Abstract

Generics (e.g., “Dogs bark”) are thought by many to lead to essentializing: to assuming that members of the same category share an internal property that causally grounds shared behaviors and traits, even without evidence. Similarly, generics are thought to increase generalizing, that is, attributing properties to other members of the same group given evidence. However, it is not clear from past research what underlies the capacity of generic language to cause increased essentializing and generalizing. Is it something special about generics, or are there broader mechanisms at work, such as the fact that generics are terms that signal high proportions? Our studies suggest there is less of a special role for generics in essentializing and generalizing than previously supposed. Study 1 (100 5-6 year-olds, 140 adults) found high proportion quantifiers (“most”, “many”) elicited essentializing about a novel social kind (Zarpies) as well as generics for adults; and elicited generalizing to an even greater extent than generics. Specifics (“this”) did not protect against either essentializing or generalizing when compared to the quantifier “some”. Study 2 (100 5-6 year-olds, 112 adults) found neither generics nor visual imagery indicating multiple instances led to essentializing in adults. While generics increased generalizing compared to specifics or visual imagery alone, there was no difference in generalizing for generics compared to repeated visual imagery accompanied by repeated specifics, again suggesting no protective role of specifics. In contrast to what others have found, there were no differences in essentializing or generalizing across conditions in either study for children.

Keywords: Generics; Quantifiers; Essentializing; Stereotype; Prejudice

Language Signalling High Proportions, not Just Generics, Leads to Essentializing for Novel Social Kinds

Generics, it is claimed, are a double-edged sword. They help children efficiently learn about the world (Cimpian & Markman, 2008; Gelman & Raman, 2003; Gelman, Star, & Flukes, 2002; Graham, Gelman, & Clarke, 2016; Graham, Nayer, & Gelman, 2011; Hollander, Gelman, & Star, 2002), but it is also suggested they play a special role in essentializing and stereotype formation in adults and children (Cimpian, 2010; Cimpian & Erickson, 2012; Cimpian, Gelman, & Brandone, 2010; Cimpian & Markman, 2011; Gelman, Ware, & Kleinberg, 2010; Rhodes, Leslie, Bianchi, & Chalik, 2018; Rhodes, Leslie, Saunders, Dunham, & Cimpian, 2018; Rhodes, Leslie, & Tworek, 2012). However, it is not clear from past research what underlies the capacity of generic language to increase essentializing. Is it something special about generics, or are there broader mechanisms at work, such as the fact that generics are terms that signal high proportions? What can we learn about the underlying mechanism causing increased essentializing by consideration of the component indicators used to measure essentializing? This paper seeks to determine whether: (1) generics in and of themselves cause essentializing in novel social kinds; (2) other high proportion quantifiers (“most”, “many”) cause essentializing as much as generics; (3) specifics protect against essentializing, rather than generics causing essentializing; and (4) imagery depicting multiple instances of the same behavior cause essentializing as much as generics.

Generics are sentences such as “Cats have tails” and “Girls like pink.” They express generalizations, but unlike quantified sentences, do not explicitly articulate information about how many members of the category have the given property (Carlson, 1977; Gelman & Tardif, 1998; Lawler, 1973; Leslie & Gelman, 2012; Pappas & Gelman, 1998). Another distinction between generics and quantifiers is that, given a striking property (carrying the West Nile virus) or a characteristic property (laying eggs), generics are often judged as true even if only a small percentage of the kind has these properties (Prasada, Khemlani, Leslie, & Glucksberg, 2013). Further, unlike specifics (e.g., “This cat has a tail”), generics are not about particular individuals, but rather are thought to carry general information relating the kind to the given property. Several studies seem to indicate generics help adults and children learn novel information about known (Gelman et al., 2002; Nguyen & Gelman, 2012), and novel kinds (Chambers, Graham, & Turner, 2008; Gelman et al., 2010; Graham et al., 2016; Graham et al., 2011; Hollander, Gelman, & Raman, 2009; Stock, Graham, & Chambers, 2009). For instance, Chambers et al. (2008) introduced novel creatures to 4- and 5-year-olds, and taught them properties of the creatures, using either generics (e.g., “Pagons are friendly”), or specifics (e.g., “These pagons are friendly.”) Children were significantly more likely to extend the properties to novel instances of the creatures when generics were used.

Psychological essentialism is the assumption that members of a category are similar at their core, such that they tend to share qualities, even with no explanation for why they should share these qualities (Gelman, 2004). Several studies demonstrate generics increase essentializing for novel animals kinds (Cimpian et al., 2010; Gelman et al., 2010), familiar social kinds when generics match their prior knowledge (Cimpian & Erickson, 2012; Cimpian & Markman, 2011), and novel social kinds (Rhodes, Leslie, Bianchi, et al., 2018; Rhodes, Leslie, Saunders, et al., 2018; Rhodes et al., 2012). For instance, Rhodes et al. (2012) had adults and children (3 and 4 years) read a picture book about a novel social category of people called “Zarpies.” The book gave information about Zarpies, using either generics (e.g., “Zarpies love to eat flowers”), specifics (e.g., “This Zarpie loves to eat flowers”), or specifics with the category label withheld (e.g., “This one loves to eat flowers.”) Participants were then asked questions to determine how much they essentialized Zarpies. For instance, they were asked whether an adopted Zarpie would have the same characteristics as their biological Zarpie mother (essentializing) or their adoptive non-Zarpie mother (non-essentializing). They found children and adults essentialized more when generics were used compared to specifics.

Some note that generics increase essentializing, and potentially stereotyping and prejudice (Goldfarb, Lagattuta, Kramer, Kennedy, & Tashjian, 2017), and have suggested the use of generic language should perhaps be avoided (Leslie, 2017; Rhodes et al., 2012; Wodak, Leslie, & Rhodes, 2015). For instance, Leslie (2017, p. 42) claims,

“However, the evidence suggests that the use of labels and generics contributes to essentialization, and so the converse may also hold: reducing the use of labels and generics for racial, ethnic, and religious groups may reduce the extent to which children grow up essentializing these groups.”

But if generics do increase essentializing of novel social kinds, what underlies their capacity to do so? Below, we first motivate an understanding of what the different component measures of essentializing indicate, following Gelman et al. (2010), and argue some may be better indicators of generalizing rather than essentializing; then we turn to plausible alternative hypotheses which the available evidence does not rule out. In the studies described later, we explore these alternative measures and hypotheses.

In Rhodes et. al. (2012) and Rhodes et. al. (2018), questions denoting generalizing and essentializing were combined to give an essentializing score (Gelman et al., 2010, used the term “category-property composite” where we use the term “generalization”). While generalizing can indicate essentializing (if one essentializes, then this may lead to generalizing), following Gelman et al. (2010), we suggest it is important to distinguish generalizing from essentializing, as one may generalize without essentializing. For instance, one may conclude all the people at the bus stop are waiting for the bus without essentializing those people. A willingness to attribute to a new Zarpie a property one already encountered in a Zarpie may indicate simply a willingness to generalize. By contrast, a willingness to attribute a *new* property indicates more of a tendency to attribute sameness to Zarpies, and so is a better indicator (though still not infallible) of essentializing. For instance, if a participant read, “Zarpies hate ice cream” in training, and then responded novel Zarpies also hate ice cream during testing, this counted as essentializing in previous studies. However, we suggest this may be a better indicator of generalizing the information learned from the book, not necessarily essentializing (Gelman et al., 2010). Additionally, if a participant read, “Zarpies climb tall fences” and then was asked, “Why is this [same] Zarpie climbing a tall fence?”, responses including generic language, e.g., “Because Zarpies like to” were counted as essentializing previously. However, again, this may be a better measure of generalizing information learned in the book, and may not indicate essentializing (Gelman et al., 2010).

Furthermore, Rhodes, Leslie, Saunders, et al. (2018) awarded an additional point to responses on the previously discussed “Why” questions for giving an internal response, which we agree is essentializing. However, they state, “Because he likes to” counts as an internalizing (and therefore essentializing) response, but it is not clear to us this is necessarily the case. Answers such as “Because he likes to” should not necessarily, we think, count as essentializing. If we do not know *why* they like to, we cannot determine whether it is due to internal or external reasons. For example, if asked why Americans watch football, the answer, “Because they like to” does not seem necessarily essentialist. In the case of Zarpies, Zarpies might like chasing their own shadows because they are genetically predisposed to be attracted to dark Zarpie-like shapes (an internal explanation, which is essentialized) (Cimpian & Markman, 2011; Noyes & Keil, 2020); or they might have learnt to chase shadows in their Zarpie gym class, and come to enjoy it over time (a practice explanation, which is external and nonessentialized) (Cimpian & Markman, 2011; Noyes & Keil, 2020). Given these considerations, it is possible generalizing, rather than essentializing, drove some existing results. Therefore, the first question we address is to what extent generics increase essentializing — rather than generalizing — for novel social kinds.

Next, we consider several plausible alternatives to the hypothesis generics have a distinctive and special connection to essentializing novel social kinds. Is it speech involving generics *per se* which causes essentializing, or are generics just one example of a type of language which does this? Quite plausibly, quantifiers such as “many” or “most” could have the same effects because they, like generics, signal high proportions (Saul, 2017). Cimpian and Erickson (2012) hypothesized quantifiers denoting high proportions, e.g., “most”, lead to generic beliefs, which in turn lead to essentializing, and indeed, found “most” was just as likely as generics to produce essentialized explanations for familiar social kinds (girls). Therefore, we expect the same to hold for novel social kinds. However, our hypothesis differs from Cimpian and Erickson as we propose any language signalling high proportions, including generics, “many”, and “most”, may lead directly to essentializing, without the need to form generic beliefs. Cimpian and Erickson support the inclusion of generic beliefs, noting high proportion quantifiers are often misremembered as generics. Indeed, English-speaking adults misremember high proportion quantifier sentences as generics around 45% of the time, while making the inverse error around 12% of the time (Gelman, Sanchez Tapia, & Leslie, 2016). However, if we follow this logic, we might only expect people to essentialize when exposed to a high proportion quantifier around half the time they essentialize when exposed to generics due to their generic-encoding error rate (45%). This seems to us to pose difficulties for their supposition that an intermediate belief of generic form is present. If instead Cimpian and Erickson argue high proportion quantifiers lead to generic beliefs, which lead to essentializing, even when correctly remembering the original statement as a high proportion quantifier, it is not clear their evidence supports the inclusion of generic beliefs. A simpler explanation (one that fits better with the error rate) is that language signalling high proportions (including generics, when framed as high proportion quantifiers) leads directly to essentializing. Our second goal, then, is to determine whether language signalling high proportions increases essentializing as much as generics.

Some have suggested we might generalize by default (Leslie, 2008; Leslie & Gelman, 2012). Therefore, a further alternative hypothesis is that generics do not cause essentializing, but rather, specifics protect against essentializing. While some papers tried to control for this possibility with a label-free condition (e.g., “This one likes to eat flowers”) (Leslie & Gelman, 2012), these controls still contain the specific marker “this.” Further, while several papers found generics are common in parents’ language, making up 3-5% of utterances (Gelman et al., 2002; Gelman & Tardif, 1998), specifics are more common than generics, with parents using up to four times as many specifics as generics (Nyhout & O'Neill, 2014). Therefore, one hypothesis is people essentialize information by default, unless specifics highlight that the information does not pertain to a category. This would mean specifics do the work of individuating, and frequent parental use of specifics might serve the function of ensuring children learn exceptions. Our third goal is, therefore, to determine whether generics (and possibly other high proportion terms) cause essentializing in novel social kinds, or whether specifics protect against essentializing.

Another possibility is it may not even be *language* signalling high proportions which increases essentializing, but representations of multiple instances. Pre-verbal infants already show bias against people of other races (Fassbender, Teubert, & Lohaus, 2016; Kelly et al., 2005; Kinzler, Shutts, DeJesus, & Spelke, 2009; Liu, Xiao, Quinn, et al., 2015; Liu, Xiao, Xiao, et al., 2015; Xiao, Quinn, et al., 2018; Xiao, Wu, et al., 2018). Additionally, 4- to 9-year-olds expected group members to conform to novel norms, in one case, eating a specific type of berry, when group members were presented in a group of three (without a label) more than when group members were presented individually (without a label) (Roberts, Ho, & Gelman, 2017). Furthermore, 5-year-olds showed intergroup bias when they were arbitrarily allocated to an orange or green group, even though group names were never used, and were only indicated by the color of their t-shirts (Dunham, Baron, & Carey, 2011). Therefore, it is possible that *seeing* multiple members of a novel social group acting the same could increase essentializing. Thus, generics, and possibly other types of language indicating high proportions such as “most” and “many”, may increase essentializing not because *language* increases essentializing, but because this type of language serves as a shortcut to a representation of high proportions.

Study 1 sought to determine to what extent: (1) generics cause essentializing (rather than only generalizing) in novel social kinds; (2) language signalling high proportions (not just generics) causes essentializing; and (3) specifics protect against essentializing novel social kinds. Adults and children read a picture book about fictional people called “Zarpies” (Rhodes, Leslie, Saunders, et al., 2018; Rhodes et al., 2012). Each book page featured a Zarpie performing an activity, such as eating flowers (see Appendix A). In both studies, we replicated the generics and specifics conditions by using generic (e.g., “Zarpies love to eat flowers.”) and specific text (e.g., “This Zarpie loves to eat flowers.”) We also added two other types of language signalling high proportions: “most” and “many”; and one type of language signalling uncertain proportions: “some.”

The goal of Study 2 was to determine whether repeated imagery causes essentializing to the same extent as generics when presenting novel social kinds. We added a label-only condition, with pictures but no explanatory text. For instance, participants might see a picture of a person eating flowers, accompanied by “Zarpie” with an arrow pointing at the person. We expected this condition to signal a single instance, like the specifics condition. We also added two conditions to signal multiple instances without using high proportion language. First, we added an additional specifics condition where each behavior repeated three times in a row (e.g., three different Zarpies all hating ice cream on three sequential pages, with every page saying, “This Zarpie hates ice cream.”), such that the language signalled one instance, but the images signalled multiple instances. Finally, we added a label-only condition where each behavior was repeated three times by different Zarpies, again, so the images indicated multiple instances.

**Study 1**

In this experiment, we exposed participants to one of five stories about Zarpies. Two conditions replicated previous research: *Generics* and *This* (specifics) (Rhodes, Leslie, Saunders, et al., 2018; Rhodes et al., 2012). We added three new conditions to determine whether differences between *Generics* and *This* were driven by *Generics*, *This*, both, or the signalling of high proportions more generally. *Most* and *Many* conditions were added as markers of high proportions, and a *Some* condition was added as a marker of an uncertain proportion (see Appendix B). Children were 5- and 6-year-olds to ensure they understood quantifiers “most,” “many” and, “some” (Brandone, Gelman, & Hedglen, 2015; Cimpian & Erickson, 2012; Gelman, Leslie, Was, & Koch, 2015; Halberda, Taing, & Lidz, 2008), so they would not potentially recall these words as generics as a potential “default” (Gelman, Leslie, Gelman, & Leslie, 2019).

**Method**

**Participants.** Previous similar research was powered to find large effect sizes (Gelman et al., 2010; Rhodes, Leslie, Saunders, et al., 2018; Rhodes et al., 2012).A G-Power analysis found 80 participants were needed for a large effect size (f = 0.4) with power at 0.80 and alpha at 0.05 (Faul, Erdfelder, Lang, & Buchner, 2007). We tested at least 80 participants in both the adult and child groups in case the patterns were different and data needed to be analysed separately. We aimed to recruit more participants than needed in case some participants did not complete the study, did not complete it properly (e.g., giving responses bearing no relation to the questions, 1 adult), were ineligible (e.g., adults for whom English was not a first language, 6), or did not want to participate (i.e., children, 0). We recruited more participants than required, especially for adults, where our study was quickly completed online. There were 140 adults: 71 males, 68 females and one identifying as neither (M age = 31.35 years, SD = 10.29, Range = 18-65). Each participant was randomly assigned to each of five conditions. All participants spoke English as a first language. Participants lived in the United Kingdom (81), the United States of America (45), or other countries (13). Participants were recruited through Prolific Academic, and were paid £1GBP (for less than 10 minutes work).

There were 100 children: 50 males and 50 females (M age = 71 months, 20 days SD = 5 months, 22 days Range = 60 months, 1 day – 84 months 20 days). Each participant was randomly assigned to each of five conditions. All participants spoke English as a first language and lived in the United Kingdom. Participants were White (45), Asian (5), Black (1), of mixed ethnicity (5), or ethnicity information was not reported (45). Parents had a postgraduate degree (13), Undergraduate degree (17), High school diploma (15), no academic qualification (1) or education information was not reported (54). Participants were recruited via email and telephone correspondence with schools. Schools were paid £5GBP for each child who participated, with which educational resources were bought. Parents signed a written consent form. The study was approved by the University of XXXX Psychology Department’s ethics committee: “Generics and Essentialising”; approval number: 003190.

**Materials.** Five picture books were created using Zarpies (using the original Rhodes, et al., 2012 books), differing for each condition by the wording. Each book consisted of the same 16 illustrated pages, showing a picture of a Zarpie displaying a distinctive physical or behavioral characteristic (e.g., “Zarpies… love to eat flowers; have stripes in their hair”; see Appendix A for the full *Generics* version of the book; see Appendix B for all text across books). The characters were diverse in terms of gender (half male, half female), race/ethnicity (four White, four Black, four Latino and four Asian) and age (four young children, four older children, four adults and four older adults). To ensure participants could visually identify Zarpies as a categorical membership, the characters were dressed in category-typical clothing, without any single feature defining category membership.

The five conditions and hence separate books were: *Generics, This, Many, Most* and *Some*. The wording of each book was modified to match the condition. In the *Generics* condition the book read e.g., “Zarpies love to eat flowers.” In the *This* condition the book read e.g., “This Zarpie loves to eat flowers.” In the *Many* condition, the book read e.g., “Many Zarpies love to eat flowers.” In the *Most* condition, the book read e.g., “Most Zarpies love to eat flowers.” Finally, in the *Some* condition, the book read, “Some Zarpies love to eat flowers.”

Thirteen test questions were divided into two categories of essentialist beliefs (Gelman et al., 2010). These included: three inheritance questions, where participants were asked whether a Zarpie child who was adopted by a non-Zarpie mother would behave in the same way as the Zarpie mother or the adoptive mother; six induction questions in which a new Zarpie displayed a property not seen in the book, and participants were asked if further novel Zarpies would display the same property; and four explanation questions where participants were asked *why* a Zarpie did something in the previous book. See Appendix C for all items. Questions were exactly the same as those in Rhodes et al. (2012).

As a control to determine whether past research findings were perhaps driven by generalizing rather than essentializing (Rhodes, Leslie, Saunders, et al., 2018; Rhodes et al., 2012), 10 test questions were included for which answers might indicate generalizing (Gelman et al., 2010). These included six induction questions asking whether or not new Zarpies had the same properties as the Zarpies in the book. Affirmative answers indicated generalizing, but not necessarily essentializing. See Appendix C for all items. The four explanation questions for which participants were asked *why* a Zarpie had done something in the previous book were also analysed for generalizing answers.

**Design.** The experiment was a between-subjects design. The independent variables were the type of language used (*Generics, This, Many, Most, Some*) and participants’ age (Adult, Child). The dependent variables were whether participants gave (1) a combination of essentialized and generalized answers (replicating Rhodes, et al., 2012); (2) essentialized answers; and (3) generalized answers (as a control).

**Procedure.** Adults from Prolific Academic accessed the book and test questions online using the Qualtrics website. They read the book independently and answered the test questions immediately after.

We verbally presented books and test questions to children during individual sessions. We read the book twice, and asked questions immediately after (following Study 2 of Rhodes, et al., 2012). Responses were recorded on an Olympus MP3 recorder whilst also being written down by the experimenter when questioning.

**Coding.**

***Essentializing.***

*Inheritance questions*. These were coded as essentializing if participants said the child had the same property as the Zarpie mother (1 point), and were coded as not essentializing if participants said the child had the same property as the adoptive mother (0 points). If a participant did not respond, no score was assigned.

*Induction of novel properties.* These were coded as essentializing if participants responded “yes” (1 point), and as not essentializing if participants said “no” (0 points). If a participant did not respond, no score was assigned.

*Explanation questions.* Explanation questions (e.g., “Why is this Zarpie climbing a tall fence?”) were given one point for essentializing if the answer was due to an inherent explanation (e.g., “His legs were *made for* climbing,” “*God made them* that way,” “Because of *DNA*”), or a trait explanation (e.g., “They are *crazy*,” “Zarpies are *strong*”) (Cimpian & Markman, 2011). Zero points were given if the answer referred to an external explanation (e.g., “He was trying to *get to the other side*,” “They chase shadows because they *think it’s a person*”), or practice explanations (e.g., “Zarpies are *taught* to jump over puddles”). The answer was not scored if this could not be inferred (e.g., “running,” “I don’t know”). In contrast to Gelman et al. (2010) and Rhodes, Leslie, Saunders, et al. (2018) who counted responses such as “Because he likes to” as essentializing, we did not for our main essentializing analyses. Since we do not know *why* they like to, we cannot determine whether it is due to internal or external reasons (see introduction). Out of a possible 960 explanations, only seven suggested trait-based essentializing (e.g., “Because he has a mental illness”), and five suggested inherent explanations (e.g., “Because he is an average Zarpie”). Additionally, these responses were spread across conditions: Generics (3); Most (3); Many (1); Some (4); This (1). Therefore, this measure was dropped from essentializing only analyses. Using multiple imputation for missing values (Béland, Pichette, & Jolani, 2016), the nine remaining essentializing items used in our essentializing only analyses showed good reliability, *KR(20)* = .73.

However, we included e.g., “Because he likes to” responses in our first analysis collapsed across essentializing and generalizing items to determine how our data compared to the original Rhodes et al. (2012) paper. Fifty-five participants’ (23%) explanations were coded for essentializing, using the Rhodes et al. (2012) coding scheme by a second coder. Agreement was acceptable, *Cohen’s kappa* = 0.77.

***Generalizing.*** This was used as a control measure for essentializing to determine whether past results showing generics increased essentializing (Rhodes, et al., 2012; Rhodes, Leslie, Saunders, et al., 2018) were perhaps led by generalizing rather than essentializing.

*Induction of familiar properties.* These were coded as generalizing if participants responded “yes” (1 point), and as not generalizing if participants responded “no” (0 points). If a participant did not respond, no score was assigned.

*Explanation questions.* These were given one point for generalizing if the answer referred to a high proportion of Zarpies or had generic form (e.g., “Because *Zarpies* climb fences,” “*Most Zarpies* like climbing,” “*They* do that to get places,” “It’s *a Zarpie* thing”); zero points if the answer referred to one instance in non-generic form, or an uncertain proportion (e.g., “Because *this Zarpie* climbs fences,” “That’s what *some Zarpies* do,” “*He* likes climbing,” “*It* wants to get to the other side”); and the answer was not scored if this could not be inferred (e.g., “running,” “I don’t know”). A second coder coded 56 participants’ (23%) explanations for generalizing. Agreement was excellent, *Cohen’s kappa* = 0.91. Using multiple imputation for missing values, the 10 generalizing items showed good reliability, *KR(20)* = .85. Using multiple imputation for missing values, the 23 original essentializing and generalizing Rhodes et al. (2012) items coded as essentializing (i.e., all 23 items tested) showed good reliability, *KR(20)* = .84.

**Results**

Children’s scores were all skewed, therefore we could not use ANOVA. Instead, we used Logit Mixed Effects Models (LMEM). LMEMs allow repeated-measures non-parametric designs with missing data, and control for random effects, while increasing power by analysing by item instead of by participant (Hoicka & Akhtar, 2011). All analyses included participant code and item as random variables.

We first ran analyses on all questions together, including using the same coding scheme for essentializing Explanation items (i.e., “Because he likes to” would be counted as essentializing), to compare our results to the original Rhodes et al. (2012) paper. Figure 1 displays the mean percentage of trials and 95% confidence intervals for essentialized/generalized responses by language type and age group. Participants did not give an answer that could be coded (as essentializing/generalizing, or not essentializing/generalizing) for 4.9% of trials. Additionally, 31% of children’s explanation responses could not be coded for generalizing, of which 72% of answers did not allow us to determine the proportion (e.g., it is unclear whether “Going to park” refers to a high or low proportion of Zarpies), and 28% were due to children not responding or saying e.g., “I don’t know.” Furthermore, 12% of children’s explanation responses could not be coded for essentializing, of which 74% were due to children not responding or saying e.g., “I don’t know”, 19% were because children did not actually answer the question, e.g., “Most Zarpies”, and 6% were because we could not determine whether the answer was essentialist or not, e.g., “I think it's because it's got a smile on its face and it.” The model for adults and children together was improved by age group (*χ2*(1) = 56.84, *p* < .0001), language type (*χ2*(4) = 49.94, *p* < .0001), and an interaction of age group and language type (*χ2*(4) = 24.63, *p* < .001). Adults essentialized/generalized significantly more often than children (*Odds-Ratio, OR* = 3.23, *p* < .0001). Due to the interaction, we ran separate analyses for adults and children.

*Figure 1*. Percentage of essentialized/generalized responses, by language type and age group. Error bars represent 95% confidence intervals.

Language type improved the model for adults (*χ2*(4) = 70.91, *p* < .0001). For pairwise comparisons, we used treatment contrasts repeatedly, resulting in 10 pairwise comparisons instead of the standard four. Therefore, we used Bonferroni corrections for pairwise comparisons, such that a significant p-value was reduced to .02. The final model (*loglik* = -1724.7, *N* = 3144) found adults essentialized/generalized significantly more often in the *Many* and *Most* conditions than the *Generics*, *Some*, and *This* conditions, as well as significantly more in the *Generics* condition than the *Some* and *This* conditions (all *OR* > 2.18, *p* < .003). No other differences were found. Adults were significantly more likely to generalize/essentialize in the *Many* and *Most* conditions (*OR* > 2.55, *p* < .013). Adults were significantly more likely to *not* generalize/essentialize in the *Some* and *This* conditions (both *OR* > 2.52, *p* < .030). Adults were equally likely togeneralize/essentialize or not in the *Generics* condition (*p* > .05).

Language type did not improve the model for children. When analysing data collapsed across language types, the final model (*loglik* = -1147.5, *N* = 2108) found children were more likely to *not* generalize/essentialize than engage in generalizing/essentializing (*OR* = 2.87, *p* < .001).

Figure 2 displays the mean percentage of trials and 95% confidence intervals for essentialized responses (using our coding scheme) by language type, question type, and age group. Participants did not give an answer that could be coded (as essentializing, or not essentializing) for 0.2% of trials. The model for adults and children together was improved by age group (*χ2*(1) = 14.78, *p* = .0001), question type (*χ2*(1) = 5.50, *p* = .0191), an interaction of age group and question type (*χ2*(1) = 47.02, *p* < .0001), and an interaction between age group, question type, and language type (*χ2*(8) = 14.78, *p* < .0001; final model: *loglik* = -1171.7, *N* = 2155). Adults essentialized significantly more often than children (*OR* = 1.64, *p* = .0203). Participants essentialized significantly more often for induction of novel property questions than inheritance questions (*OR* = 2.65, *p* = .0198). Given the above interactions, we ran separate analyses for adults and children.

The model for adults was improved by question type (*χ2*(1) = 6.83, *p* = .0090), language type (*χ2*(4) = 14.56, *p* = .0058), and an interaction of question type and language type (*χ2*(4) = 28.56, *p* < .0001). The final model (*loglik* = -686.6, *N* = 1260) found adults essentialized significantly more often for induction of novel property questions than inheritance questions (*OR* = 8.13, *p* = .0012). We used Bonferroni corrections for pairwise comparisons. Adults essentialized significantly more often in the *Most* than *Many* and *Some* conditions (*OR* > 3.60, *p* < .0200). No other differences were found. Given the above interaction, we ran separate analyses for adults’ induction of novel property questions, and adults’ inheritance questions.

Language type improved the model for adults’ induction of novel property questions (*χ2*(4) = 16.78, *p* = .0021). We used Bonferroni corrections for pairwise comparisons. The final model (*loglik* = -449.4, *N* = 840) found adults essentialized significantly more often in the *Generics* and *Many* conditions than the *Some* and *This* conditions, and significantly more often in the *Most* condition than the *This* condition (all *OR* > 4.34, *p* < .019). No other differences were found. Adults were equally likely to essentialize or not essentialize within all conditions (all *p* ≥ .1000).

Language type did not improve the model for adults’ inheritance questions (*χ2*(4) = 7.58, *p* = .1083). The final model (*loglik* = -204.8, *N* = 420) found adults were significantly more likely to *not* essentialize based on inheritance questions (*OR* = 6.87, *p* < .0001).

The model for children was not improved by any variables or interactions. The final model (*loglik* = -483.0, *N* = 895) found children were significantly more likely to *not* essentialize overall (*OR* = 3.62, *p* < .0001).

*Figure 2*. Percentage of essentialized and generalized responses, by language type, question type, and age group. Error bars represent 95% confidence intervals.

Figure 2 displays the mean percentage of trials and 95% confidence intervals for generalized responses by language type, question type, and age group. Participants did not give an answer that could be coded (as generalizing, or not generalizing) for 8.9% of trials. The model for adults and children together was improved by age group (*χ2*(1) = 74.82, *p* < .0001), language type (*χ2*(4) = 73.46, *p* < .0001), an interaction of age group and language type (*χ2*(4) = 23.51, *p* = .0001), and an interaction of question type and language type (*χ2*(5) = 96.85, *p* < .0001). The final model (*loglik* = -1051.2, *N* = 2203) found adults generalized significantly more often than children (*OR* = 3.80, *p* = .0029). Given the above interaction, we ran separate analyse for adults and children.

Language type improved the model for adults (*χ2*(4) = 91.42, *p* < .0001). Using pairwise comparisons with Bonferroni corrections, the final model (*loglik* = -671.9, *N* = 1336) found adults generalized significantly more often in the *Many* than all other conditions, in the *Most* condition than the *Generics*, *Some*, and *This* conditions, and in the *Generics* than the *Some* and *This* conditions (all *OR* > 1.15, *p* < .0012). No other differences were found, therefore we did not test Familiar Property and Explanation questions separately. Adults were significantly more likely to generalize than not in the *Many* and *Most* conditions (both *OR* > 6.36, *p* < .0001). Adults were equally likely to generalize as to not generalize in the *Generics* condition (*p* > .6249). Adults were significantly more likely to *not* generalize in the *This* and *Some* conditions (both *OR* > 4.56, *p* < .0456).

The model for children was improved by question type (*χ2*(1) = 5.02, *p* = .0251), and an interaction between question type and language type (*χ2*(8) = 33.62, *p* < .0001). The final model (*loglik* = -403.5, *N* = 867) found no overall effect of language type. Given the above interaction, we ran separate analyses for each question type.

Language type did not improve the models for children’s induction of familiar property questions, nor their explanation items (both *χ2*(4) < 7.37, *p* > .1177). The final models (induction of familiar property questions: *loglik* = -297.4, *N* = 594; explanation items: *loglik* = -67.7 *N* = 273) found children were significantly more likely to *not* generalize across conditions for both question types (both *OR* > 5.22, *p* < .0001).

**Discussion**

We found the same results as Rhodes et. al. (2012) that adults essentialize/generalize more when generics are used compared to specifics. However, unlike Rhodes et. al. (2012), and Rhodes, Leslie, Saunders, et al. (2018), this result did not hold for the children we studied when powering for a large effect size (the same power level of past research). We consider possible explanations for this in the general discussion. We also found not only did the high proportion quantifiers “most” and “many” increase essentializing/generalizing for adults compared to specifics, as well as “some”, they also did so to a greater extent than generics. Therefore, high proportion quantifiers “most” and “many” seem to be more powerful than generics in leading adults to essentialize/generalize about novel social kinds.

When we broke questions down to separate essentializing and generalizing types, Study 1 found adults essentialized more when generics, “many” or “most” were used compared to specifics, and when generics or “many” were used compared to “some”, for induction of novel property questions only. There were no differences between generics, “many” or “most” nor between specifics and “some.” Therefore, our study seems to show language signalling high proportions increases essentializing for novel social kinds. It is not the case, in this study, that generics are unique in having this effect. Furthermore, specifics alone do not appear to protect against essentializing. Interestingly, in our study, there were no differences for children (when powering for a large effect size), who had very low essentializing rates overall, a result which would seem to indicate children in our sample did not tend to essentialize novel social kinds, regardless of the type of language used, similar to another recent research finding on 4- and 5-year-olds (Noyes & Keil, 2020).

Study 1 also found adults generalized more when “many” or “most” were used compared to generics, suggesting that, if anything, high proportion quantifiers are better suited to learning about novel social kinds than generics. Generics, “many,” and “most” also led to more generalizing than specifics and “some.” In the case of children in our study, again, language did not affect how much they generalized novel social kinds (when powering for a large effect size). Altogether, our results suggest our finding that “many” and “most” lead to more essentializing/generalizing than generics were led by generalizing, not essentializing, further emphasizing the need to treat these constructs separately.

**Study 2**

The goal of Study 2 was to determine whether only *linguistic descriptions* indicating high proportions increases essentializing, or whether this also happens through *visual representations* of multiple instances. Participants were exposed to one of five stories. We again included the original *Generics* and *This* (specifics) conditions. We added three new conditions to determine whether linguistic descriptions, or the representation of multiple instances, drove Study 1’s results. A label-only condition in which the same pictures were used, but the text only read “Zarpie” on each page, with an arrow pointing to the Zarpie, was included as a visual marker of a single instance. A *This* condition, in which three different Zarpies did the same actions, was included to indicate multiple instances, despite the language indicating one instance each time. Finally, a label-only condition in which three different Zarpies did the same actions was included to cumulatively indicate multiple instances.

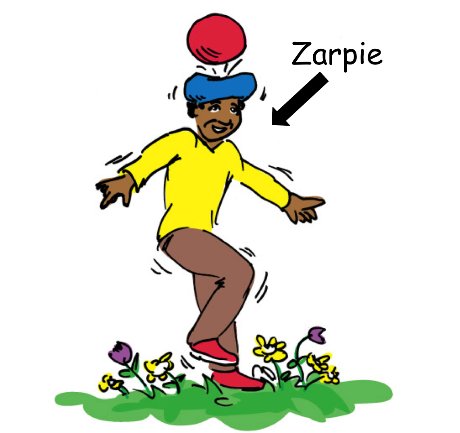
**Method**

**Participants.** Sample size was determined as in Study 1. There were 112 adults: 67 males and 45 females (M age = 31.45 years, SD = 10.34, Range = 18-67). All participants spoke English as a first language, while four additional participants who did not speak English as a first language were excluded. Each participant was randomly assigned to each of five conditions. Participants lived in the United Kingdom (51), the United States of America (35), or other countries (26). Participants were recruited through Prolific Academic, and were paid one British pound in the *Generic, This*, and *Label-only* conditions (for less than 10 minutes work), and two British pounds in the *ThisX3* and *Label-onlyX3* conditions (for less than 20 minutes works, since the book was three times as long).

There were 100 children: 50 males and 50 females (M age = 70 months, 27 days SD = 5 months, 17 days; Range = 60 months, 26 days – 85 months 6 days). Each participant was randomly assigned to each of five conditions. One additional participant did not answer any questions, so was excluded. All participants spoke English as a first language and all lived in the United Kingdom. Participants were White (28), Asian (2), of mixed ethnicity (1), or ethnicity information was not reported (68). Parents had a postgraduate degree (12), Undergraduate degree (11), High school diploma (8), or education information was not reported (68). Participants were recruited and paid as in Study 1. Parents signed a written consent form.

**Materials.** Five picture books were created using the characters Zarpies. The original *Generics* and *This* condition picture books from Study 1 were used. Three new picture books were used for Study 2: one with 16 illustrated pages (*Label-only*), and the other two with 48 illustrated pages (*ThisX3*; *Label-onlyX3*). Novel illustrations were created for the 48-page books matching the style of the original books (Rhodes et al., 2012).

The five conditions and hence separate books were: *Generics*, *ThisX3*, *Label-onlyX3*, *This*, and *Label-only* (see Figure 3 for an example from the *Label-onlyX3* book). The wording of each book was modified to match the condition. The *Generics* and *This* conditions were the same as Study 1. In the *ThisX3* condition, the book had the same sentences as in the *This* condition, repeated three times (following Roberts, et al., 2017), and each page had a different Zarpie doing the same thing. In the *Label-only* condition, the narrative of the original *This*/*Generics* book was replaced by the word “Zarpie” with an arrow pointing at the Zarpie character, but no additional text was given to describe what the Zarpie was doing. Finally, the *Label-onlyX3* condition was similar to the *Label-only* condition, but there were three pictures of different Zarpies doing the same thing. An Olympus MP3 recorder was used for all trials to record children’s responses for coding. The test questions were the same as Study 1.

**** **** 

*Figure 3*. Example from the *Label-onlyX3* book. Participants saw pages one by one, all in a row. Children were simply told “Zarpie” while the experimenter pointed to each person.

**Design.** The experiment was a between-subjects design. The independent variables were the book type (*Generics, ThisX3, Label-onlyX3, This, Label-only*), and age group (Adult, Child). The dependent variables were the same as Study 1.

**Procedure.** Same as Study 1. When testing children in both *Label-only* conditions, the experimenter pointed where the arrow pointed (at the Zarpie character), and said “Zarpie.”

**Coding.** Same as Study 1. Again, out of a possible 848 explanations, only eight suggested trait-based essentializing (e.g., “Is it because he's naughty”), and none suggested inherent explanations. Additionally, these responses were spread across conditions: *Generics* (0); *ThisX3* (0); *Label-onlyX3* (3); *This* (3); *Label-only* (2). Therefore, this measure was dropped from analyses. Missing items were replaced using maximum likelihood estimates. The nine remaining essentializing items did not show good reliability, *KR(20)* = .67 (Kline, 2000). We therefore tested reliability for inheritance items and items involving induction of novel properties separately. The three inheritance items did not show good reliability, *KR(20)* = .61, therefore these items were not analysed. Interestingly, these items only showed Alpha Cronbach scores of .62 in Rhodes et al. (2012), suggesting they perhaps should not have been included in that study either. The six items involving induction of novel properties showed good reliability, *KR(20)* = .78, therefore these items were analysed alone for essentializing. The results did not change whether the inheritance items were included or not. A second coder coded 47 participants’ (22%) explanation questions for generalizing. Agreement was excellent, *Cohen’s kappa* = 0.91. The 10 generalizing items showed good reliability *KR(20)* = .77. We still used “Because he likes to” type responses in our first analysis to see how our results mirror the original Rhodes et al. (2012) study. A second coder coded 45 participants’ (21%) explanation questions for essentializing (including “He likes to” type responses). Agreement was acceptable, *Cohen’s kappa* = 0.70. The 23 original Rhodes et al. (2012) essentializing and generalizing items showed good reliability *KR(20)* = .80.

**Results**

Children’s scores were skewed, therefore we could not use ANOVA. Therefore, LMEM was used with participant code and item as random variables.

Figure 4 displays the mean percentage of trials and 95% confidence intervals for essentialized/generalized responses by book type and age group. Participants did not give an answer that could be coded (as essentializing/generalizing, or not essentializing/generalizing) for 4.5% of trials. Additionally, 22% of children’s explanation responses could not be coded for generalizing, of which 50% of answers did not allow us to determine the proportion (e.g., it is unclear whether “To be high” refers to a high or low proportion of Zarpies), and 50% were due to children not responding or saying e.g., “I don’t know.” Furthermore, 13% of children’s explanation responses could not be coded for essentializing, of which 86% were due to children not responding or saying e.g., “I don’t know”, 8% were because we could not determine whether the answer was essentialist or not, e.g., “Because it’s got blue hair”, 4% were because children did not actually answer the question, e.g., “Because he”, and 2% were inaudible. The model for adults and children together was improved by age group (*χ2*(1) = 62.31, *p* < .0001), book type (*χ2*(4) = 29.71, *p* < .0001), and an interaction of age group and book type (*χ2*(4) = 12.00, *p* = .01735). The final model (*loglik* = -2317.5, *N* = 4659) found adults essentialized/generalized significantly more often than children (*OR* = 2.59, *p* = .0018). Due to the interaction of age and book type, we ran separate analyses for adults and children.

*Figure 4*. Percentage of essentialized/generalized responses by book type and age group. Error bars represent 95% confidence intervals.

Book type improved the model for adults (*χ2*(4) = 32.63, *p* < .0001). For pairwise comparisons, we used Bonferroni corrections. The final model (*loglik* = -1332.2, *N* = 2548) found adults essentialized/generalized significantly more often in the *Generics* condition than all other conditions (all *OR* > 2.82, *p* < .0006). No other differences were found. Adults were equally likely togeneralize/essentialize or not in the *Generics* and *ThisX3* conditions (both *p* > .05). Adults were significantly more likely to *not* generalize/essentialize in all other conditions (all *OR* > 2.97, *p* < .0180).

Book type did not improve the model for children. When analysing data collapsed across book types, the final model (*loglik* = -910.5, *N* = 2111) found children were more likely to *not* generalize/essentialize than engage in generalizing/essentializing (*OR* = 6.83, *p* < .0001). We also examined whether children were more likely to not respond or say e.g., “I don’t know” in the non-verbal conditions. However, descriptive statistics showed no clear pattern, with children giving a non-responses as follows: Generics (13%), Label-onlyX3 (20%), ThisX3 (8%), Label-only (9%), and This (5%), such that the Generics condition had the second highest percentage of non-responses, above that of the Label-only condition, both of which were relatively low.

Figure 5 displays the mean percentage of trials and 95% confidence intervals for essentialized responses by book type and age group for induction of novel property questions only (since inheritance questions were not reliable, although including them did not change the results). Participants did not give an answer that could be coded (as essentializing, or not essentializing) for 1.4% of trials. Age group improved the model (*χ2*(1) = 52.63, *p* < .0001). However, there was no effect of book type, nor an interaction between age group and book type. The final model (*loglik* = 595.6, *N* = 1260) found adults essentialized significantly more often than children (*OR* = 11.13, *p* < .0001). Participants essentialized and did not essentialize equally overall, *p* = .162. Running adults’ and children’s data separately did not lead to significant results.

Figure 5 displays the mean percentage of trials and 95% confidence intervals for generalized responses by book type, question type, and age group. Participants did not give an answer that could be coded (as generalizing, or not generalizing) for 11.5% of trials. LMEM was used with participant code and item as random variables. The model for adults and children together was improved by age group (*χ2*(1) = 83.94, *p* < .0001), question type (*χ2*(1) = 14.31, *p* < .0001), an interaction of age group and question type (*χ2*(1) = 5.54, *p* = .0186), book type (*χ2*(4) = 26.48, *p* < .0001), an interaction of question type and book type (*χ2*(4) = 69356 *p* < .0001), an interaction of age group and book type (*χ2*(4) = 11.51, *p* = .0214), and a 3-way interaction between age group, question type, and book type (*χ2*(4) = 10.24, *p* = .0366; final model: *loglik* = -739.7, *N* = 1978). Due to the interactions, we ran separate analyses for each age group.

*Figure 5*. Percentage of essentialized (novel property items only) and generalized responses, by book type and age group. Error bars represent 95% confidence intervals.

The model for adults was improved by question type (*χ2*(1) = 9.71, *p* = .0018), book type (*χ2*(4) = 37.67, *p* < .0001), and an interaction between question type and book type (*χ2*(4) = 24.32, *p* < .0001).The final model (*loglik* = -439.4, *N* = 1111) found adults generalize significantly more often for induction of familiar property questions than explanation questions (*OR* = 5.39 *p* < .0001). Pairwise comparisons using Bonferroni corrections found adults generalized significantly more often in the *Generics* condition than all other conditions (all *OR* > 3.58, *p* < .0179). Due to the interaction, we ran separate analyses for each question type.

Book type improved the model for adults’ induction of familiar property questions (*χ2*(4) = 16.13, *p* = .0029, *loglik* = -336.0, *N* = 672). Pairwise comparisons using Bonferroni corrections found adults generalized significantly more often in the *Generics* condition than the *Label-onlyX3, Label-only*, and *This* conditions (all *OR* > 8.35, *p* < .0012), but not the *ThisX3* condition. No other differences were found. Adults were significantly more likely to generalize than not in the *Generics* condition (*OR* = 13.13 *p* = .0390). Adults were equally likely to generalize and not generalize in all other conditions (all *p* > .5370).

Book type improved the model for adults’ explanation items (*χ2*(4) = 12.18, *p* = .0161, *loglik* = -77.3, *N* = 439). Pairwise comparisons using Bonferroni corrections found adults generalized significantly more often in the *Generics* condition than all other conditions (all *OR* > 371.30, *p* < .0015). No other differences were found. Adults were equally likely to generalize and not generalize in the *Generics* condition (*p* = .4640). Adults were significantly more likely to *not* generalize in all other conditions (all *OR* > 29.36, *p* < .0001).

The model for children’s generalizing was improved by question type (*χ2*(1) = 11.13, *p* = .0008), and an interaction of question type and book type (*χ2*(8) = 47.56, *p* < .0001; final model: *loglik* = -264.9, *N* = 867). Pairwise comparisons using Bonferroni corrections found no significant differences across conditions. Due to the interaction, we ran separate analyses for each question type.

Book type improved neither the models for children’s induction of familiar property questions, nor for children’s explanation items. The final models (familiar property: *loglik* = -235.1, *N* = 541; explanation: *loglik* = -32.3, *N* = 326), found children were significantly more likely to *not* generalize overall (both *OR* > 7.90, *p* < .0039).

**Collapsing Data over Studies 1 and 2**

We collapsed the children’s data from Studies 1 and 2 for the *Generics* and *This* conditions only to determine whether we could obtain the same results as Rhodes et al. (2012) with a smaller effect size. A G-Power analysis found 80 participants would accommodate a medium to large effect size (*Cohen’s d* = 0.64) with power at 0.80 and alpha at 0.05 (Faul et al., 2007). Book type did not improve the model. The final model (*loglik* = -784.1, *N* = 1652), found children were significantly more likely to *not* essentialize/generalize overall (*OR* = 4.56, *p* < .0001).

Since we had *N* = 31 British adult participants in each of the *Generics* and *This* conditions across Studies 1 and 2, as well as *N* = 12 American adult participants in the *Generics* condition, and *N* = 15 American adult participants in the *This* condition, we examined whether there might be general cultural differences in adults using the original (Rhodes, Leslie, Bianchi, et al., 2018; Rhodes et al., 2012) coding scheme. The model for essentializing/generalizing was improved by language type, (*χ2*(1) = 25.72, *p* < .0001), but not country, or an interaction of language type and country. The final model (*loglik* = -1148.8, *N* = 1998), found British and American adults were significantly more likely to generalize/essentialize in the *Generics* than *This* condition (*OR* = 2.92, *p* < .0001). British and American adults were equally likely to generalize/essentialize as not across conditions (*p* > .05).

**Discussion**

We again found the same results as Rhodes et. al. (2012), that adults essentialize/generalize more when generics are used compared to specifics. However, again, unlike Rhodes et. al. (2012), and Rhodes, Leslie, Saunders, et al. (2018), children in our sample did not seem to essentialize more when generics were used as compared with specifics (when powering for a large effect size). With adults, we found verbal descriptions of multiple instances are important to essentializing/generalizing, and multiple visual representations do not have the same impact.

Study 2 found neither adults nor children in our sample essentialized more when generics or visual representations of multiple instances were used compared to specifics or visual representations of single instances (when powering for a large effect size). This study shows visual representations of multiple instances do not lead to more essentializing than visual representations of single instances. This study also indicates specifics alone are not protective against essentializing. The finding in Study 1 that adults essentialize more (at least in the context of novel properties) for generics versus specifics was not replicated. As in Study 1, children had very low essentializing rates overall, suggesting the children in our sample do not tend to essentialize novel social kinds, regardless of the type of language or images used.

Study 2 also found adults generalized more when generics were used compared to all other conditions, except when specifics were repeated three times in the contexts of extending familiar properties. Therefore, specifics may lose some of their protective ability, compared to generics, when repeated only three times. This suggests that when specifics are repeated, they may sometimes be as effective as generics in increasing generalizations. In the case of the children in our sample, again, language and images did not affect how much they generalized novel social kinds (when powering for a large effect size).

When collapsing our generics and specifics data across children from Studies 1 and 2, increasing our power did not change our results using the original coding scheme (Rhodes, Leslie, Bianchi, et al., 2018; Rhodes et al., 2012), suggesting a very large sample is needed to find an effect. However, looking at the graphs across Studies 1 and 2, there is no clear pattern in any direction, suggesting even larger samples would not necessarily lead to a significant result with our sample. Furthermore, we examined whether there might be a general cultural difference in adults by collapsing our British and American samples across Studies 1 and 2 for the generics and specific conditions. We found no difference across countries, suggesting that while the British 5- and 6-year-olds in our sample do not generalize/essentialize as American children do, adults behave in a similar way. Therefore, while we do not pinpoint exactly when British children shift to using generics and specifics to make judgments about groups of people, we know there is a general shift from childhood to adulthood.

**General Discussion**

Our results provide evidence that generics may not play a special role in essentializing novel social kinds, for either adults or children. Study 1 found adults essentialized more about novel social kinds when generics, “most,” and “many” were used, compared to specifics, and when generics or “many” were used, compared to “some.” Therefore, for novel social kinds, high proportion quantifiers increased essentializing as much as generics did. Children in our sample, surprisingly, did not distinguish conditions (when powering for a large effect size). Furthermore, neither adults nor children in our sample distinguished conditions in terms of essentializing in Study 2, such that the findings comparing the generics and specifics conditions were not the same as Study 1 (again, when powering for a large effect size).

The results of several previous studies indicate generics lead children to essentialize familiar animals (Cimpian & Markman, 2009; Cimpian & Scott, 2012), novel artefacts (Cimpian & Cadena, 2010), familiar social kinds (Cimpian & Erickson, 2012), and lead children and adults to essentialize novel social kinds (Rhodes, Leslie, Bianchi, et al., 2018; Rhodes, Leslie, Saunders, et al., 2018; Rhodes et al., 2012). Our results for essentializing about novel social kinds appear weaker than past research (Rhodes, Leslie, Bianchi, et al., 2018; Rhodes, Leslie, Saunders, et al., 2018; Rhodes et al., 2012), especially for children. Importantly, we measured essentializing differently from previous studies (see introduction). One possibility is that previous results were led by participants’ generalizing rather than essentializing responses. Our results for adults are in keeping with this hypothesis: Study 1 found generics, and language signalling high proportions, significantly increased generalizing for adults, and Study 2 found generics significantly increased generalizing compared to specifics and visual imagery. However, surprisingly, neither study found generics increased generalizing for children. Unexpectedly, our results do not build on the fairly strong consensus that generics lead to essentializing and generalizing for children.

When analysing data in the same way as Rhodes et al. (2012) and Rhodes, Leslie, Saunders, et al. (2018), we confirmed the findings that generics led to greater essentializing/generalizing (combined) than specifics for adults, but this was not true for the children in our study. We also found lower levels of essentializing/generalizing in children overall than previous research, consistent with recent research findings on 4- and 5-year-olds (Noyes & Keil, 2020). One reason for this may be that the children in our study were 5- and 6-year-olds, however other studies also tested 5- and 6-year-olds (Rhodes, Leslie, Saunders, et al., 2018), suggesting the results are unlikely age-specific. A second possibility is that our sample was powered for a large effect size, and the effect may be smaller. However, past research was also powered for a large effect size (Rhodes, Leslie, Bianchi, et al., 2018; Rhodes et al., 2012), and the graphs in both studies suggest children do not essentialize/generalize more in the generics than specifics conditions. A third possibility is that the difference is cultural in nature. The original studies tested American children, while the current studies tested British children. Our data suggest no differences between American and British adults. However, this could be due to developmental cultural differences in how children learn; cognitive, linguistic, or social skills; their understanding of morality; or a variety of other reasons. This is not the only domain in which British children appear to develop differently from North American children. For instance, British children lag behind American children in terms of language development (Hamilton, Plunkett, & Schafer, 2000). Cross-linguistic study of the acquisition path of quantifiers at age 5 indicates that along four primary dimensions of the meaning and use of quantifiers, there are robust similarities in the order of acquisition of quantifiers across 31 languages (Katsos et al., 2016), therefore it is unlikely quantifiers held British children back. However, British children may be worse at understanding and using generic language as a grouping mechanism, and specific language as an individuating mechanism. Furthermore, British children peek and lie about peeking much less often than North American children, which could be driven by differences in moral development (Alloway, McCallum, Alloway, & Hoicka, 2015). Perhaps British children see essentializing social kinds as a moral issue in a way that American children do not. Still further, there are differences in findings in American and British children’s ability to infer causal relations from patterns of events (McCormack, Butterfill, Hoerl, & Burns, 2009; Sobel, Tenenbaum, & Gopnik, 2004). Future research should investigate whether the results we have found truly are due to cultural developmental differences, and if so, what factors may affect such cultural differences.

**High Proportion Quantifiers and Generics Act Alike**

Our findings suggest when adults essentialize novel social kinds, this is not due to generics *per se* but to language signalling high proportions. Study 1 therefore demonstrates the importance of using appropriate controls when determining how language affects social cognition. We suggest generics may not be a special type of language especially suited to serve this function for novel social kinds, in either adults or children.

Cimpian and Erickson (2012) found children’s causal attributions were more essentialized when they learned novel information about known social kinds (e.g., girls), whether generics or “most” was used. Our Study 1 results converge with these findings. As discussed in the introduction, if adults were essentializing from “many” and “most” via essentialist beliefs, we would still only expect adults to essentialize from high proportion quantifiers at a fraction of the rate that they do so when generics are used based on their generic-encoding error rate for “many” and “most.” However in our study, and the Cimpian and Erickson study, there was no difference in the levels of essentializing between generics and high proportion quantifiers, even though only 13% of adults’ explanation responses in the Most condition, and 9% in the Many condition, involved generics in our study (versus 46% in the Generics condition). If encoding high proportion quantifiers as generics is not an important step to essentializing, it is unclear why generic beliefs are a necessary step.

If instead the argument is that high proportion quantifiers always lead to generic beliefs, even if the high proportion quantifier is remembered as such, then it is not clear why the belief should be labelled “generic.” Rather, perhaps signalling high proportions, through high proportion quantifiers or generics, leads directly to stronger concept formation more generally, and hence, essentializing.

Past research also found generics help children generalize novel information about familiar animals and artefacts (Cimpian & Markman, 2008; Gelman et al., 2002; Nguyen & Gelman, 2012); and novel animals and artefacts (Chambers et al., 2008; Gelman & Bloom, 2007; Gelman et al., 2010; Graham et al., 2016; Graham et al., 2011; Hollander et al., 2009; Stock et al., 2009). However, our control measure in Study 1 found language indicating high proportions, including “most” and “many,” serve this function even better than generics for adults, whereas none of these types of language serve this function for children in our sample. Therefore, even in the case of generalizing information, our results suggest generics are not unique in encouraging generalizing in our sample.

While high proportion quantifiers do lead adults to essentialize as much as generics about novel social kinds, one interesting point is that generics may still be the default language to express essentialism. Rhodes et al. (2012) found parents used significantly more generics after they were induced to have essentialist beliefs, and this was not the case for quantified language. However, they did not distinguish between high, low, and indefinite proportion quantifiers, making it unclear whether it was possible parents also increase high proportion quantifiers when essentialist beliefs are induced. Future research should examine this possibility. Furthermore, a recent naturalistic study based on 26 CHILDES corpora found adults and children actually used more generics in reference to less essentialized categories (inanimates) than more essentialized categories (animates) (Mehrotra & Perfors, 2019), suggesting generics may not be the default language to express essentialism after all.

A final consideration is that it may not be language signalling high proportions alone that leads to essentializing. Noyes and Keil (2020) recently found that from 6 years children and adults essentialized novel social kinds when generics were paired with biological information (“Vawnsies feel sick when they drink milk”), but not when paired with cultural information (“Vawnsies believe that fish talk to God”), suggesting the contextual information attached to the generic matters. Four- to 5-year-olds did not distinguish these contexts, tending *not* to essentialize either context, suggesting understanding internal explanations could be important to using generics. Future research should consider the combination of high proportion quantifiers and contextual information on essentializing judgments.

**Are Specifics Protective?**

Past research suggested generics increase essentializing (Cimpian, 2010; Rhodes et al., 2012), but it was also compatible with the claim that instead, specifics protect against essentializing (Rhodes, Leslie, Bianchi, et al., 2018). Our results provide some evidence this is not the case. In Study 1, specifics were no better than the quantifier “some” at preventing essentializing in adults or children: “some” is not a specific but rather expresses an uncertain quantity similar to “at least one” which is compatible with “all”, and though it often pragmatically implicates “not all” (Katsos & Bishop, 2011; Skordos & Papafragou, 2016), this is again not specific and is compatible with e.g., “most.” Furthermore, in Study 2, specifics were no better than the label-only conditions in preventing essentializing. Our results were the same for our control measure, generalizing. Additionally, there was no difference in how much adults generalized in terms of extending familiar properties to novel individuals in Study 2 when generics or repeated specifics were used, suggesting repeating specifics may inhibit any potential protective mechanism anyway. Therefore, the support for specifics being protective against essentializing or generalizing is poor. Thus, perhaps the default is to avoid essentializing or generalizing unless language indicating high proportions is used.**Acknowledgments**

We thank children and schools for participating, Sarah-Jane Leslie for sharing illustrations, Ray Drainville for creating novel illustrations, and the Centre for XXXX at the University of XXXX for project funding.

Funding: This work was supported by grants from the Centre for XXXX at the University of XXXX, to XXXX and XXXX.

**References**

Alloway, T. P., McCallum, F., Alloway, R. G., & Hoicka, E. (2015). Liar, liar, working memory on fire: Investigating the role of working memory in childhood verbal deception. *Journal of Experimental Child Psychology, 137*, 30-38. doi:10.1016/j.jecp.2015.03.013

Béland, S., Pichette, F., & Jolani, S. (2016). Impact on Cronbach’s α of simple treatment methods for missing data. *The Quantitative Methods for Psychology, 12*(1), 57-73. doi:10.20982/tqmp.12.1.p057

Brandone, A. C., Gelman, S. A., & Hedglen, J. (2015). Children's developing intuitions about the truth conditions and implications of novel generics versus quantified statements. *Cognitive Science, 39*(4), 711-738. doi:10.1111/cogs.12176

Carlson, G. N. (1977). *Reference to kinds in English.* University of Massachussetts, Amherst.,

Chambers, C. G., Graham, S. A., & Turner, J. N. (2008). When hearsay trumps evidence: How generic language guides preschoolers' inferences about unfamiliar things. *Language and Cognitive Processes, 23*(5), 749-766. doi:10.1080/01690960701786111

Cimpian, A. (2010). The impact of generic language about ability on children's achievement motivation. *Developmental Psychology, 46*(5), 1333-1340. doi:10.1037/a0019665

Cimpian, A., & Cadena, C. (2010). Why are dunkels sticky? Preschoolers infer functionality and intentional creation for artifact properties learned from generic language. *Cognition, 117*(1), 62-68. doi:10.1016/j.cognition.2010.06.011

Cimpian, A., & Erickson, L. C. (2012). The effect of generic statements on children's causal attributions: Questions of mechanism. *Developmental Psychology, 48*(1), 159-170. doi:10.1037/a0025274

Cimpian, A., Gelman, S. A., & Brandone, A. C. (2010). Theory-based considerations influence the interpretation of generic sentences. *Language and Cognitive Processes, 25*(2), 261-276. doi:10.1080/01690960903025227

Cimpian, A., & Markman, E. M. (2008). Preschool children's use of cues to generic meaning. *Cognition, 107*(1), 19-53. doi:10.1016/j.cognition.2007.07.008

Cimpian, A., & Markman, E. M. (2009). Information learned from generic language becomes central to children's biological concepts: Evidence from their open-ended explanations. *Cognition, 113*(1), 14-25. doi:10.1016/j.cognition.2009.07.004

Cimpian, A., & Markman, E. M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child Development, 82*(2), 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*(3), 419-433. doi:10.1016/j.cognition.2012.02.003

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of “minimal” group affiliations in children. *Child Development, 82*(3), 793-811. doi:10.1111/j.1467-8624.2011.01577.x

Fassbender, I., Teubert, M., & Lohaus, A. (2016). The development of preferences for own-race versus other-race faces in 3-, 6- and 9-month-old Caucasian infants. *European Journal of Developmental Psychology, 13*(1), 152-165. doi:10.1080/17405629.2015.1073585

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. doi:10.3758/bf03193146

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*(9), 404-409. doi:10.1016/j.tics.2004.07.001

Gelman, S. A., & Bloom, P. (2007). Developmental changes in the understanding of generics. *Cognition, 105*(1), 166-183. doi:10.1016/j.cognition.2006.09.009

Gelman, S. A., Leslie, S.-J., Gelman, R., & Leslie, A. (2019). Do children recall numbers as generic? A strong test of the generics-as-default hypothesis. *Language Learning and Development, 15*(3), 217-231. doi:10.1080/15475441.2019.1571418

Gelman, S. A., Leslie, S.-J., Was, A. M., & Koch, C. M. (2015). Children's interpretations of general quantifiers, specific quantifiers and generics. *Language, Cognition and Neuroscience, 30*(4), 448-461. doi:10.1080/23273798.2014.931591

Gelman, S. A., & Raman, L. (2003). Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development, 74*(1), 308-325. doi:10.1111/1467-8624.00537

Gelman, S. A., Sanchez Tapia, I., & Leslie, S.-J. (2016). Memory for generic and quantified sentences in Spanish-speaking children and adults. *Journal of Child Language, 43*(6), 1231-1244. doi:10.1017/S0305000915000483

Gelman, S. A., Star, J. R., & Flukes, J. E. (2002). Children's use of generics in inductive inferences. *Journal of Cognition and Development, 3*(2), 179-199. doi:10.1207/S15327647JCD0302\_3

Gelman, S. A., & Tardif, T. (1998). A cross-linguistic comparison of generic noun phrases in English and Mandarin. *Cognition, 66*(3), 215-248. doi:10.1016/S0010-0277%2898%2900021-3

Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of generic language on category content and structure. *Cognitive Psychology, 61*(3), 273-301. doi:10.1016/j.cogpsych.2010.06.001

Goldfarb, D., Lagattuta, K. H., Kramer, H. J., Kennedy, K., & Tashjian, S. M. (2017). When your kind cannot live here: How generic language and criminal sanctions shape social categorization. *Psychological Science, 28*(11), 1597-1609. doi:10.1177/0956797617714827

Graham, S. A., Gelman, S. A., & Clarke, J. (2016). Generics license 30-month-olds’ inferences about the atypical properties of novel kinds. *Developmental Psychology, 52*(9), 1353–1362. doi:10.1037/dev0000183

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*(2), 493-507. doi:10.1111/j.1467-8624.2010.01572.x

Halberda, J., Taing, L., & Lidz, J. (2008). The development of “most” comprehension and its potential dependence on counting ability in preschoolers. *Language Learning and Development, 4*(2), 99-121. doi:10.1080/15475440801922099

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language, 27*(3), 689-705. doi:10.1017/s0305000900004414

Hoicka, E., & Akhtar, N. (2011). Preschoolers joke with jokers, but correct foreigners. *Developmental Science, 14*(4), 848-858. doi:10.1111/j.1467-7687.2010.01033.x

Hollander, M. A., Gelman, S. A., & Raman, L. (2009). Generic language and judgements about category membership: Can generics highlight properties as central? *Language and Cognitive Processes, 24*(4), 481-505. doi:10.1080/01690960802223485

Hollander, M. A., Gelman, S. A., & Star, J. (2002). Children's interpretation of generic noun phrases. *Developmental Psychology, 38*(6), 883-894. doi:10.1037/0012-1649.38.6.883

Katsos, N., & Bishop, D. V. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition,, 120*(1), 67-81. doi:10.1016/j.cognition.2011.02.015

Katsos, N., Cummins, C., Ezeizabarrena, M. J., Gavarró, A., Kraljević, J. K., Hrzica, G., . . . Van Hout, A. (2016). Cross-linguistic patterns in the acquisition of quantifiers. *Proceedings of the National Academy of Sciences, 113*(33), 9244-9249. doi:10.1073/pnas.1601341113

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., . . . Pascalis, O. (2005). Three‐month‐olds, but not newborns, prefer own‐race faces. *Developmental Science, 8*(6), F31-F36. doi:10.1111/j.1467-7687.2005.0434a.x.

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27*(4), 623-634. doi:10.1521/soco.2009.27.4.623

Kline, P. (2000). *Handbook of psychological testing* (2nd ed.). London: Routledge.

Lawler, J. M. (1973). Studies in English generics. *University of Michigan Papers in Linguistics, 1*(1), 1-184.

Leslie, S.-J. (2008). Generics: Cognition and acquisition. *Philosophical Review, 117*(1), 1-47. doi:10.1215/00318108-2007-023

Leslie, S.-J. (2017). The original sin of cognition: Fear, prejudice, and generalization. *The Journal of Philosophy, 114*(8), 393-421. doi:10.5840/jphil2017114828

Leslie, S.-J., & Gelman, S. A. (2012). Quantified statements are recalled as generics: Evidence from preschool children and adults. *Cognitive Psychology, 64*(3), 186-214. doi:10.1016/j.cogpsych.2011.12.001

Liu, S., Xiao, N. G., Quinn, P. C., Zhu, D., Ge, L., Pascalis, O., & Lee, K. (2015). Asian infants show preference for own-race but not other-race female faces: the role of infant caregiving arrangements. *Frontiers in Psychology Vol 6 2015, ArtID 593, 6*. doi:10.3389/fpsyg.2015.00593

Liu, S., Xiao, W. S., Xiao, N. G., Quinn, P. C., Zhang, Y., Chen, H., . . . Lee, K. (2015). Development of visual preference for own- versus other-race faces in infancy. *Developmental Psychology, 51*(4), 500-511. doi:10.1037/a0038835

McCormack, T., Butterfill, S., Hoerl, C., & Burns, P. (2009). Cue competition effects and young children’s causal and counterfactual inferences. *Developmental Psychology, 45*(6), 1563-1575. doi:10.1037/a0017408

Mehrotra, S., & Perfors, A. (2019). *Generic noun phrases in child speech.* Paper presented at the 41st Annual Conference of the Cognitive Science Society, Montreal.

Nguyen, S. P., & Gelman, A. (2012). Generic language facilitates children's cross-classification. *Cognitive Development, 27*(2), 154-167. doi:10.1016/j.cogdev.2012.01.001

Noyes, A., & Keil, F. C. (2020). There is no privileged link between kinds and essences early in development. *Proceedings of the National Academy of Sciences*, 1-3. doi:10.1073/pnas.2003627117

Nyhout, A., & O'Neill, D. K. (2014). Storybooks aren't just for fun: Narrative and non-narrative picture books foster equal amounts of generic language during mother-toddler book sharing. *Frontiers in Psychology Vol 5 2014, ArtID 325, 5*. doi:10.3389/fpsyg.2014.00325

Pappas, A., & Gelman, S. A. (1998). Generic noun phrases in mother-child conversations. *Journal of Child Language, 25*(1), 19-33. doi:10.1017/s0305000997003292

Prasada, S., Khemlani, S., Leslie, S.-J., & Glucksberg, S. (2013). Conceptual distinctions amongst generics. *Cognition, 126*(3), 405-422. doi:10.1016/j.cognition.2012.11.010

Rhodes, M., Leslie, S.-J., Bianchi, L., & Chalik, L. (2018). The role of generic language in the early development of social categorization. *Child Development, 89*(1), 148-155. doi:10.1111/cdev.12714

Rhodes, M., Leslie, S.-J., Saunders, K., Dunham, Y., & Cimpian, A. (2018). How does social essentialism affect the development of inter-group relations? *Developmental Science, 21*(1), e12509. 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, 109*(34), 13526-13531. doi:10.1073/pnas.1208951109

Roberts, S. O., Ho, A. K., & Gelman, S. A. (2017). Group presence, category labels, and generic statements influence children to treat descriptive group regularities as prescriptive. *Journal of Experimental Child Psychology, 158*, 19-31. doi:10.1016/j.jecp.2016.11.013

Saul, J. (2017). Are generics especially pernicious? *Inquiry*, 1-17. doi:10.1080/0020174X.2017.1285995

Skordos, D., & Papafragou, A. (2016). Children’s derivation of scalar implicatures: Alternatives and relevance. *Cognition, 153*(6-18). doi:10.1016/j.cognition.2016.04.006

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science, 28*(3), 303-333. doi:10.1207/s15516709cog2803\_1

Stock, H. R., Graham, S. A., & Chambers, C. G. (2009). Generic language and speaker confidence guide preschoolers' inferences about novel animate kinds. *Developmental Psychology, 45*(3), 884-888. doi:10.1037/a0015359

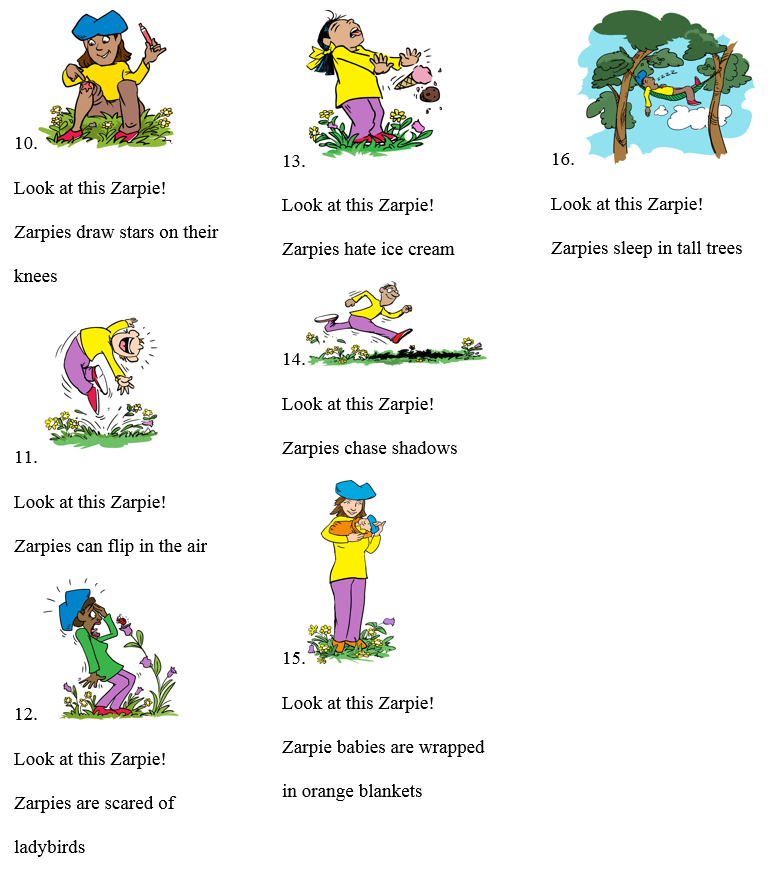
Wodak, D., Leslie, S.-J., & Rhodes, M. (2015). What a loaded generalization: Generics and social cognition. *Philosophy Compass, 10*(9), 625-635. doi:10.1111/phc3.12250

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2018). Older but not younger infants associate own-race faces with happy music and other-race faces with sad music. *Developmental Science, 21*(e12537), 1-10. doi:10.1111/desc.12537

Xiao, N. G., Wu, R., Quinn, P. C., Liu, S., Tummeltshammer, K. S., Kirkham, N. Z., . . . Lee, K. (2018). Infants rely more on gaze cues from own-race than other-race adults for learning under uncertainty. *Child Development, 89*(3), e229–e244. doi:10.1111/cdev.12798

**Appendix A**

Generics book. The same images were shown for all other books in Studies 1 and 2, however wording was different (see Appendix B).



**Appendix B**

Text for each condition in Studies 1 and 2. Images were the same for all conditions, except the ThisX3 and ArrowX3 conditions in Study 2 had three different images of different Zarpies doing the same thing.

Book text, by condition, Study 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Page | Generics | Most | Many | Some | This |
| 1 | Look at this Zarpie! Zarpies love to eat flowers. | Look at this Zarpie! Most Zarpies love to eat flowers. | Look at this Zarpie! Many Zarpies love to eat flowers. | Look at this Zarpie! Some Zarpies love to eat flowers. | Look at this Zarpie! This Zarpie loves to eat flowers. |
| 2 | Look at this Zarpie! Zarpies have stripes in their hair. | Look at this Zarpie! Most Zarpies have stripes in their hair. | Look at this Zarpie! Many Zarpies have stripes in their hair. | Look at this Zarpie! Some Zarpies have stripes in their hair. | Look at this Zarpie! This Zarpie has stripes in her hair. |
| 3 | Look at this Zarpie! Zarpies can bounce a ball on their heads. | Look at this Zarpie! Most Zarpies can bounce a ball on their heads. | Look at this Zarpie! Many Zarpies can bounce a ball on their heads. | Look at this Zarpie! Some Zarpies can bounce a ball on their heads. | Look at this Zarpie! This Zarpie can bounce a ball on his heads. |
| 4 | Look at this Zarpie! Zarpies like to sing. | Look at this Zarpie! Most Zarpies like to sing. | Look at this Zarpie! Many Zarpies like to sing. | Look at this Zarpie! Some Zarpies like to sing. | Look at this Zarpie! This Zarpie likes to sing. |
| 5 | Look at this Zarpie! Zarpies climb tall fences. | Look at this Zarpie! Most Zarpies climb tall fences. | Look at this Zarpie! Many Zarpies climb tall fences. | Look at this Zarpie! Some Zarpies climb tall fences. | Look at this Zarpie! This Zarpie climbs tall fences. |
| 6 | Look at this Zarpie! Zarpies flap their arms when they’re happy. | Look at this Zarpie! Most Zarpies flap their arms when they’re happy. | Look at this Zarpie! Many Zarpies flap their arms when they’re happy. | Look at this Zarpie! Some Zarpies flap their arms when they’re happy. | Look at this Zarpie! This Zarpie flaps her arms when she is happy. |
| 7 | Look at this Zarpie! Zarpies have freckles on their feet. | Look at this Zarpie! Most Zarpies have freckles on their feet. | Look at this Zarpie! Many Zarpies have freckles on their feet. | Look at this Zarpie! Some Zarpies have freckles on their feet. | Look at this Zarpie! This Zarpie has freckles on his feet. |
| 8 | Look at this Zarpie! Zarpies hop over puddles. | Look at this Zarpie! Most Zarpies hop over puddles. | Look at this Zarpie! Many Zarpies hop over puddles. | Look at this Zarpie! Some Zarpies hop over puddles. | Look at this Zarpie! This Zarpie hops over puddles. |
| 9 | Look at this Zarpie! Zarpies hate walking in the mud. | Look at this Zarpie! Most Zarpies hate walking in the mud. | Look at this Zarpie! Many Zarpies hate walking in the mud. | Look at this Zarpie! Some Zarpies hate walking in the mud. | Look at this Zarpie! This Zarpie hates walking in the mud. |
| 10 | Look at this Zarpie! Zarpies draw stars on their knees. | Look at this Zarpie! Most Zarpies draw stars on their knees. | Look at this Zarpie! Many Zarpies draw stars on their knees. | Look at this Zarpie! Some Zarpies draw stars on their knees. | Look at this Zarpie! This Zarpie draws stars on her knees. |
| 11 | Look at this Zarpie! Zarpies can flip in the air. | Look at this Zarpie! Most Zarpies can flip in the air. | Look at this Zarpie! Many Zarpies can flip in the air. | Look at this Zarpie! Some Zarpies can flip in the air. | Look at this Zarpie! This Zarpie can flip in the air. |
| 12 | Look at this Zarpie! Zarpies are scared of ladybirds. | Look at this Zarpie! Most Zarpies are scared of ladybirds. | Look at this Zarpie! Many Zarpies are scared of ladybirds. | Look at this Zarpie! Some Zarpies are scared of ladybirds. | Look at this Zarpie! This Zarpie is scared of ladybirds. |
| 13 | Look at this Zarpie! Zarpies hate ice cream | Look at this Zarpie! Most Zarpies hate ice cream | Look at this Zarpie! Many Zarpies hate ice cream | Look at this Zarpie! Some Zarpies hate ice cream | Look at this Zarpie! This Zarpie hates ice cream |
| 14 | Look at this Zarpie! Zarpies chase shadows. | Look at this Zarpie! Most Zarpies chase shadows. | Look at this Zarpie! Many Zarpies chase shadows. | Look at this Zarpie! Some Zarpies chase shadows. | Look at this Zarpie! This Zarpie chases shadows. |
| 15 | Look at this Zarpie! Zarpie babies are wrapped in orange blankets. | Look at this Zarpie! Most Zarpie babies are wrapped in orange blankets. | Look at this Zarpie! Many Zarpie babies are wrapped in orange blankets. | Look at this Zarpie! Some Zarpie babies are wrapped in orange blankets. | Look at this Zarpie! This Zarpie baby is wrapped in an orange blanket. |
| 16 | Look at this Zarpie! Zarpies sleep in tall trees. | Look at this Zarpie! Most Zarpies sleep in tall trees. | Look at this Zarpie! Many Zarpies sleep in tall trees. | Look at this Zarpie! Some Zarpies sleep in tall trees. | Look at this Zarpie! This Zarpie sleeps in tall trees. |

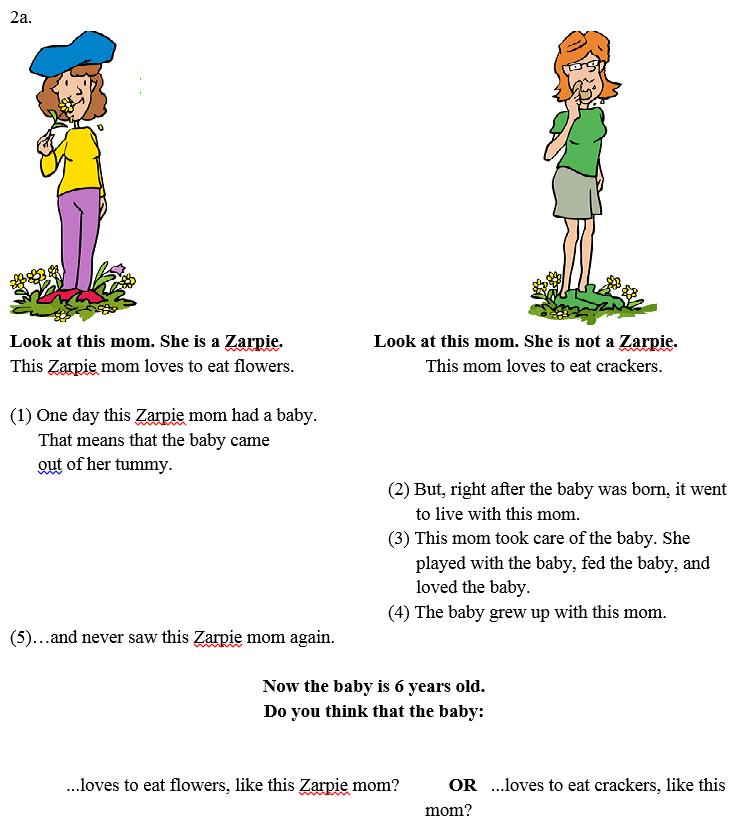
Book text, by condition, Study 2

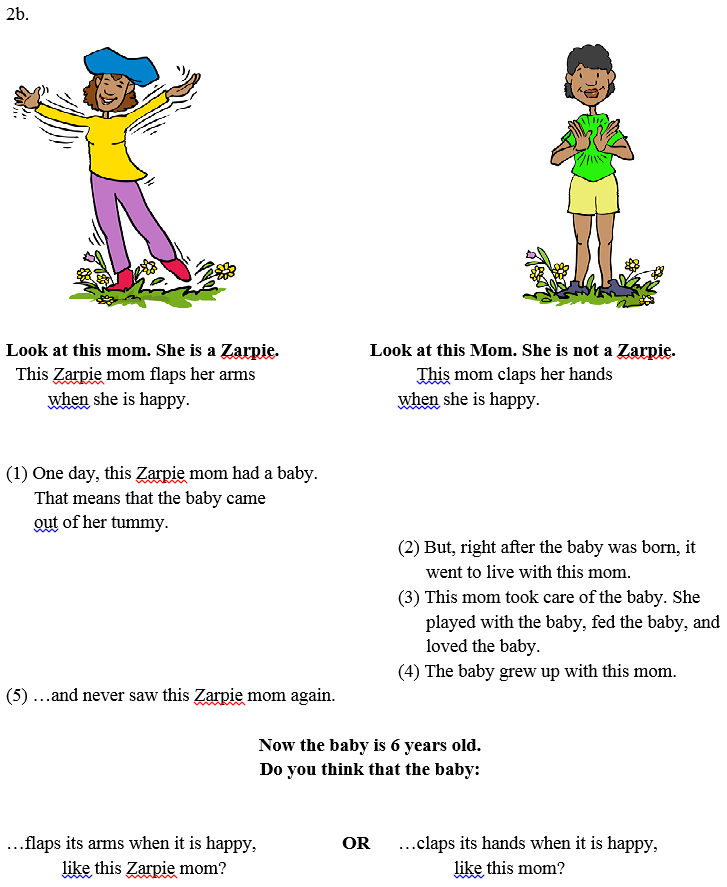
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Page | Generics | ThisX3 (repeated with each of 3 images) | This | ArrowX3  (repeated with each of 3 images) | Arrow |
| 1 | Look at this Zarpie! Zarpies love to eat flowers. | Look at this Zarpie! This Zarpie loves to eat flowers. | Look at this Zarpie! This Zarpie loves to eat flowers. | Zarpie | Zarpie |
| 2 | Look at this Zarpie! Zarpies have stripes in their hair. | Look at this Zarpie! This Zarpie has stripes in his/her hair. | Look at this Zarpie! This Zarpie has stripes in her hair. | Zarpie | Zarpie |
| 3 | Look at this Zarpie! Zarpies can bounce a ball on their heads. | Look at this Zarpie! This Zarpie can bounce a ball on his/her heads. | Look at this Zarpie! This Zarpie can bounce a ball on his heads. | Zarpie | Zarpie |
| 4 | Look at this Zarpie! Zarpies like to sing. | Look at this Zarpie! This Zarpie likes to sing. | Look at this Zarpie! This Zarpie likes to sing. | Zarpie | Zarpie |
| 5 | Look at this Zarpie! Zarpies climb tall fences. | Look at this Zarpie! This Zarpie climbs tall fences. | Look at this Zarpie! This Zarpie climbs tall fences. | Zarpie | Zarpie |
| 6 | Look at this Zarpie! Zarpies flap their arms when they’re happy. | Look at this Zarpie! This Zarpie flaps his/her arms when she is happy. | Look at this Zarpie! This Zarpie flaps her arms when she is happy. | Zarpie | Zarpie |
| 7 | Look at this Zarpie! Zarpies have freckles on their feet. | Look at this Zarpie! This Zarpie has freckles on his/her feet. | Look at this Zarpie! This Zarpie has freckles on his feet. | Zarpie | Zarpie |
| 8 | Look at this Zarpie! Zarpies hop over puddles. | Look at this Zarpie! This Zarpie hops over puddles. | Look at this Zarpie! This Zarpie hops over puddles. | Zarpie | Zarpie |
| 9 | Look at this Zarpie! Zarpies hate walking in the mud. | Look at this Zarpie! This Zarpie hates walking in the mud. | Look at this Zarpie! This Zarpie hates walking in the mud. | Zarpie | Zarpie |
| 10 | Look at this Zarpie! Zarpies draw stars on their knees. | Look at this Zarpie! This Zarpie draws stars on his/her knees. | Look at this Zarpie! This Zarpie draws stars on her knees. | Zarpie | Zarpie |
| 11 | Look at this Zarpie! Zarpies can flip in the air. | Look at this Zarpie! This Zarpie can flip in the air. | Look at this Zarpie! This Zarpie can flip in the air. | Zarpie | Zarpie |
| 12 | Look at this Zarpie! Zarpies are scared of ladybirds. | Look at this Zarpie! This Zarpie is scared of ladybirds. | Look at this Zarpie! This Zarpie is scared of ladybirds. | Zarpie | Zarpie |
| 13 | Look at this Zarpie! Zarpies hate ice cream | Look at this Zarpie! This Zarpie hates ice cream | Look at this Zarpie! This Zarpie hates ice cream | Zarpie | Zarpie |
| 14 | Look at this Zarpie! Zarpies chase shadows. | Look at this Zarpie! This Zarpie chases shadows. | Look at this Zarpie! This Zarpie chases shadows. | Zarpie | Zarpie |
| 15 | Look at this Zarpie! Zarpie babies are wrapped in orange blankets. | Look at this Zarpie! This Zarpie baby is wrapped in an orange blanket. | Look at this Zarpie! This Zarpie baby is wrapped in an orange blanket. | Zarpie | Zarpie |
| 16 | Look at this Zarpie! Zarpies sleep in tall trees. | Look at this Zarpie! This Zarpie sleeps in tall trees. | Look at this Zarpie! This Zarpie sleeps in tall trees. | Zarpie | Zarpie |

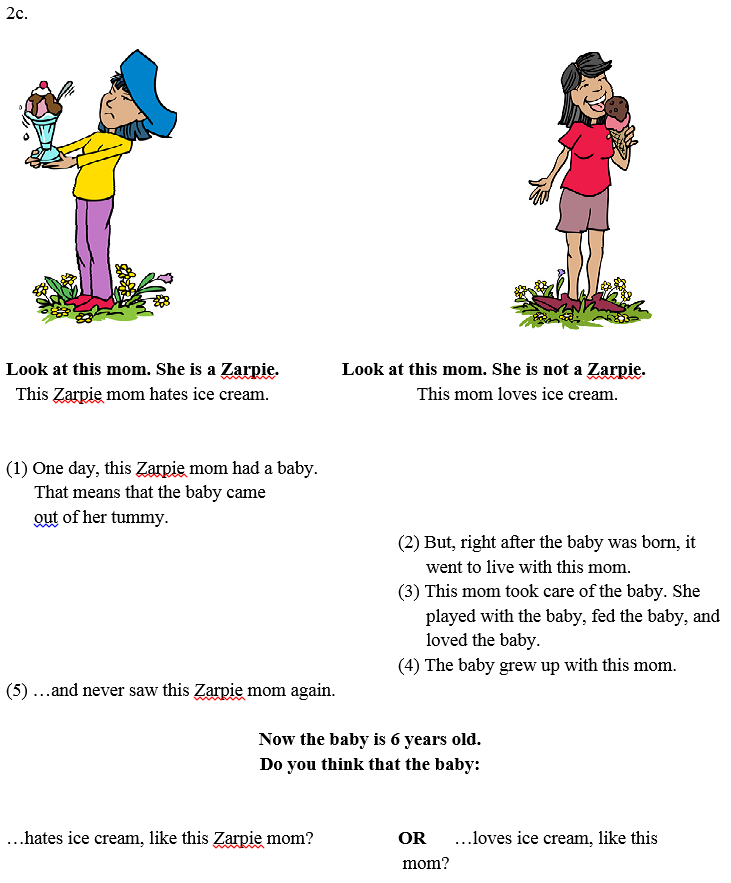
**Appendix C**

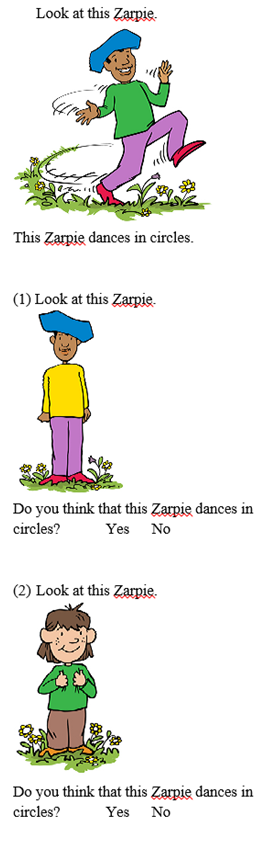
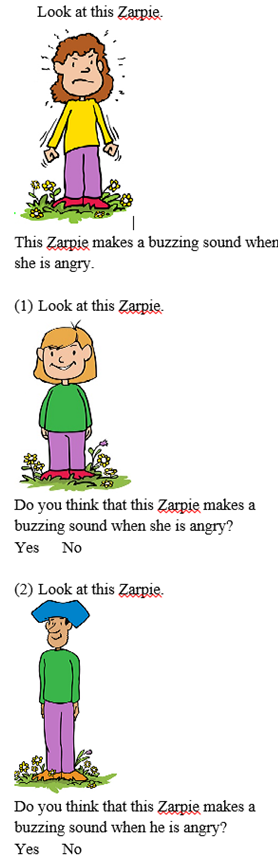
Questions for all conditions

**Essentializing: Inheritance Questions**



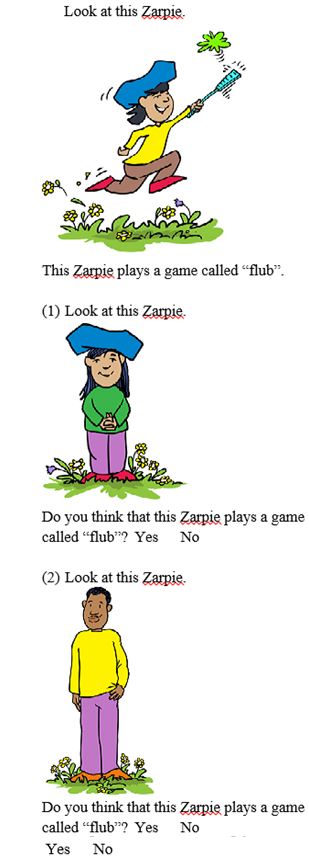




**Essentializing: Induction of Novel Properties Questions**

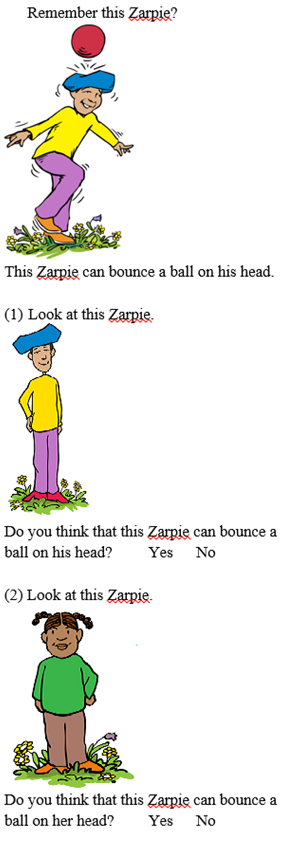
a.

b..v

****

c.v

**Generalizing: Induction of Familiar Properties Questions.**



c.v

**Generalizing: Explanation Questions (also used for Essentializing to replicate the original Rhodes, et al., 2012 study).**