I have never seen or heard about two cultures that deal with gender in the exact same way." This

means that the gender binary Americans are familiar with is not what anthropologists call a cultural universal, a feature that is present in all known human societies. If it can be demonstrated that a certain behavior or practice is a cultural universal, this is strong evidence that it arises from our basic human nature. For example, language and music are found in every known human society, and we conclude from this that being speakers and music-makers are an inherent part of what it means to be human. Dr. Pretchik argues that this is not so for the man-woman binary.

"Male and female gender concepts don't even exist in every society I know of," Dr. Pretchik continued, "And even among societies that do distinguish between men and women there is a remarkable degree of diversity in the specific content of what it means to be male or female across cultures." According to Dr. Pretchik's research, there is tremendous cross-cultural variability in gender differences in appearance and dress, occupation, personal preferences and tastes, personality traits, and styles of communication and interaction with others. For example, there are cultures where women wear their hair shorter than men, cultures in which men tend to pursue more cooperative occupations and pastimes while women tend to pursue more competitive activities, and cultures where men tend to do more listening than women in interpersonal communication. Thus, the typical differences between men and women that Americans recognize may be completely reversed in other cultures.

Dr. Pretchik also notes that some societies recognize genders beyond the man-woman binary, saying "Without question there are societies that have concepts of a 'third gender.' This, along with the other evidence I have been discussing, clearly shows that the man-woman gender binary and the associated beliefs about each gender that are common in America are just one way among an infinite variety of ways that a society can deal with gender. That is the anthropological and historical fact, and I wouldn't be surprised at all if America has a very different approach to gender 100 years from now."

SI.3.3 Pilot Study to Study 3

We initially conducted a version of Study 3 (see main text) with a preregistered sample of $N = 300$. We deviated from the preregistration and doubled the sample size to $N = 600$. Because this was a major deviation from the preregistration, we are reporting this initial study here as a pilot and conducted a new version of it that strictly adhered to its preregistered recruitment and analysis plan (see Study 3 in the main text).

Method

Participants

We initially planned to collect data from 300 participants but doubled the sample size to increase power to detect small effects. This was a deviation from the preregistration. We included "Phase" (initial sample of 300 vs. additional sample of 300) as a random effect in all statistical models. The final sample was thus 600 English-speaking participants: all U.S. adults; $M_{age} = 37.69$; 55% women, 44% men, 1% gender non-binary, and < 1% unspecified; and 69% self-identified as "White" or "Caucasian" and the remaining 31% self-identified as a range of other race and ethnicities (e.g., "Latina," "indigenous," and "mixed"). Prior to exclusions, 693 participants completed the survey, and 93 were excluded according to preregistered criteria as detailed below. A sensitivity power analysis revealed that this sample was sufficiently powered to detect a small-to-medium effect of condition on pronoun endorsement ($d = 0.23$ on a two-tailed independent-samples t test, power = 80%, $\alpha = .05$; Cohen, 1992; Faul et al., 2007).

Procedure and Materials

Methods were generally identical to Study 3 reported in the main text. Exceptions are detailed below.

In terms of exclusions, two coders blind to condition read participants' open-ended responses about the purpose of the studies (elicited at the end of each survey) and identified participants who expressed suspicion about the cover story ($n = 60$) or provided nonsense responses ($n = 2$). Coder agreement was strong ($\kappa = .84$), and discrepancies were resolved via discussion. According to the preregistered criteria, these participants were excluded alongside participants who failed additional preregistered attention checks ($n = 31$).

In the first survey, participants read an article designed to activate either high gender essentialist ideas or low gender essentialist ideas. Participants also completed the same gender essentialism items as in Study 3 ($\alpha = .90$).

In the second survey, participants again read information about personal pronouns and completed a comprehension check question. In Studies 1, 2a, and 2b this comprehension check question was used as an initial screening criterion, and respondents who failed this check were immediately redirected to the end of the survey without completing any measures. In Study 3 reported in the main text, participants who answered incorrectly were excluded from analyses. However, in the present pilot study, we retained participants even if they failed this check. Instead, we planned to conduct a robustness analysis where we included responses to this question (correct vs. incorrect) as an additional factor in the models. Next, using the same statements from Study 1, participants reported their endorsement of binary gender ($\alpha = .85$), race ($\alpha = .90$), and identity-neutral ($\alpha = 0.84$) pronouns.

At the end, participants were thanked for their participation, debriefed about the full purpose of the study, and compensated $1.41. Median completion time was 16 minutes.

Results and Discussion

The analytic approach was similar to Study 3 reported in the main text. Exceptions are highlighted below.

Manipulation Check

The essentialism manipulation worked as intended. We computed a mixed effects model with gender essentialism predicted by essentialism condition with a random intercept for phase. Gender essentialism was higher in the high essentialism condition ($M = 6.24$, $SD = 1.61$) compared to the low essentialism condition ($M = 5.30$, $SD = 1.83$), $B = 0.94$, $SE = 0.14$, $p < .001$, $\beta = 0.52$.

Relationships Between Gender Essentialism and Pronoun Endorsement

Individuals who were higher in gender essentialism showed stronger endorsement of binary gender pronouns and race pronouns and weaker endorsement of identity-neutral pronouns. We computed a mixed effects model with endorsement of pronouns predicted by gender essentialism, pronoun type, and their interaction terms. This model also included condition as a control variable as well as a random intercept for participant nested within phase. This model exhibited singular fit due to the random intercepts explaining very little variance, but estimates were identical in an ordinary least-squares regression. For statistics, see Table S8.

Experimental Effect of Condition on Pronoun Endorsement

There was weak evidence for an experimental effect of the essentialism articles on endorsement of pronouns. We computed a linear mixed-effects model with condition, pronoun type, and their interaction terms predicting pronoun endorsement, with random intercepts for participant and participant nested within phase. We compared this model to an identical model without the two interaction terms to determine whether the overall condition × pronoun type interaction was significant. There was a significant difference in fit between these two models, $\chi^2 (2) = 7.95$, $p = .019$, suggesting that the condition × pronoun type interaction was significant.

Examining the simple slopes, participants in the high-essentialism condition showed non-significantly greater endorsement of binary gender pronouns compared to those in the low-essentialism condition, $B = 0.17$, $SE = 0.12$, $p = .155$, $\beta = 0.10$, as well as significantly greater endorsement of race pronouns, $B = 0.40$, $SE = 0.12$, $p = .001$, $\beta = 0.22$, and non-significantly lower endorsement of identity-neutral pronouns, $B = -0.09$, $SE = 0.12$, $p = .473$, $\beta = -0.05$. (These are the results that we meta-analyzed in the main text with those of Study 3.) The effect on race pronouns is notable and consistent with prior work that has found that manipulating gender essentialism sometimes "spills over" to increase essentialism in other domains besides gender (Christy et al., 2019).

Following the preregistered analysis plan, we tested and found that performance on the comprehension check significantly moderated these results; that is, adding the condition × pronoun type × comprehension check (pass, fail) interaction terms to the model explained significantly more variance, $\chi^2 (2) = 30.22$, $p < .001$. The predicted experimental results were stronger among participants who passed the comprehension check question ($n = 541$; 90.2% of the total sample). That is, among these participants, those in the high (vs. low) essentialism condition endorsed binary gender pronouns significantly more, $B = 0.25$, $SE = 0.13$, $p = .046$, $\beta = 0.14$, race pronouns significantly more, $B = 0.39$, $SE = 0.13$, $p = .002$, $\beta = 0.22$, and identity-neutral pronouns non-significantly less, $B = -0.11$, $SE = 0.13$, $p = .369$, $\beta = -0.06$ (Figure S3).

Mediation of Condition Effect on Pronoun Endorsement

In the preregistered multilevel moderated mediation model, we found that the manipulation impacted endorsement of pronouns through essentialism (Figure S4). We found that the indirect effect of the manipulation on pronoun endorsement via gender essentialism differed by pronoun type. The indirect effect differed for binary gender pronouns compared to identity-neutral pronouns, $B = 0.50$, $SE = .09$, $p < .001$, and for race pronouns compared to identity-neutral pronouns, $B = 0.49$, $SE = .09$, $p < .001$. For both binary gender and race pronouns, the indirect effects of the manipulation were positive: Participants in the high (vs. low) essentialism condition showed greater gender essentialism, which in turn related to greater endorsement of binary gender and race pronouns (*ab*s for gender pronouns = 0.25, $p < .001$, and race pronouns = 0.24, $p < .001$). In contrast, the indirect effect of the manipulation was negative for identity-neutral pronouns: Participants in the high (vs. low) essentialism condition showed greater gender essentialism, which in turn related to lower endorsement of identity-neutral pronouns ($ab = -0.25$, $p < .001$).

Table S8

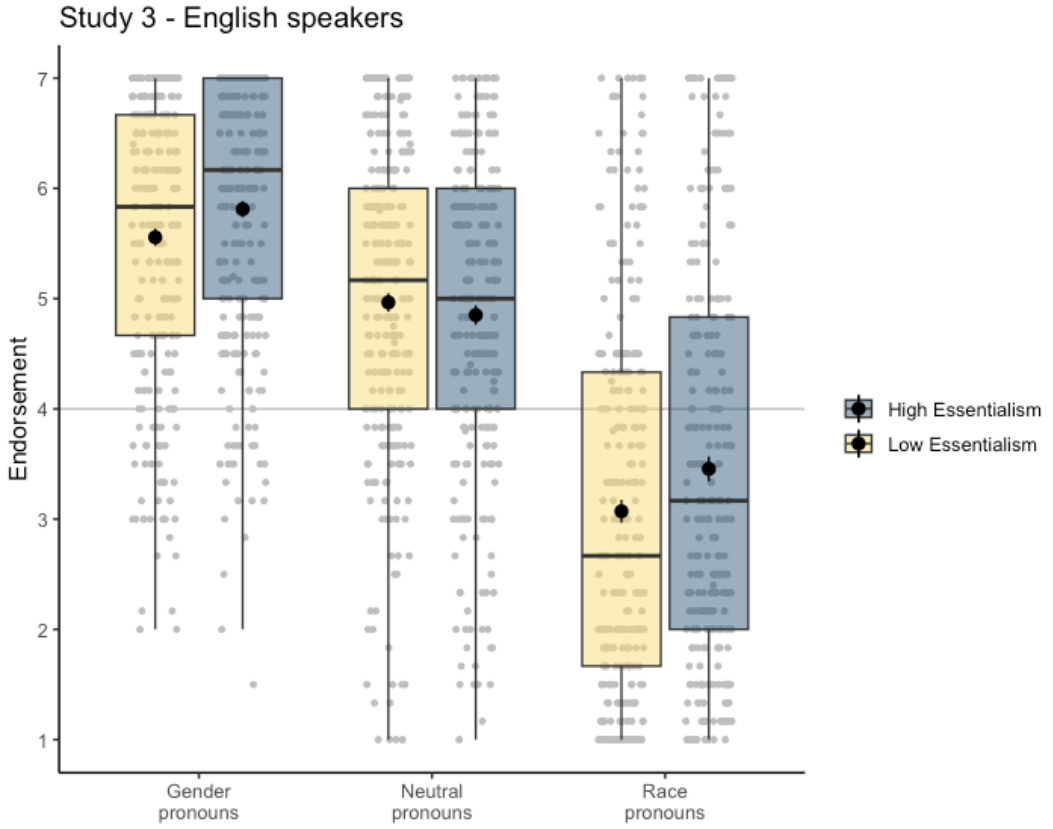
Relationship Between Essentialism and Endorsement of Three Pronoun Types

Term			
	χ^2 (*df*)	*B* (*SE*)	β
GE × Pronoun Type	175.64 (2) ***		
GE × Neutral (= 1) vs. Gender (= 0) Pronouns		−0.54 (0.04) ***	−0.54
GE × Race (= 1) vs. Gender (= 0) Pronouns		−0.01 (0.04)	−0.01
GE × Neutral (= 1) vs. Race (= 0) Pronouns		−0.53 (0.04) ***	−0.53
Simple slope of GE for Gender Pronouns		0.27 (0.03) ***	0.27
Simple slope of GE for Race Pronouns		0.26 (0.03) ***	0.26
Simple slope of GE for Neutral Pronouns		−0.27 (0.03) ***	−0.27

Note. GE = Gender essentialism. The beta coefficient (β) is a measure of effect size derived from a model with standardized GE and standardized pronoun endorsement. Note that the unstandardized and standardized coefficients are identical because GE and pronoun endorsement have similar standard deviations.
*$p < .05$. **$p < .01$. ***$p < .001$.

Figure S3

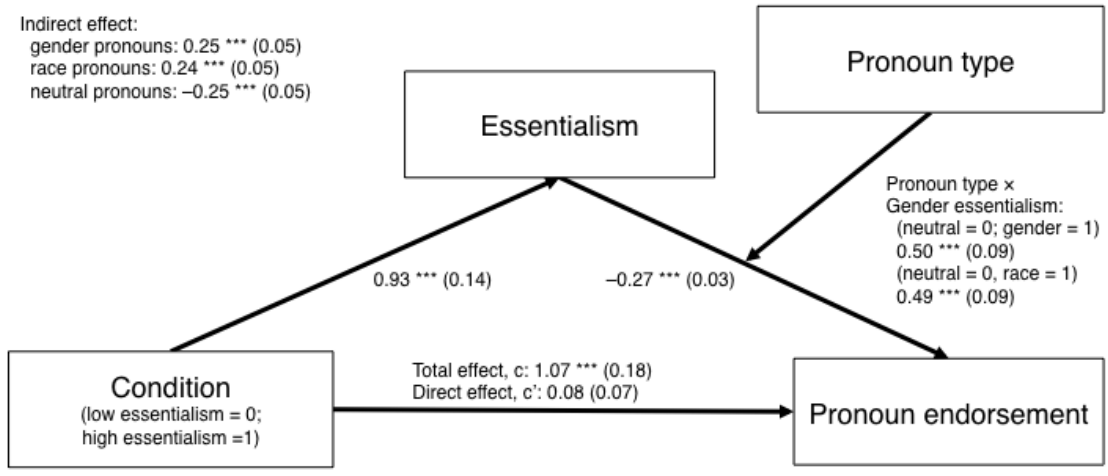
Experimental Effect of Condition on Pronoun Endorsement in the Study 3 Pilot



Note. Results are shown for the subset of participants (n = 541 out of 600) who passed the initial comprehension check question. Boxplots with means (black dots), medians (horizontal lines in each of the boxplots), and standard error bars. Raw data are underlaid (gray dots).

Figure S4

Multilevel Moderated Mediation of the Effect of Condition on Pronoun Endorsement Through Gender Essentialism in the Study 3 Pilot



Note. Unstandardized coefficients are reported with *SE*s in parentheses. Gender essentialism is mean centered, and the essentialism condition and pronoun type are dummy coded as specified. *** $p < .001$.