Title

When Your Kind Cannot Live Here: How Generic Language and Criminal Sanctions Shape Social Categorization

Journal

Psychological Science, 28(11)

ISSN

0956-7976

Authors

Goldfarb, Deborah
Lagattuta, Kristin Hansen
Kramer, Hannah J
et al.

Publication Date

2017-11-01

DOI

10.1177/0956797617714827

Peer reviewed

Research Article

Psychological Science
2017, Vol. 28(11) 1597–1609

DOI: 10.1177/0956797617714827
www.psychologicalscience.org/PS

When Your Kind Cannot Live Here: How Generic Language and Criminal Sanctions Shape Social Categorization

Deborah Goldfarb, Kristin Hansen Lagattuta, Hannah J. Kramer, Katie Kennedy, and Sarah M. Tashjian
Department of Psychology and Center for Mind and Brain, University of California, Davis

Abstract

Using generic language to describe groups (applying characteristics to entire categories) is ubiquitous and affects how children and adults categorize other people. Five-year-olds, 8-year-olds, and adults (N = 190) learned about a novel social group that separated into two factions (citizens and noncitizens). Noncitizens were described in either generic or specific language. Later, the children and adults categorized individuals in two contexts: criminal (individuals labeled as noncitizens faced jail and deportation) and noncriminal (labeling had no consequences). Language genericity influenced decision making. Participants in the specific-language condition, but not those in the generic-language condition, reduced the rate at which they identified potential noncitizens when their judgments resulted in criminal penalties compared with when their judgments had no consequences. In addition, learning about noncitizens in specific language (vs. generic language) increased the amount of matching evidence participants needed to identify potential noncitizens (preponderance standard) and decreased participants' certainty in their judgments. Thus, generic language encourages children and adults to categorize individuals using a lower evidentiary standard regardless of negative consequences for presumed social-group membership.

Keywords

social cognition, language, decision making, legal processes, social structure

Received 8/25/16; Revision accepted 5/22/17

As 20 million refugees flee war, famine, and hate crimes, does willingness to deliver punitive sanctions for improper border crossings depend on whether people conceptualize noncitizens as a homogeneous generic group rather than as specific individuals? In the midst of the current global refugee crisis, governments have called upon their citizens to report suspected illegal immigrants (Aliverti, 2015), and this has resulted in searches, harassment, and threats of jail and deportation solely because individuals exhibit traits (e.g., physical or behavioral attributes) believed to be characteristic of noncitizens (Holley, 2016; K. Lyons, 2016; Potok, 2016). Children's and adults' perceptions of who is a noncitizen may relate to how they learn about, and thus conceptualize, that social group. The current research analyzed how two features—language (generic vs. specific) and context (criminal vs. noncriminal)—influence how children and adults categorize individuals on the basis of varying amounts of evidence.

Referencing social-group members in generic language (e.g., "males are strong") as opposed to specific language (e.g., "Jacob is strong") carries substantial weight (Cimpian & Markman, 2011; Rhodes, Leslie, & Tworek, 2012). When generic, rather than specific, language is used to define a characteristic of a social or a nonsocial category, children and adults more often assume that the characteristic is an inherent, stable, and functional feature of all group members (Cimpian & Markman, 2011; Gelman, Ware, & Kleinberg, 2010; Rhodes et al., 2012). Moreover, when

Corresponding Author:
Deborah Goldfarb, University of California, Davis, Department of Psychology, One Shields Ave., Davis, CA 95616
E-mail: dagoldfarb@ucdavis.edu

category members are described as having a particular attribute in generic, rather than specific, language, children and adults are more likely to conclude that individuals with that feature belong to that group (Gelman et al., 2010; Graham, Nayer, & Gelman, 2011). The influence of language genericity is not equally robust across age, however. Compared with older individuals, 3- to 5-year-olds require more exposure to generic input to exhibit the effect, and preschoolers sometimes respond equivalently to generic and specific statements, revealing that they do not always distinguish between them (Cimpian & Scott, 2012; Gelman et al., 2010).

Although societies sometimes sanction individuals strictly on the basis of social-group membership (e.g., governments blocking residency), it remains an open empirical question whether the influence of language genericity on categorization varies as a function of whether the categorization will lead to such consequences. Simply marking category boundaries (e.g., associating different groups with uniquely colored T-shirts) elicits in-group preferences and increases children's and adults' willingness to reject or withhold resources from outsiders (Callahan & Ledgerwood, 2016; Diesendruck & Menahem, 2015; Killen, Mulvey, & Hitti, 2013; Rhodes, Leslie, Saunders, Dunham, & Cimpian, 2017; Tajfel, 1982). Still, recent studies suggest that there are limits to the effects of language genericity when categorization invokes negative outcomes. Learning about a novel out-group in generic, rather than specific, language does not increase 4- to 6-year-olds' negative attitudes toward the out-group (Rhodes et al., 2017). Also, children and adults less often endorse threatening generic statements when the statements concern humans than when they concern objects (Tasimi, Gelman, Cimpian, & Knobe, 2016). The current study tested (a) whether 5-year-olds, 8-year-olds, and adults would identify fewer individuals as noncitizens when there were criminal penalties for those identified than when there were no consequences, and (b) if the genericity of the language in which participants learned about noncitizens would moderate any such effect.

Researchers typically assess the influence of generic (vs. specific) language on categorization by presenting exemplars that either do or do not exhibit a single feature (e.g., striped feet; Gelman et al., 2010). Everyday social categorization, however, often requires evaluating multiple pieces of evidence, some of which are consistent and some of which are inconsistent with learned group characteristics. We expanded theory and methods in this line of research by testing whether language genericity alters the evidentiary standard needed to declare an individual a social-group member (in this case, a noncitizen). Participants classified a series of novel creatures as citizens or noncitizens after viewing three pieces of evidence about each individual; the number of characteristics that matched what participants had previously heard about noncitizens varied from zero to three.

We further investigated whether evidentiary standards vary by context and age. Children and adults may have a more stringent threshold for declaring individuals as noncitizens when those identified face criminal penalties. Previous research suggests alternative possibilities for age-related changes in the evidentiary threshold. Compared with children, adults assume more coherent category structures (Keil & Batterman, 1984; Murphy & Medin, 1985), which would potentially decrease the evidence they need to label individuals as noncitizens. In contrast, because 4- to 5-year-olds have difficulty recognizing indeterminacy (i.e., they tend to treat a single piece of positive evidence as sufficient; Fay & Klahr, 1996), they may require less evidence about group membership than do older individuals.

Finally, we measured participants' certainty and decision speed for all categorization selections. Judgment and decision-making researchers commonly use these metrics as indicators of decision confidence (Kiani, Corthell, & Shadlen, 2014). Extending these methods to our investigation of social categorization allowed us to go beyond dichotomous judgments. That is, we tested not only which variables (language, context, amount of evidence, and age) influenced participants' identification of category members, but also how these variables influenced how sure they felt about their choices and the amount of time they took to deliberate.

Central Hypotheses

We predicted that participants who learned about a novel group of noncitizens via generic rather than specific language would (a) identify more individuals as potential noncitizens, (b) require a lower evidentiary standard for such identifications, (c) report more certainty in their decisions, and (d) categorize individuals more quickly. Regarding developmental patterns, we hypothesized that, as found in prior studies, the effects of language genericity would strengthen with age and that evidentiary standards for labeling individuals as noncitizens would either increase or decrease with age. We predicted that the presence of criminal penalties for presumed noncitizens would (a) reduce the identification of potential noncitizens and (b) increase the evidentiary standard for such identification. Finally, we explored whether language genericity modified any effects of criminal context.

Method

Participants

Participants (N = 190) included sixty-six 5-year-olds (M = 5.57 years, SD = 0.76; 32 females, 34 males), sixty-two 8-year-olds (M = 8.40 years, SD = 0.84; 33 females,

29 males), and 62 adults (M = 22.16 years, SD = 6.99; 33 females, 29 males). This sample size was determined prior to data collection, and it is in line with other studies that have detected moderate effects of language and social-group membership in similar samples (Diesendruck, Salzer, Kushnir, & Xu, 2015; Gelman et al., 2010). We chose these age groups because they typically reveal differences in effects of generic language, social categorization decisions, and endorsement of punishment (Brown, 2011; Gelman et al., 2010; Killen et al., 2013). We discontinued recruitment at our predetermined stopping point.

The sample was composed of 58% non-Hispanic Caucasians (5-year-olds: n = 46; 8-year-olds: n = 45; adults: n = 19), 15% Asians (5-year-olds: n = 3; 8-year-olds: n = 2; adults: n = 23), 10% Hispanics (5-year-olds: n = 5; 8-year-olds: n = 3; adults: n = 11), 2% African Americans (5-year-olds: n = 1; adults: n = 2), and 15% participants of "other" races or ethnicities (5-year-olds: n = 11; 8-year-olds: n = 11; adults: n = 7). One 8-year-old participant's ethnicity was not reported. Most (94%) 5- and 8-year-olds' parents had a college degree. The parents of the 5- and 8-year-olds earned an average of \$80,000 to \$90,000 a year; the adults reported that their parents earned an average of \$60,000 to \$70,000 a year.

All participants were typically developing. Nine were excluded from analyses: 2 because they failed to comply with the study's procedures (5-year-olds: n = 1; 8-year-olds: n = 1) and 7 because the experimenter failed to ask all of the comprehension-check questions (8-year-olds: n = 2; adults: n = 5).

The children were recruited from a list of previous participants, at local farmers' markets and schools, and through fliers, listserv e-mails, and referrals from other participants. The adults were recruited through undergraduate courses at a large research university.

Materials and procedure

The study took place in a single research session (45 min) in a quiet room. A female experimenter administered the tasks in the order in which they are described here. All age groups performed the tasks on a laptop with a customized keyboard that allowed participants to affix study-specific pictures to response keys. We used DirectRT (Jarvis, 2012) to record participants' decisions and their reaction times (in milliseconds). The children received \$10.00 for participating, and the adults received course credit. (Participants completed additional measures that were not part of this project, and results for those measures will be reported elsewhere.)

Training. Participants completed a series of training exercises to ensure that they understood the response scales. First, the experimenter taught participants how to use the categorization keys. Four animal images appeared on the screen, and participants pressed the keyboard picture indicating where each animal lives: land (dog, lion) or water (dolphin, octopus). Participants achieved 100% accuracy before progressing.

Next, the experimenter instructed participants on how to utilize a 4-point pictorial certainty scale (also on the keyboard) composed of green circles increasing from small ("not at all sure you are right") to large ("very sure you are right"). After learning what each circle represented, participants saw four images (i.e., car, chair, hat, and mug) ranging in pixelation from clear to distorted. Participants identified each image and rated their decision certainty by pressing the associated keyboard picture. Following this practice, participants also had to point correctly to the circle that represented each level of certainty (asked in random order) before they continued with the study.

Background narrative. The background narrative for the experimental tasks (see Fig. 1)[1] was presented on the laptop computer. Participants learned that a society of creatures called Twiggums lived in a valley on a far-away island and that, one day, some of the Twiggums moved to the mountains to start a new community (complete with a unique flag and home colors; Callahan & Ledgerwood, 2016) and renamed themselves Zuttles. Critically, Twiggums (who remained citizens of the valley) and Zuttles (noncitizens) were physically indistinguishable and originated from the same group. The only differentiating information about the two groups that the narrator mentioned was their label (Twiggums vs. Zuttles) and their status in the valley (Twiggums = citizens; Zuttles = noncitizens). Zuttles and Twiggums had anthropomorphic features (i.e., upright posture, two eyes, one nose, one mouth), and all varied slightly from each other in body weight, height, and shade of orange. We used novel creatures to reduce the possibility that preconceived notions or experiences would affect participants' judgments (Killen, 2007; Rhodes et al., 2012). We did not provide clothing, age, socioeconomic-status, or gender cues to remove the possibility that participants would interpret this information as indicative of group membership (Dunham, Baron, & Carey, 2011).

Next, a narrator (a prerecorded female voice played through the laptop) told participants about Zuttles in either generic or specific language (see Fig. 2). Participants saw 12 Zuttles, each paired with a psychological characteristic (liking or hating a food) or a behavioral characteristic (being good or bad at an activity). The narrator described the characteristic in either *generic* language (e.g., "Zuttles like apples") or *specific* language (e.g., "This Zuttle, Dax, likes apples"). The attribute was crossed out in red when Zuttles hated the food or were bad at the activity. All characteristics presented were

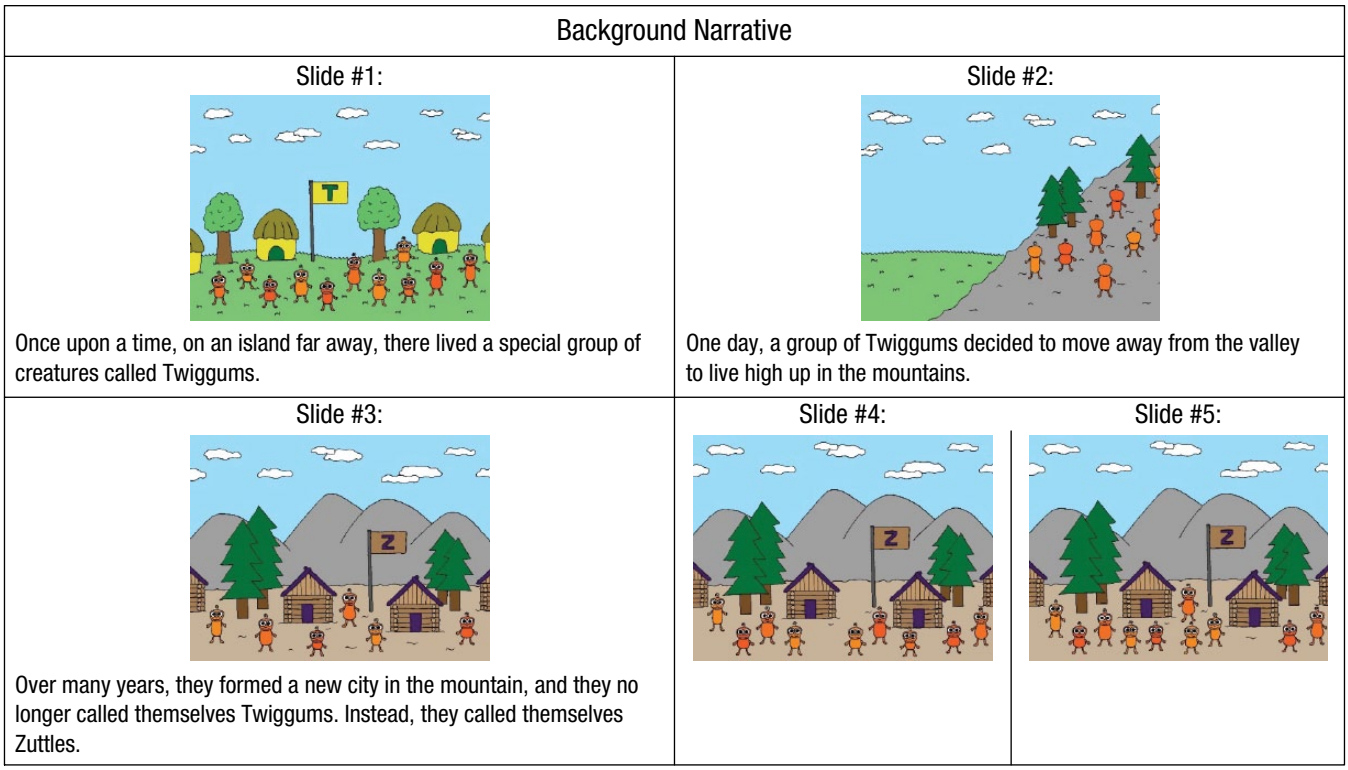


Fig. 1. Slides presented to all participants at the start of the study to introduce Zuttles and Twiggums. The text beneath each image corresponds to the audio that played while that slide was displayed.

ones that could be shared widely (e.g., liking apples) and were not deviant or harmful.

Given the variety of the evidence presented and its nonexclusivity (e.g., participants were not told that "only" Zuttles liked the food), there was no right answer as to whether an individual was or was not a Zuttle.

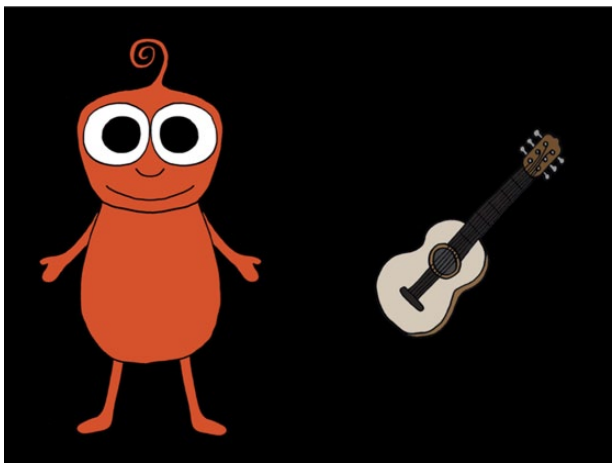


Fig. 2. Sample slide showing a Zuttle characteristic. The voice-over accompanying this slide was "Zuttles are good at guitar" in the generic-language condition and "This Zuttle, Zonti, is good at guitar" in the specific-language condition.

This design mirrors real life, in which there are no correct answers when classifying individuals into social constructs: Stereotypes are often based on preferences or abilities neither unique to a social group nor inherently negative. Moreover, even if a person's characteristics match what is believed to be true about a group, this does not guarantee that he or she is actually a member of that group (e.g., an adult who is good at sports, likes action movies, and serves in the military could be male or female). The four types of characteristics (i.e., liked foods, hated foods, activities Zuttles were good at, activities Zuttles were bad at) were presented in separate blocks, but the block order was randomized. Language condition was manipulated between participants (via random assignment) to ensure that characteristics were equally weighted across conditions (Sutherland & Cimpian, 2015) and because generic language can contaminate later blocks with specific language (Cimpian & Scott, 2012).

Categorization task. After the background information was presented, participants completed the categorization task. In the *noncriminal context*, participants were told that we wanted help deciding whether creatures were Zuttles or Twiggums. The instructions reinforced the point that this categorization had no consequences.

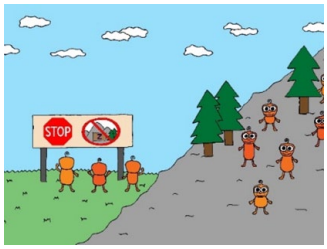
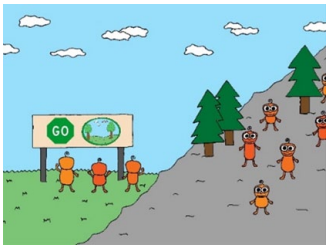
Manipulation of Context	
Slide Introducing the Criminal Context	Slide Introducing the Noncriminal Context
	
If the criminal context was first: Well, after time had passed, some of the Zuttles decided that they wanted to move back to the valley. But there was a problem with this. The Twiggums, you see, had passed a law that said, "No Zuttles are allowed to be in the valley." In fact, the Twiggums put signs all over the valley showing this rule. Twiggum police recently found out that some Zuttles broke this rule and live in the valley. The police have several suspects who they think might be Zuttles. We need your help finding the Zuttles. If you decide the creature is a Zuttle, that one will be sent to jail for 100 days and then forced back to the mountains. Remember, Twiggums and Zuttles look exactly alike. There is no way to tell if creatures are Twiggums or Zuttles just by looking at them. **If the criminal context was second:** Now imagine that the Twiggums passed a new law that said, "No Zuttles are allowed in the valley." In fact, the Twiggums put signs all over the valley showing this rule. Twiggum police recently found out that some Zuttles broke this rule and live in the valley. The police have several suspects who they think might be Zuttles. We need your help finding the Zuttles. If you decide the creature is a Zuttle, that one will be sent to jail for 100 days and then forced back to the mountains. Remember, Twiggums and Zuttles look exactly alike. There is no way to tell if creatures are Twiggums or Zuttles just by looking at them.	**If the noncriminal context was first:** Well, after time had passed, some of the Zuttles decided that they wanted to move back to the valley. We need your help finding the Zuttles. If you decide the creature is a Zuttle, nothing happens. We just need some help making this decision. Remember, Twiggums and Zuttles look exactly alike. There is no way to tell if creatures are Twiggums or Zuttles just by looking at them. **If the noncriminal context was second:** Now imagine that the Twiggums decided to get rid of the law. All creatures are allowed to live in the valley. We need your help finding the Zuttles. If you decide the creature is a Zuttle, nothing happens. We just need some help making this decision. Remember, Twiggums and Zuttles look exactly alike. There is no way to tell if creatures are Twiggums or Zuttles just by looking at them.

Fig. 3. Slides presented to participants before each context of the categorization task: criminal context (left) and noncriminal context (right). The text beneath each image corresponds to the audio that played while that slide was displayed; for each context, the audio varied depending on whether that context was presented first or second.

In the *criminal context*, participants learned that the Twiggums made laws forbidding Zuttles from entering the valley and the police needed help deciding which creatures were noncitizens (Zuttles). Research shows that illegal border crossings are considered wrong by the majority (82%) of 5- to 10-year-olds (Brown, 2011); thus, children as young as 5 should recognize this as a crime (we also verified this understanding in our sample by asking control questions, discussed later in this section). The seriousness of entering the valley illegally was conveyed by stating that all participant-identified Zuttles (noncitizens) would be jailed for 100 days and then forced back to the mountains (deported). The order of the two contexts was counterbalanced across participants. During the switch between contexts, participants learned that the law was either enacted or withdrawn, such that border crossing had become criminal or noncriminal, respectively (see Fig. 3).

For both the criminal and the noncriminal contexts, participants completed 14 test trials, each showing an image of a new creature paired with a unique trio of characteristics. Across trials, the number of exact matches to what participants had previously learned about Zuttles (zero, one, two, or three exact matches) was varied (see Fig. 4). Each creature was presented individually, with the three characteristics displayed above it. The narrator named the three characteristics (e.g., "This one is good at piano, . . ."). During the task, participants could refer to a memory cue sheet with the 12 previously learned characteristics (see Fig. 4); this

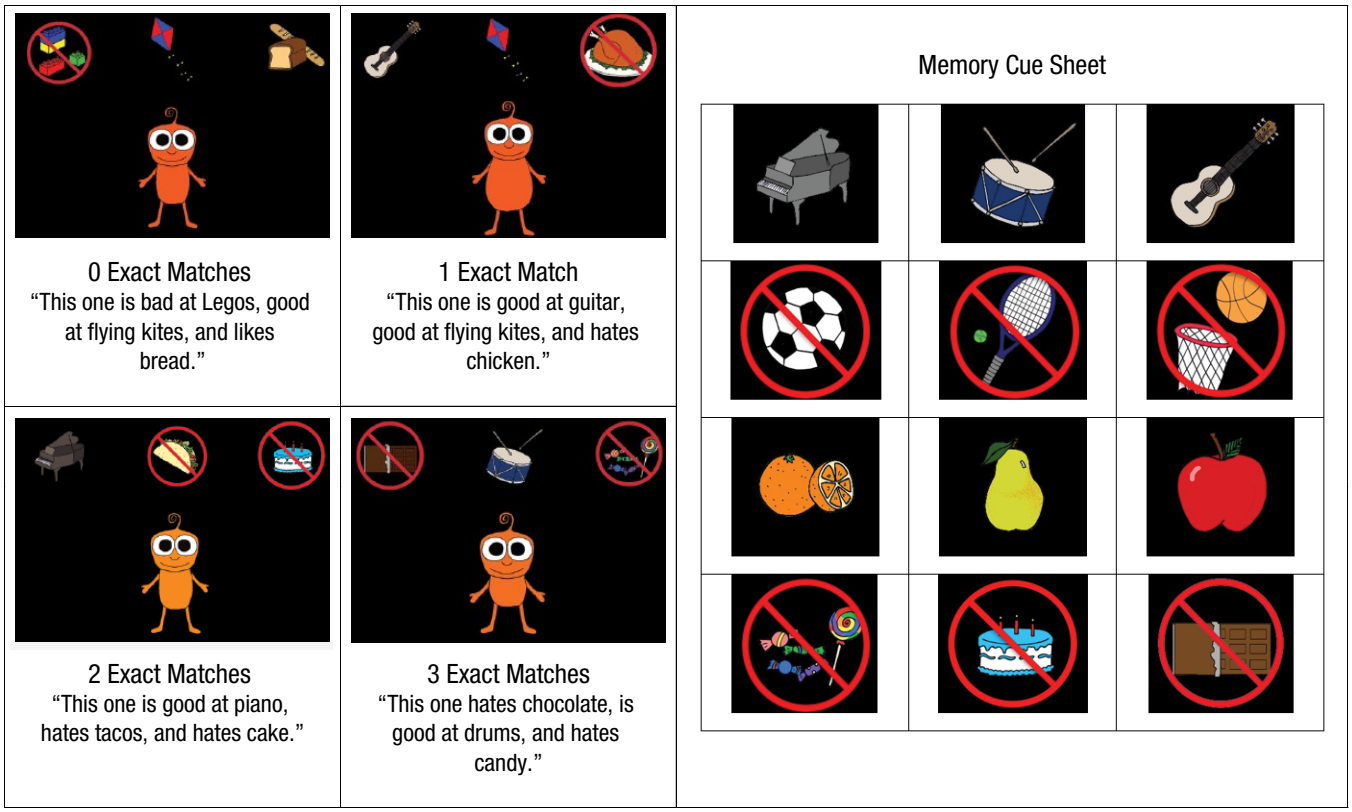


Fig. 4. Illustration of the materials used for testing. The slides on the left exemplify the manipulation of the amount of evidence presented to participants during the categorization task (zero exact matches, one exact match, two exact matches, or three exact matches). Throughout testing, a clearly visible memory cue sheet (right) pictorially presented all the information that participants had previously been taught about Zuttles.

eliminated memory demands and ensured that any effects of language condition were not the result of increased recall for the presented information in the generic-language condition (Gülgöz & Gelman, 2015). Because there were no predetermined Zuttles (noncitizens), all or none of the creatures shown in the test trials could be Zuttles, and chance identification was 50%.

On each trial, participants pressed a picture on the keyboard to indicate whether the creature was a Zuttle (noncitizen; i.e., picture of a *Z* in the noncriminal context, picture of a jail in the criminal context) or a Twiggum (citizen; i.e., picture of a *T* in the noncriminal context, picture of a valley in the criminal context; see Fig. 5). The left/right position of the pictures was counterbalanced across participants. After each decision, participants received a visual and auditory reminder of their choice. In the criminal context, the prerecorded narrator said in a neutral tone, "You sent that one to the jail," while a picture of a jail was displayed, or "You sent that one to the valley," while a picture of a valley was displayed. In the noncriminal context, the narrator said, "You said that one was a Zuttle," while a picture of a *Z* was displayed, or "You said that one was a Twiggum," while a picture of a *T* was displayed. After this reminder, participants rated the certainty of their decision (i.e., "How sure are you of your choice?") using the scale described earlier.

Fig. 5. Image of the keyboard participants used to make categorization decisions (dichotomous choice between two pictorial stickers) and to report their certainty in their choices (4-point circle scale). The pictorial stickers for the noncriminal context are shown here. The *Z* and *T* were replaced with jail and valley pictures, respectively, for the criminal context.

To help solidify the differences between the contexts and engage participants, the experimenter had the participants affix the appropriate response pictures to the keyboard before the first categorization set was presented and again before starting the second categorization set (i.e., after the context switch). The experimenter also verified that the participants understood the procedures and the premises of the contexts, including the legality of crossing the border, by asking a series of control questions (i.e., "Are Zuttles allowed to live in the valley?" "Which button do you press if you think it is a Zuttle?" "Which button do you press if you think it is a Twiggum?"). Participants proceeded to the categorization task once they had responded correctly to all three questions. Finally, the experimenter reminded the participants of the memory cue sheet before presenting each categorization set. These procedures ensured that the participants understood (a) which creatures were citizens (Twiggums) and which were noncitizens (Zuttles), (b) that border crossing was prohibited only in the criminal context, and (c) that creatures believed to be noncitizens would be jailed and deported in the criminal context (the response keys had pictures of a jail and a valley), but would face no consequences in the noncriminal context (the response keys had pictures of a "Z" and a "T"; Fig. 5).

Results

We analyzed the effects of language (generic vs. specific), age, context (criminal vs. noncriminal), and amount of evidence (zero, one, two, or three exact matches to previously learned characteristics) on three dependent variables: proportion of individuals identified as Zuttles (noncitizens), certainty of decisions, and speed of categorization. Language and age were between-subjects variables, and context and amount of evidence were within-subjects variables. For simplicity, we refer to Zuttles as noncitizens and Twiggums as citizens throughout the Results and Discussion sections, regardless of the context (i.e., criminal vs. noncriminal).

As noted previously, there were no predetermined noncitizens or citizens in the categorization tasks. Thus, when we report participants' "identification" of noncitizens, this does not signify "correct" judgments. Rather, these identification rates indicate how many individuals participants *declared* to be noncitizens.

Proportion of individuals identified as noncitizens

We conducted a 2 (language: generic, specific) × 3 (age: 5-year-olds, 8-year-olds, adults) × 2 (context: criminal, noncriminal) × 4 (evidence: zero, one, two, or three exact matches) repeated measures analysis of covariance (ANCOVA) on the proportion of trials in which participants identified creatures as noncitizens; order (criminal context first, noncriminal context first) was included as a covariate. This analysis revealed main effects for language, $F(1, 183) = 13.70$, $p < .001$, $\eta_p^2 = .07$; age, $F(2, 183) = 12.44$, $p < .001$, $\eta_p^2 = .12$; and evidence, $F(3, 549) = 54.47$, $p < .001$, $\eta_p^2 = .23$. These main effects were qualified by a Language × Context interaction, $F(1, 183) = 3.97$, $p = .048$, $\eta_p^2 = .02$; a Language × Evidence interaction, $F(3, 549) = 2.93$, $p = .033$, $\eta_p^2 = .02$; a Language × Age interaction, $F(2, 183) = 3.28$, $p = .040$, $\eta_p^2 = .04$; and an Age × Evidence interaction, $F(6, 549) = 7.19$, $p < .001$, $\eta_p^2 = .07$. Despite our predictions otherwise, there was no overall effect of context and no significant Context × Evidence interaction.

Language × Context interaction. As expected, participants in the generic-language condition categorized more creatures as noncitizens than did those in the specific-language condition, in both the criminal context ($p < .001$) and the noncriminal context ($p = .021$). Figure 6 shows, however, that the effect of language was nearly twice as large in the criminal context as in the noncriminal context. In addition, criminal penalties reduced identification of potential noncitizens only when participants learned about noncitizens as specific individuals ($p = .001$; $p > .250$ in the generic-language condition). Participants in the generic-language condition identified noncitizens at a rate that was above chance in both the noncriminal context, $t(91) = 6.84$, $p < .001$, and the criminal context,

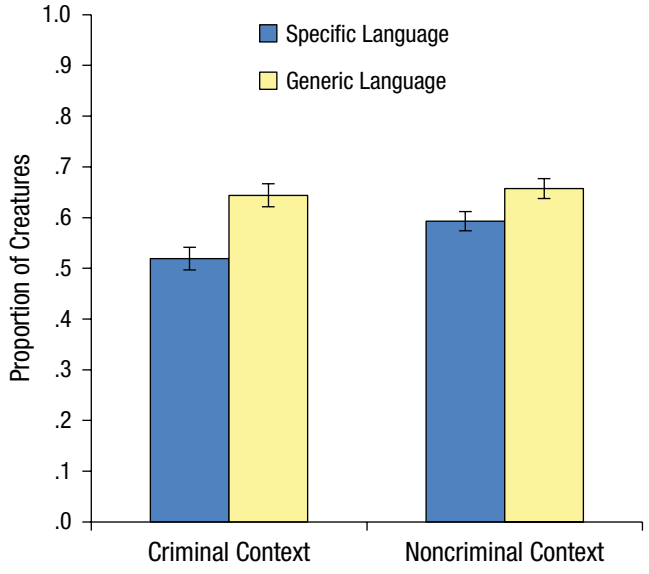


Fig. 6. Proportion of creatures identified as noncitizens as a function of language (generic, specific) and context (criminal, noncriminal). Error bars represent ±1 *SE*.

$t(91) = 6.11$, $p < .001$. In contrast, participants in the specific-language condition identified noncitizens at an above-chance rate only when there were no criminal consequences for noncitizens—noncriminal context: $t(97) = 4.93$, $p < .001$; criminal context: $t(97) = 0.88$, $p > .250$.

Language × Evidence interaction. As anticipated, language genericity also affected the evidentiary standard participants required to categorize individuals as noncitizens (Fig. 7). Compared with participants in the specific-language condition, those in the generic-language condition labeled more creatures as noncitizens when either one or three pieces of evidence matched what they had previously learned (*p*s < .007), and the pattern was in the same numerical direction when there were zero or two exact matches. In both language conditions, participants identified creatures as noncitizens at a rate that was significantly below chance when there were no pieces of matching evidence—specific language: $t(97) = -8.16$, $p < .001$; generic language: $t(91) = -6.59$, $p < .001$. When there was a single piece of matching evidence, only participants in the generic-language condition judged creatures to be noncitizens at a rate that was above chance level—specific language: $t(97) = -0.17$, $p > .250$; generic language: $t(91) = 5.21$, $p < .001$. When the preponderance standard was met (i.e., when there were two or three exact matches), participants in both language conditions identified creatures as noncitizens at an above-chance rate—specific language: $t(97)$s > 6.09, *p*s < .001; generic language: $t(91)$s > 6.94, *p*s < .001.

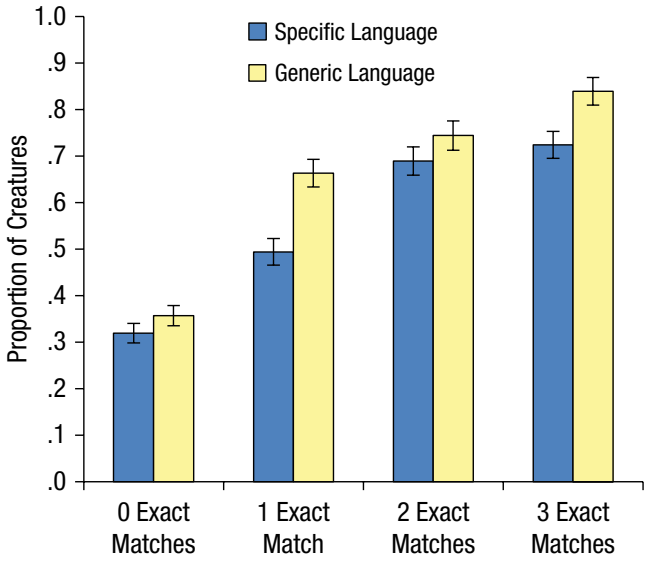


Fig. 7. Proportion of creatures identified as noncitizens as a function of language (generic, specific) and evidence (zero, one, two, or three pieces of matching evidence). Error bars represent ±1 *SE*.

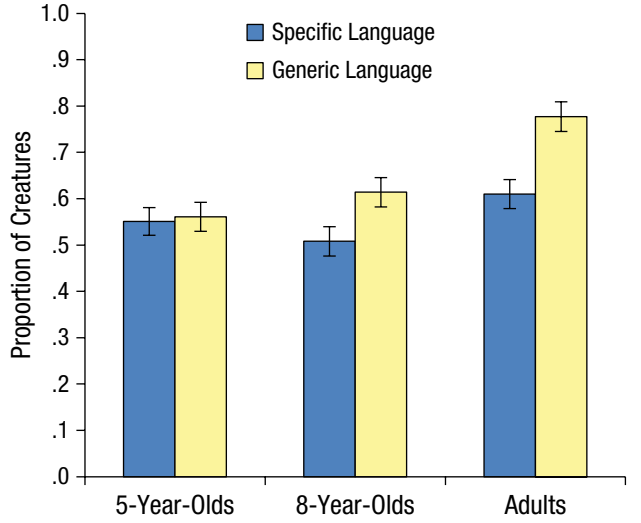


Fig. 8. Proportion of creatures identified as noncitizens as a function of language (generic, specific) and age group (5-year-olds, 8-year-olds, adults). Error bars represent ±1 *SE*.

Language × Age interaction. Previous research has shown that the effects of generic language increase with age, and our data reflected the same pattern. The adults and 8-year-olds (*p*s < .018), but not the 5-year-olds ($p > .250$), identified significantly more individuals as noncitizens in the generic-language condition compared with the specific-language condition (see Fig. 8). The Language × Age interaction does not undermine the Language × Context and Language × Evidence interactions, as those two-way interactions were not qualified by three-way interactions with age (*p*s > .250).

Age × Evidence interaction. All age groups labeled more individuals as noncitizens when there were three exact matches in the evidentiary trio than when there was just one exact match (*p*s < .005), and also identified more individuals as noncitizens when there was one or more exact matches compared with when there were no matches (*p*s < .001). With increasing age, participants required less evidence to categorize creatures as noncitizens (see Fig. 9). That is, when all three pieces of evidence exactly matched what participants had previously learned, the 5-year-olds identified fewer noncitizens than did the 8-year-olds and adults (*p*s < .013). The adults also categorized more creatures as noncitizens than did the children when there was one or two exact matches (*p*s < .002). All three age groups identified creatures as noncitizens at a rate that was above chance when the preponderance of the evidence pointed in that direction—5-year-olds: $t(65)$s > 3.24, *p*s < .002; 8-year-olds: $t(61)$s > 3.97, *p*s < .001; adults: $t(61)$s > 10.32, *p*s < .001. Only the adults labeled creatures as noncitizens at an above-chance rate when there was a single piece of

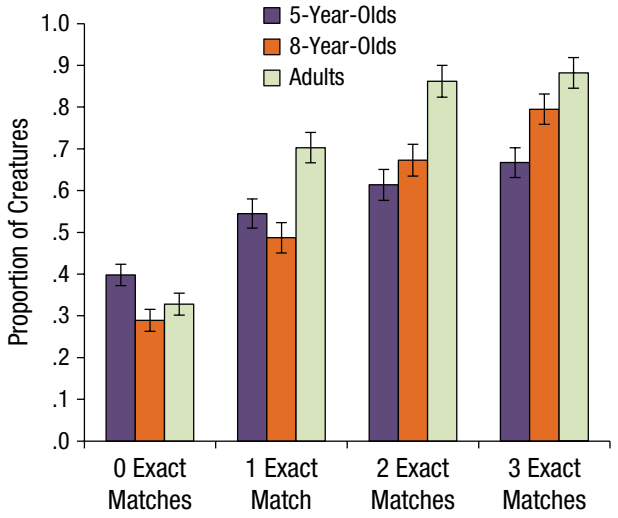


Fig. 9. Proportion of creatures identified as noncitizens as a function of age group (5-year-olds, 8-year-olds, adults) and evidence (zero, one, two, or three pieces of matching evidence). Error bars represent ±1 *SE*.

matching evidence—adults: $t(61) = 5.57$, $p < .001$; 5-year-olds: $t(65) = 1.28$, $p = .204$; 8-year-olds: $t(61) = -0.32$, $p > .250$. All three age groups categorized creatures as noncitizens at a rate that was below chance when none of the evidence matched—5-year-olds: $t(65) = -3.73$, $p < .001$; 8-year-olds: $t(61) = -7.49$, $p < .001$; adults: $t(61) = -7.81$, $p < .001$.

Categorization certainty

A parallel repeated measures ANCOVA on average certainty ratings yielded main effects for language, $F(1, 183) = 4.62$, $p = .033$, $\eta_p^2 = .03$, and evidence, $F(3, 549) = 36.35$, $p < .001$, $\eta_p^2 = .17$. These main effects were qualified by a Language × Evidence interaction, $F(3, 549) = 6.70$, $p < .001$, $\eta_p^2 = .04$, and an Age × Evidence interaction, $F(6, 549) = 13.96$, $p < .001$, $\eta_p^2 = .13$.

Language × Evidence interaction. Results for this interaction were consistent with our prediction. Participants in the generic-language condition were more certain in their decisions than were participants in the specific-language condition when there were three ($p < .001$) or two ($p = .019$) exact matches, but not when there were one or no exact matches (ps > .250; Fig. 10).

Age × Evidence interaction. As illustrated in Figure 11, the adults were less certain than the children on trials with zero or one piece of matching evidence (ps < .011). There were no age-related differences when there were two exact matches (ps > .250). The 5-year-olds expressed less decision certainty when given three pieces of matching evidence than did the 8-year-olds and adults (ps < .031).

Categorization reaction time

A parallel repeated measures ANCOVA on average reaction times (in seconds) yielded main effects for age group, $F(2, 183) = 29.13$, $p < .001$, $\eta_p^2 = .24$; context, $F(1, 183) = 11.61$, $p = .001$, $\eta_p^2 = .06$; and evidence, $F(3, 549) = 8.32$, $p < .001$, $\eta_p^2 = .04$. These main effects were qualified by a Language × Evidence interaction, $F(3, 549) = 3.60$, $p = .013$, $\eta_p^2 = .02$, and a Context × Block Order interaction, $F(1, 183) = 46.41$, $p < .001$, $\eta_p^2 = .20$.[2] With increasing age, participants categorized more quickly (5-year-olds: $M = 11.17$ s, 95% confidence interval, CI = [10.47, 11.87]; 8-year-olds: $M = 9.09$ s, 95% CI = [8.37, 9.81]; adults: $M = 7.29$ s, 95% CI = [6.57, 8.01]; ps < .001). As illustrated in Figure 12, contrary to our predictions, participants in the generic-language condition did not categorize the creatures faster than did participants in the specific-language condition at any level of evidence (ps > .111). Although participants were generally faster to categorize as the amount of matching evidence increased, this linear change was more pronounced in the generic-language condition than in the specific-language condition.

Discussion

Language genericity altered several aspects of social categorization in addition to rates of identification. Participants who learned about a novel social group (noncitizens) in specific rather than generic language classified fewer unknown individuals as noncitizens, and this effect was increased when those identified faced criminal penalties. Participants in the specific-language condition also utilized a higher evidentiary standard to declare an individual a noncitizen and reported less decision certainty when the evidence

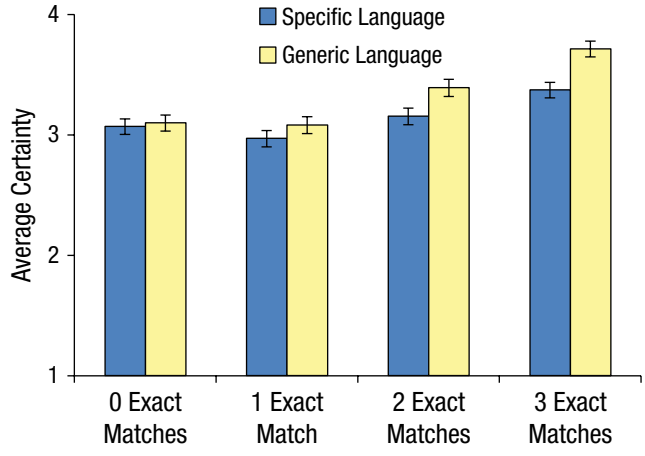


Fig. 10. Average certainty as a function of language (generic, specific) and evidence (zero, one, two, or three pieces of matching evidence). Error bars represent ±1 *SE*.

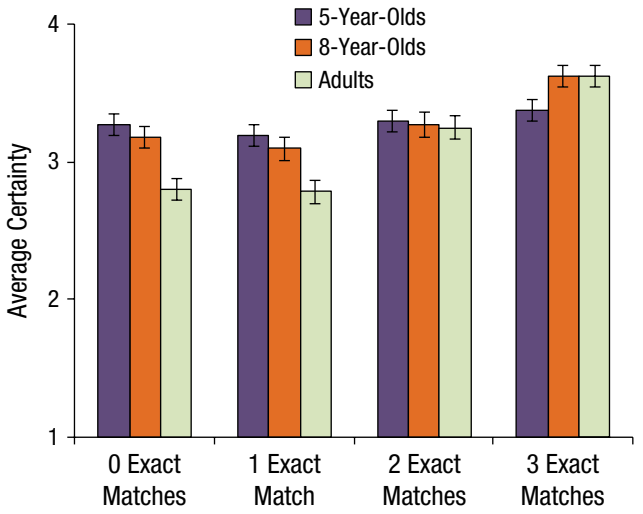


Fig. 11. Average certainty as a function of age group (5-year-olds, 8-year-olds, adults) and evidence (zero, one, two, or three pieces of matching evidence). Error bars represent ±1 *SE*.

weighed in favor of such categorization. Compared with children, adults needed less evidence to declare that unknown individuals were noncitizens, categorized individuals faster, reported lower certainty when the evidence was weak, and expressed greater certainty when all the evidence matched noncitizen exemplars. As found in prior work, generic language had a stronger effect in adults than in young children. In this section, we elucidate these findings further and provide directions for future research.

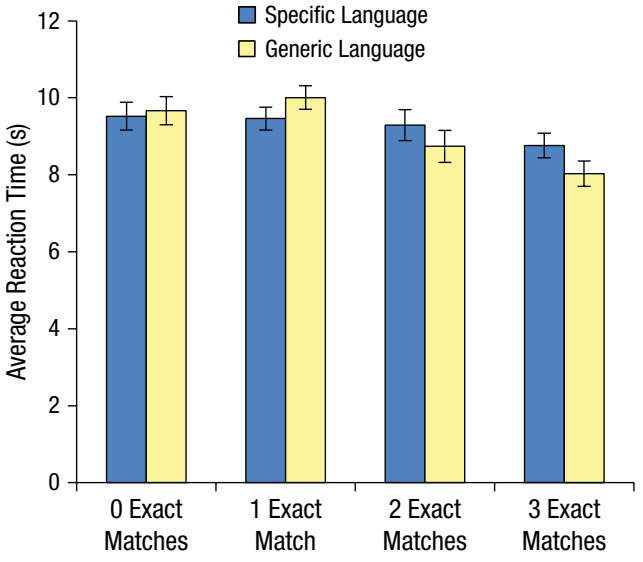


Fig. 12. Average reaction time as a function of language (generic, specific) and evidence (zero, one, two, or three pieces of matching evidence). Error bars represent ±1 *SE*.

Language genericity and criminal sanctions

Learning about group members in generic rather than specific language increased the frequency at which participants declared individuals to be noncitizens, and this effect nearly doubled when those identified faced jail and deportation. Thus, the criminal context magnifies the well-established effect of language genericity on categorization (e.g., Gelman et al., 2010). Criminal sanctions reduced identification of individuals as noncitizens only when participants learned about this social group as individuals. Notably, the generic-language effect persisted despite our stringent test: Irrespective of language condition, all participants received brief training on group characteristics, received the same memory cue sheet, and heard the category label ("Zuttle") the same number of times during training.

Several interrelated factors may have driven these effects of language and context. Generic language can make social boundaries appear more inherent (Kraus & Keltner, 2013), and thus potentially encourage participants to justify laws regulating groups. People help single identified victims more than groups (Kogut & Ritov, 2005), so perhaps first learning about noncitizens as named individuals increased participants' empathy or affiliation toward noncitizens, reducing the number of individuals they later identified for punishment. Even without confirming the exact mechanism (or mechanisms) underlying these effects, these data support reframing refugee crises as affecting individuals with unique stories, rather than amorphous groups (Diamond, 2016). Reducing generic language may elicit more sensitive decision making, especially when combined with additional strategies for changing stereotypes about marginalized groups (Gaertner et al., 2000).

Evidentiary threshold

Generic language reduced the evidentiary threshold from a preponderance of evidence to a single piece of matching evidence. Generic language thus bolsters category structures by lowering the tipping point for identifying category members. Extrapolating this finding to the real world suggests that when governments call on citizens to report suspected "noncitizens" (Aliverti, 2015), they may encourage misidentifications based on minimal proof. Examples of utilizing a lowered evidentiary threshold to classify individuals in a social group include taunting "build that wall" to an opposing team solely because of the team's perceived ethnicity (Holley, 2016) and alerting security because an airplane passenger is writing in a suspicious language, math (Rampell, 2016).

By taking a developmental perspective, we gained insight into age-related changes in the formation of social category structures. Contrary to predictions that young children would utilize a simple positive-capture strategy (i.e., "if anything matches, say 'noncitizen'"), our data support the alternative hypothesis: Children use a *higher* evidentiary standard than adults—regardless of language condition. Compared with the children in our study, the adults may have more often believed that the presented characteristics were defining features, such that each characteristic was sufficient to categorize an individual (Deng & Sloutsky, 2015; Keil & Batterman, 1984). The adults may have more often assumed that the presented characteristics were exclusive to noncitizens and doubled the weight of each matching attribute (i.e., it was evidence *for* noncitizenship and *against* citizenship). Further, the adults' greater knowledge base may have facilitated more cohesive theories about how evidentiary trios, consisting of weak and indeterminate evidence, could still indicate noncitizen status (Murphy & Medin, 1985).

Decision confidence

Language genericity altered not only the criteria for identifying group members, but also decision confidence. Across language conditions, participants reported higher certainty and categorized individuals faster when there were two or three pieces of matching evidence. These findings suggest that more than a single piece of matching evidence led to firmer beliefs (K. E. Lyons & Ghetti, 2011). Generic (vs. specific) language further boosted participants' certainty when the evidence favored noncitizenship. These data strengthen claims that generic information is perceived as universally known (Cimpian & Scott, 2012) and inherent to knowledge systems (Sutherland & Cimpian, 2015). They further reveal that language genericity may influence the stability of category structure: Lower certainty (from specific language) indicates room for flexibility and change in social-group concepts, whereas higher certainty (from generic language) signals more rigid and durable social categories.

Our results also confirm prior studies finding that recognition of indeterminacy improves with age (Lagattuta & Sayfan, 2011), as the children in our study expressed greater confidence than adults when the evidence was weak. In contrast, when all three characteristics matched previous exemplars, the 8-year-olds and adults reported higher certainty than the 5-year-olds. This exaggerated confidence (nearing "very sure" on our certainty scale) indicates that beliefs about social categories can be influenced by the *illusion-of-validity* heuristic (Kahneman & Tversky, 1973), a bias found to increase between childhood and adulthood in other social judgment tasks (Lagattuta & Sayfan, 2013). Participants' certainty for all judgments should have been mitigated by the fact that the features presented are ones that are widely shared (e.g., liking apples and pears, being bad at tennis).

Extending the paradigm

This paradigm opens avenues for future work. Our group criterion was citizenship, a legal construct not tied to appearance, preferences, or abilities (e.g., one does not like hamburgers immediately upon being sworn in as an American citizen). We described noncitizens with benign characteristics and never raised alarm about this group. We provided no rationale for why they left their place of origin or why they returned. To ensure that categorization was uncontaminated by group membership (Roberts & Gelman, 2015), we created a scenario in which participants were third-party observers who learned about novel (rather than familiar) social categories. Researchers could manipulate these experimental features to test further how language genericity and context shape social categorization.

Conclusion

Categorical thinking aids cognitive processing by simplifying and organizing complex information (see Macrae & Bodenhausen, 2000). Here, we have shown that speaking in generic terms precipitates assumptions about the stability and uniformity of social categories: When children and adults learn about a target group in generic (rather than specific) language, they require less evidence to label individuals as members of that group, ignore negative consequences to those so identified, and express higher certainty about their choices. Despite the rightful concerns that we, and other researchers, have raised about generic language, discussing members of social groups in specific language does not eradicate biases about those groups. The rates for identifying presumed noncitizens and decision confidence were still quite high even when participants learned about group members as named individuals (Figs. 6 and 10). Thus, excitement surrounding the benefits of specific over generic language needs to be tempered by recognition of how effortlessly beliefs about social categories coalesce in people's minds.

Action Editor

Brian P. Ackerman served as action editor for this article.

Author Contributions

D. Goldfarb and K. H. Lagattuta developed the study concept. D. Goldfarb, K. H. Lagattuta, and K. Kennedy designed the study. All the authors contributed to testing and data collection. D. Goldfarb, K. H. Lagattuta, and H. J. Kramer analyzed and interpreted the data, and also wrote the manuscript with assistance from K. Kennedy and S. M. Tashjian. All the authors approved the final version of the manuscript for submission.

Acknowledgments

We thank the participants, as well as the undergraduate research assistants for helping with stimulus creation, recruitment, data collection, and data entry. We also thank Karen Hjortsvang Lara for her feedback on the study.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

An internal grant from the University of California, Davis, to K. H. Lagattuta funded this study. H. J. Kramer was supported by the National Institutes of Health (Grant 201302291).

Notes

1. Note that the illustrations in Figures 1 through 4 differ from the actual stimuli in the size and spacing of the slides' components. Exact stimuli are available upon request from the corresponding author.

2. The Context × Block Order interaction did not affect any of the interpreted interactions. Generally, order (0 = noncriminal context first, 1 = criminal context first) negatively predicted reaction time for all levels of evidence in the noncriminal context (bs < –1.62, ps < .006). Participants were faster in sorting when they had the noncriminal condition second, regardless of the amount of evidence. Similar results were found in the criminal context, but in this case, participants were faster in their second sort only when one or three pieces of evidence matched (bs > 1.05, ps < .043), but not zero or two (bs < 0.88, ps > .220).

References

Aliverti, A. (2015). Enlisting the public in the policing of immigration. *British Journal of Criminology*, *55*, 215–230. doi:10.1093/bjc/azu102

Brown, C. S. (2011). American elementary school children's attitudes about immigrants, immigration, and being an American. *Journal of Applied Developmental Psychology*, *32*, 109–117. doi:10.1016/j.appdev.2011.01.001

Callahan, S. P., & Ledgerwood, A. (2016). On the psychological function of flags and logos: Group identity symbols increase perceived entitativity. *Journal of Personality and Social Psychology*, *110*, 528–550. doi:10.1037/pspi0000047

Cimpian, A., & Markman, E. M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child Development*, *82*, 471–492. doi:10.1111/j.1467-8624.2010.01525.x

Cimpian, A., & Scott, R. M. (2012). Children expect generic knowledge to be widely shared. *Cognition*, *123*, 419–433. doi:10.1016/j.cognition.2012.02.003

Deng, W. (S.), & Sloutsky, V. M. (2015). The development of categorization: Effects of classification and inference training on category representation. *Developmental Psychology*, *51*, 392–405. doi:10.1037/a0038749

Diamond, A. (2016, November 7). 'They were just like us, and they lost everything': Can empathy for refugees be taught? *The Atlantic*. Retrieved from http://www.theatlantic.com/international/archive/2016/11/msf-refugees-empathy-exhibit/506780/

Diesendruck, G., & Menahem, R. (2015). Essentialism promotes children's inter-ethnic bias. *Frontiers in Psychology*, *6*, Article 1180. doi:10.3389/fpsyg.2015.01180

Diesendruck, G., Salzer, S., Kushnir, T., & Xu, F. (2015). When choices are not personal: The effect of statistical and social cues on children's inferences about the scope of preferences. *Journal of Cognition and Development*, *16*, 370–380. doi:10.1080/15248372.2013.848870

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development*, *82*, 793–811. doi:10.1111/j.1467-8624.2011.01577.x

Fay, A. L., & Klahr, D. (1996). Knowing about guessing and guessing about knowing: Preschoolers' understanding of indeterminacy. *Child Development*, *67*, 689–716. doi:10.1111/j.1467-8624.1996.tb01760.x

Gaertner, S. L., Dovidio, J. F., Banker, B. S., Houlette, M., Johnson, K. M., & McGlynn, E. A. (2000). Reducing intergroup conflict: From superordinate goals to decategorization, recategorization, and mutual differentiation. *Group Dynamics: Theory, Research, and Practice*, *4*, 98–114. doi:10.1037/1089-2699.4.1.98

Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology*, *61*, 273–301. doi:10.1016/j.cogpsych.2010.06.001

Graham, S. A., Nayer, S. L., & Gelman, S. A. (2011). Two-year-olds use the generic/nongeneric distinction to guide their inferences about novel kinds. *Child Development*, *82*, 493–507. doi:10.1111/j.1467-8624.2010.01572.x

Gülgöz, S., & Gelman, S. A. (2015). Children's recall of generic and specific labels regarding animals and people. *Cognitive Development*, *33*, 84–98. doi:10.1016/j.cogdev.2014.05.002

Holley, P. (2016, November 17). White Texas teens chant 'build that wall' at Hispanics during high school volleyball match. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/early-lead/wp/2016/11/17/white-texas-teens-chant-build-that-wall-at-hispanics-during-high-school-volleyball-match/?utm_term=.6488ffa77804

Jarvis, B. G. (2012). DirectRT (Version 2012) [Computer software]. New York, NY: Empirisoft.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251. doi:10.1037/h0034747

Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal*

of Verbal Learning and Verbal Behavior, *23*, 221–236. doi:10.1016/S0022-5371(84)90148-8

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, *84*, 1329–1342. doi:10.1016/j.neuron.2014.12.015

Killen, M. (2007). Children's social and moral reasoning about exclusion. *Current Directions in Psychological Science*, *16*, 32–36. doi:10.1111/j.1467-8721.2007.00470.x

Killen, M., Mulvey, K. L., & Hitti, A. (2013). Social exclusion in childhood: A developmental intergroup perspective. *Child Development*, *84*, 772–790. doi:10.1111/cdev.12012

Kogut, T., & Ritov, I. (2005). The "identified victim" effect: An identified group, or just a single individual? *Journal of Behavioral Decision Making*, *18*, 157–167. doi:10.1002/bdm.492

Kraus, M. W., & Keltner, D. (2013). Social class rank, essentialism, and punitive judgment. *Journal of Personality and Social Psychology*, *105*, 247–261. doi:10.1037/a0032895

Lagattuta, K. H., & Sayfan, L. (2011). Developmental changes in children's understanding of future likelihood and uncertainty. *Cognitive Development*, *26*, 315–330. doi:10.1016/j.cogdev.2011.09.004

Lagattuta, K. H., & Sayfan, L. (2013). Not all past events are equal: Biased attention and emerging heuristics in children's past-to-future forecasting. *Child Development*, *84*, 2094–2111. doi:10.1111/cdev.12082

Lyons, K. (2016, June 26). Racist incidents feared to be linked to Brexit result. *The Guardian*. Retrieved from https://www.theguardian.com/politics/2016/jun/26/racist-incidents-feared-to-be-linked-to-brexit-result-reported-in-england-and-wales

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*, 1778–1787. doi:10.1111/j.1467-8624.2011.01649.x

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, *51*, 93–120. doi:10.1146/annurev.psych.51.1.93

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316. doi:10.1037/0033-295X.92.3.289

Potok, M. (2016). *The year in hate and extremism*. Retrieved from https://www.splcenter.org/fighting-hate/intelligence-report/2016/year-hate-and-extremism

Rampell, C. (2016, May 7). Ivy League economist ethnically profiled, interrogated for doing math on American Airlines flight. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/rampage/wp/2016/05/07/ivy-league-economist-interrogated-for-doing-math-on-american-airlines-flight/?utm_term=.cc5e2a1fbc70

Rhodes, M., Leslie, S. J., Saunders, K., Dunham, Y., & Cimpian, A. (2017). How does social essentialism affect the development of inter-group relations? *Developmental Science*. Advance online publication. doi:10.1111/desc.12509

Rhodes, M., Leslie, S. J., & Tworek, C. M. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *USA*, *109*, 13526–13531. doi:10.1073/pnas.1208951109

Roberts, S. O., & Gelman, S. A. (2015). Do children see in black and white? Children's and adults' categorizations of multiracial individuals. *Child Development*, *86*, 1830–1847. doi:10.1111/cdev.12410

Sutherland, S. L., & Cimpian, A. (2015). Children show heightened knew-it-all-along errors when learning new facts about kinds: Evidence for the power of kind representations in children's thinking. *Developmental Psychology*, *51*, 1115–1130. doi:10.1037/a0039463

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*, 1–39. doi:10.1146/annurev.ps.33020182.000245

Tasimi, A., Gelman, S. A., Cimpian, A., & Knobe, J. (2016). Differences in the evaluation of generic statements about human and non-human categories. *Cognitive Science*. Advance online publication. doi:10.1111/cogs.12440