Received: 1 April 2021 | Revised: 8 March 2022 | Accepted: 15 April 2022

DOI: 10.1111/sode.12609

ORIGINAL ARTICLE

WILEY

Children's inductive inferences about individuals with gender category uncertainty

Andrea C. Yuly-Youngblood[1] | **Janet J. Boseovski**[2]

[1]Department of Psychology, University of North Carolina at Greensboro, Greensboro, NC 27402-6170, USA

[2]Department of Psychology, University of North Carolina at Greensboro, Greensboro, NC 27402-6170, USA

Correspondence
Janet J. Boseovski, Department of Psychology, University of North Carolina at Greensboro, 296 Eberhart Building, PO Box 26170, Greensboro, NC 27402-6170, USA.
Email: jjboseov@uncg.edu

[Corrections made on 26 May 2022 after first online publication: The order of the names in the byline has corrected in this version.]

Abstract

Little research focuses on children's reasoning about people whose gender is perceived as uncertain. Five- to 8-year-olds viewed a target with a gender uncertain appearance. The target had trait or preference similarities with a character from a binary, specified gender (i.e., boy, girl) and appearance similarities with another character that had an uncertain gender. Half of the participants heard gender uncertainty labels (i.e., "We're not sure about this person. This person doesn't look like a boy or a girl.") or gender specification labels (i.e., "This person is a boy.") for each character. The remaining participants heard miscellaneous character information. Children inferred whether the target favored the same novel activity as the character with similar traits or preferences, but with a specified gender, or the gender uncertain character. Seven- and 8-year-olds made more trait-based predictions than 5- and 6-year-olds, but both age groups made unsystematic predictions in response to preferences, suggesting that some viewed preferences and others viewed gender as reflective of broader similarities among people in this context. A follow-up study with 5- and 6-year-olds conveyed gender uncertainty more directly through identification (i.e., "This person *is not* a boy or a girl.") and revealed that children consistently made preference-based, but not trait-based, predictions about the target. The present findings reveal that children do not reason about gender uncertainty in the same way that they reason about

the binary gender categories, which highlights the need to further investigate children's understanding of gender beyond the typical boy and girl dichotomy.

KEYWORDS
categorization, decision-making, diversity, gender, social cognition

1 | INTRODUCTION

Gender is a potent influence on children's social decision making (e.g., Ruble et al., 2006). For example, gender categories guide children's attitudes toward others, including how much they like and want to befriend people (e.g., Halim, 2016; Ruble et al., 2006). Importantly, preschoolers use gender categories inductively (Taylor et al., 2009): they infer that a boy with girl-like perceptual features shares properties with other boys (Gelman et al., 1986). Preschoolers also use category labels over other relevant information (i.e., traits) to make predictions about people (Diesendruck & haLevi, 2006). Similarly, young children prioritize gender categories over preferences to make inferences about others, although this dissipates in middle childhood (Martin, 1989).

The majority of developmental research centers on the binary gender categories (Dunham & Olson, 2016). To expand research beyond the binary gender categories, one can investigate how children react to individuals who *do not appear* to belong to a binary gender category and for whom they *do not have a label that confirms a binary gender category*. Thus, gender is *uncertain* via appearance and label. Gender uncertainty challenges children to go beyond categorical thinking. Viewing categories as a continuum, rather than with strict boundaries, leads children to assume more within-group and fewer between-group differences, which holds implications for the development of stereotyping (Master et al., 2012). Further, gender uncertainty can elucidate what information children value when category distinctions are unclear. We examined 5- to 8-year-olds' inferences about individuals with appearances and labels that denoted uncertainty about gender category membership. In the context of a story, an outside observer speculated about a target's appearance and claimed uncertainty about the target's gender category (Study 1) or noted that the target was not a boy or girl (Study 2). Previous research on how children use gender labels to form inductive inferences enabled us to predict how children might reason about gender uncertainty.

1.1 | Gender category labels

Broadly, children use categories to guide their predictions about their social worlds (Markman, 1989). They use category labels over perceptual similarities to make inferences about natural kinds, such as animals and objects (Gelman & Markman, 1986). Around preschool age, they prioritize gender category labels over conflicting appearance information: they predict that a boy with long hair likes to play with trucks, rather than dolls (Gelman et al., 1986). Given that children rely on gender category labels to navigate their social surroundings, an unfamiliar gender uncertainty label could instigate confusion.

Despite a potential lack of familiarity, 5- and 6-year-olds might accept and subsequently use a gender uncertainty label to guide their predictions about others. Children recognize multiple identities, and therefore different labels, for single entities by 4 years of age (Doherty & Perner, 2020; Perner et al., 2011). Consequently, 5- and 6-year-olds will likely accept a gender uncertainty label, even if a child initially categorized a gender uncertain character as a boy or a girl. Given that young children use gender category labels inductively, 5- and 6-year-olds will likely predict that someone with a gender uncertainty label shares properties with similarly labeled (i.e., gender uncertain) others, instead of people with specified gender labels (i.e., boys, girls). Therefore, 5- and 6-year-olds may use the individual's

gender uncertainty label to make inductive inferences about that person. If gender uncertainty is only shown through appearance and *without* a gender uncertainty label, 5- and 6-year-olds will likely assume that the individual belongs to a known, binary gender category and subsequently use the binary category label to make inductive inferences about the individual.

It is likely that 7- and 8-year-olds will also accept a gender uncertainty label. Children are thought to have an interpretive theory of mind by 7 years of age (e.g., Carpendale & Chandler, 1996). Specifically, children recognize that people can have access to the same information yet form different beliefs about that information: for example, an ambiguous drawing might be interpreted as a duck by one person and a rabbit by another person (Carpendale & Chandler, 1996). Upon viewing a gender uncertain individual, a 7- or 8-year-old will likely understand that one person could believe the individual is a boy, another person could believe the individual is a girl, and yet another person could believe the individual's gender is uncertain. Accordingly, a 7- or 8-year-old might readily accept gender uncertainty as one possible label for a gender uncertain individual, even if the 7- or 8-year-old initially categorized the gender uncertain individual as a boy or a girl.

Although 7- and 8-year-olds will likely accept a gender uncertainty label, this does not imply that they will actively *use* that label to make predictions about a gender uncertain person. Children's reliance on gender category labels for induction diminishes by 9 years of age (Taylor et al., 2009), mirrored by flexible gender attitudes (Ruble et al., 2006). Thus, older children might believe a gender uncertainty label is insufficient for assuming property similarities between people and therefore they may not use the gender uncertainty label for their inductive inferences about a gender uncertain person. If gender uncertainty is only shown through appearance and *without* a gender uncertainty label, 7- and 8-year-olds might assign a gender uncertain person to a binary gender category, but not necessarily *use* that category label for their predictions. Instead, 7- and 8-year-olds will likely rely on other information, such as traits or preferences, to guide their predictions about a gender uncertain person.

1.2 | Traits as sources for inductive inferences

By preschool age, children use traits to inform their social decision making, akin to their reliance on gender category labels. Children view traits as stable and use trait information to make behavioral predictions and other inferences (Boseovski & Lee, 2006; Liu et al., 2007). In one study, preschoolers saw a target who had similar appearance features (e.g., hair style) with one character and a similar trait (e.g., shy) as another character (Heyman & Gelman, 2000). They inferred that the target preferred the same novel activity as the character with a similar trait, rather than appearance. Thus, children believe that traits, but not appearance features, indicate additional common properties between people. In the context of gender uncertainty, 5- and 6-year-olds might assume that a gender uncertain person shares properties with someone that has a common trait label, rather than someone with common appearance characteristics, at least when gender uncertainty is only shown through appearance and *without* a gender uncertainty label. This will likely persist among 7- and 8-year-olds, as trait-based induction continues through childhood and is prevalent in adulthood (e.g., Diesendruck & haLevi, 2006; Gonzalez et al., 2010; Heyman & Gelman, 1999).

Other research investigates how children's inductive inferences change when provided with both trait information and gender category labels (Diesendruck & haLevi, 2006; Pillow et al., 2019), which can inform whether children's trait-based predictions will persist if they are provided with a gender uncertainty label. With a similar paradigm to the one detailed above, 5- and 6-year-olds used social categories instead of traits to make property predictions about a target character, while adults made trait-based predictions (Diesendruck & haLevi, 2006). However, if only analyzing gender categories, children's predictions about the target were at chance, while adult's predictions revealed trait-based reasoning. Thus, children view gender category labels and traits as similarly suggestive of additional property similarities between individuals (Diesendruck & haLevi, 2006). Given that children have sophisticated knowledge about traits and the implications of gender labels, it follows that they view both sources as sufficient for making inferences about others. Still, a lack of familiarity with gender uncertainty in comparison to the binary gender categories might lead

5- and 6-year-olds to overlook appearance and label commonalities between two gender uncertain individuals in favor of trait-based commonalities between a gender uncertain person and a person from a binary gender category. This may persist for 7- and 8-year-olds given that trait-based induction is prevalent across development. Other information, such as preferences, might also guide children's inferences in a gender uncertainty context.

1.3 | Personal preferences as sources for inductive inferences

Akin to the relevance of traits to children's social decision making, preschoolers rely on the stereotypes associated with each gender category to predict others' preferences (e.g., Conry-Murray, 2015). Still, 4-year-olds recognize that adherence to gender stereotypes is a personal choice, and they prioritize fairness over gender typicality concerns (Conry-Murray & Turiel, 2012). This does not necessarily mean that children will readily infer that people have counter-stereotypical preferences, particularly in public settings (Conry-Murray, 2013; Conry-Murray et al., 2020). Thus, young children recognize and value preferences, but do not always prioritize those preferences over other pertinent contextual information, such as someone's gender category and associated stereotypes, to guide their inferences about people.

There is evidence that children prioritize preferences over or as much as gender category membership with development (Biernat, 1991; Martin, 1989), following older children's flexible gender attitudes and the use of information sources beyond gender category membership for inductive reasoning (Taylor et al., 2009). For example, in one study, children viewed characters with gender stereotypical, counter-stereotypical, or neutral preferences (Martin, 1989). Three- to 5-year-olds used each character's gender category, rather than each character's provided preferences, to make inferences about each character. Conversely, 6- to 10-year-olds used each character's provided preferences and gender category to make inferences about each character. However, another study found that when preferences are not provided, 8-year-olds make inferences that follow gender norms (Conry-Murray, 2017). Thus, older children do not completely disregard an individual's gender category for induction, but simultaneously consider other information sources (e.g., personal preferences) if explicitly given. It is unclear whether this relation between preferences and gender categories extends to gender uncertainty.

Consistent with findings focused on the binary gender categories, 5- and 6-year-olds will likely rely on gender uncertainty information (i.e., appearance, label) to guide predictions about gender uncertain individuals, while 7- and 8-year-olds will likely engage in preference-based predictions. Younger children may assume that two gender uncertain people share properties with one another despite dissimilar preferences, rather than assume that a gender uncertain person shares properties with someone that has similar preferences but belongs to a binary gender category. Conversely, older children will infer that a gender uncertain person shares properties with people that have similar preferences, even though those people belong to a binary gender category. Given children's reliance on gender labels detailed above, the purported age difference will likely be stronger when a gender uncertainty label is provided, compared to when gender uncertainty is only conveyed through appearance.

1.4 | The current studies

We present data from two studies: our main study (Study 1) and a follow-up study (Study 2). Both studies incorporated a paradigm from past literature (e.g., Heyman & Gelman, 2000) to determine what information children use to make predictions about gender uncertain targets. Across studies, 5- to 8-year-olds predicted whether gender uncertain targets had the same properties as other gender uncertain people or people from a binary gender category (i.e., boy or girl). Both studies aimed to determine whether commonalities in the form of gender uncertainty, or commonalities in the form of traits and preferences, drove children's inferences about gender uncertain targets. Piloting suggested that children did not indiscriminately choose the character from a binary gender category (see Supporting Information).

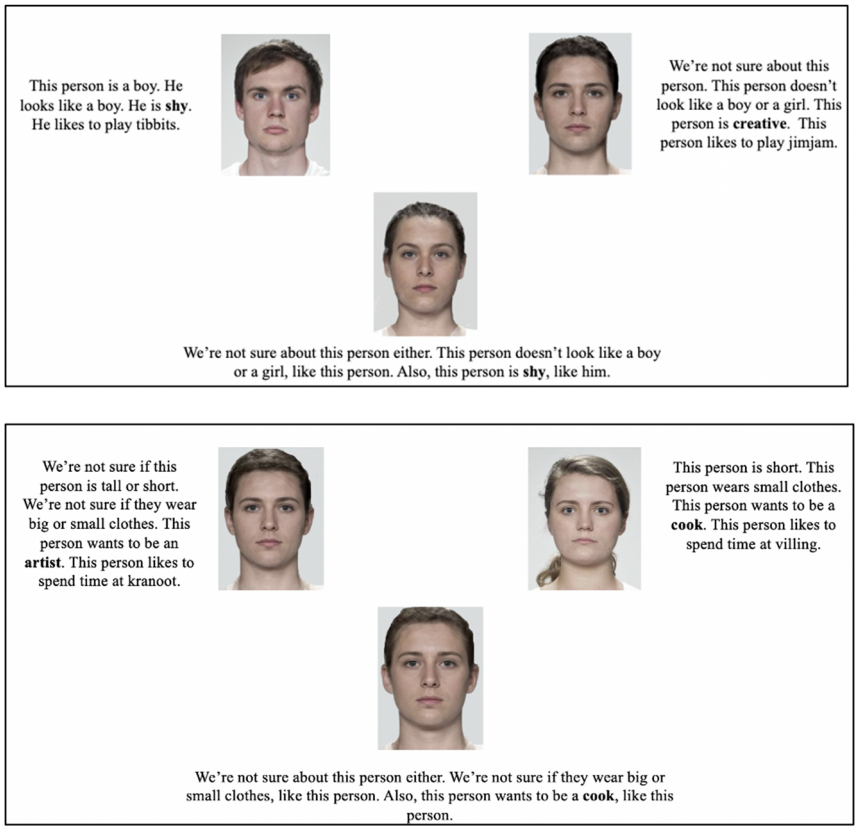

FIGURE 1 Example stimuli for Study 1. The top example illustrates *trait* information that a *boy* participant would receive in the *labeled condition*. Left to right: gender-specified character (man), gender uncertain target, gender uncertain character. The bottom example illustrates *preference* information that a *girl* participant would receive in *unlabeled condition*. Left to right: gender uncertain character, gender uncertain target, gender-specified character (woman).

Study 1 consisted of two vignettes. Each included a target with uncertain gender that had appearance similarities with another character whose appearance also denoted gender uncertainty (gender uncertain character), but trait or preference similarities (e.g., both characters described as shy or both wanted to be cooks) with another character whose appearance specified a gender category (gender-specified character). Further, half of the sample heard a label for each character (labeled condition). This included a known label for the gender-specified character (e.g., "This person is a girl. This person looks like a girl.") and a gender uncertainty label for the target and gender uncertain character (i.e., "We're not sure about this person. This person doesn't look like a boy or a girl."). The other half of the sample received miscellaneous information instead of gender labels (unlabeled condition), such as "This person is short. This person wears small clothes," for the gender-specified character and "We're not sure if this person is tall or short. We're not sure if they wear big or small clothes," for the target and gender uncertain character. See Figure 1.

Children inferred which character favored the same novel activity as the target (Novel Activity Question) and who the target should befriend (Friendship Question). The Novel Activity Question involved properties without gender associations. Conversely, the Friendship Question assessed whether children upheld same-gender friendship preferences, as past research suggests that children are more likely to infer that two same-gender individuals are friends, compared to two individuals of dissimilar gender (e.g., Maccoby, 1990). Compared to novel activities, friendship seems more applicable to children's real-world behaviors and attitudes.

Our hypotheses included the following:

We anticipated an interaction between age and labeling condition for both the Novel Activity and Friendship Questions. We anticipated that 5- and 6-year-olds would make more trait- and preference-based predictions in the unlabeled condition compared to the labeled condition. Further, we anticipated that 7- and 8-year-olds would make trait- and preference-based predictions regardless of labels. We hypothesized an interaction between age and labeling condition regardless of the similarity type between the target and gender-specified character (i.e., trait vs. preference).

We also expected an interaction between age and similarity type (i.e., trait vs. preference) for both the Novel Activity and Friendship Questions. For traits, we expected 5- and 6-year-olds, along with 7- and 8-year-olds, to systematically make trait-based predictions. For preferences, we expected 5- and 6-year-olds to systematically make predictions that centered on mutual gender uncertainty, but 7- and 8-year-olds to systematically make preference-based predictions.

For exploratory purposes, we asked children to report their liking for, desire for affiliation with, and trait attributions for each character. These exploratory measures investigated children's general attitudes about each character. We did not set any hypotheses for these measures. Further, a Gender Uncertainty Check determined whether children believed the targets belonged to a binary gender category. This primarily mattered for the unlabeled condition. Without a label to indicate gender uncertainty, participants could potentially categorize the target into a binary gender category.

2 | METHOD

2.1 | Participants

We conducted an a priori power analysis on G*Power to detect a within-between interaction with a mixed ANOVA. It revealed a sample size of 106 with a medium effect size of .25 (Cohen's *f*), alpha = .05, and 80% power. Past findings with a similar paradigm (i.e., Diesendruck & haLevi, 2006) report large effect sizes that ranged from $\eta^2 = .293$ to $\eta^2 = .127$. Because our work investigated an unexplored topic (i.e., gender uncertainty), we opted to be more conservative and use a medium effect size.

In total, 111 children participated (58 girls, 53 boys). We excluded two children for failing memory check questions (one 5-year-old, one 7-year-old). Failure entailed three or more repetitions of the information pertinent to any memory check question. We excluded three additional children due to experimenter error (two 6-year-olds, one 7-year-old). Responses from 53 5- and 6-year-olds ($M = 5.49$, $SD = .51$) and 53 7-and 8-year-olds ($M = 7.51$, $SD = .51$) were analyzed. Recruitment occurred via a database of families who agreed to participate in research and via day care or after-school programs in Greensboro, North Carolina and Fairfax, Virginia in the United States. Children participated at daycares/after-school programs or our laboratory. The sample was 72.7% White, 15.1% Black, 7.5% mixed race, and 4.7% unreported. From the sample, 8.5% did not report household income and 46.2% reported over $90,000.

2.2 | Materials

Stimuli include trios of photos with a gender-specified character that was a man (for boy participants) or woman (for girl participants), a gender uncertain character, and a gender uncertain target. The women had long hair in ponytails and the men had short hair. The gender uncertain individuals had short hair, but other feminine features (e.g., eyebrow shape). Children saw separate trios for each similarity type (trait, preference). See Appendix A and B.

Men's and women's faces were combined on the "Make an Average" feature on faceresearch.org. This created 8 gender uncertain faces and 11 undergraduates rated which faces were the most androgynous. Our final stimuli included the faces rated as most androgynous.

2.3 | Design

A 2 (age group: 5.0-6.9 vs. 7.0-8.9) x 2 (labeling condition: labeled vs. unlabeled) x 2 (similarity type: trait vs. preference similarities) mixed design was used, with labeling condition and age group as between-subject variables and similarity type as a within-subject variable. Participants in the labeled condition received known gender labels (i.e., girl, boy) or uncertain gender labels (i.e., "We're not sure…") for each character. Participants heard non-gender information (i.e., certainty or uncertainty about height and clothing size) in the unlabeled condition. Regardless of labeling condition, children received a vignette that established trait similarities between the target and gender-specified character *and* a vignette that established preference similarities between the target and gender-specified character. Gender-specified characters matched participants' reported gender.

2.4 | Procedure

Children were tested one-on-one by a researcher in a quiet room. Parents provided written informed consent prior to participation. All children provided verbal or written assent. In the lab, parents sat in an adjacent room and heard the entire session. In daycares or after-school programs, parents were not present.

Participants viewed a face trio and an accompanying vignette. One vignette included trait information and the other included preference information. Similarity type order was counterbalanced. The labels "shy" and "creative" established trait similarities. Although shyness in boys is rated as less socially acceptable than shyness in girls (Doey et al., 2014) and shyness is rated as a feminine trait (Liben & Bigler, 2002), some work suggests that children are not more likely to rate girls as shy compared to boys (Rubin et al., 1993, 2006) and gender differences reveal only a small effect (Else-Quest et al., 2006). "Cook" and "artist" established preference similarities, as children consider both gender neutral occupations (Liben & Bigler, 2002). Character traits and preferences were randomized. Character descriptions also included labeled gender information or unlabeled information, dependent upon each participant's assigned condition. See Figure 1 for examples and Appendix for the full set of information presented.

A memory check after each vignette ensured proper comprehension (e.g., "Did I say this person is shy or creative?"). Character order for the memory questions (e.g., which character was asked about first) was randomized.

Participants then answered the Novel Activity Question (e.g., "Does this person [point to target] like to play tibbits like this person who is shy [point to gender-specified character] or does this person like to play jimjam like this person who is creative [point to gender uncertain character]?") and Friendship Question (i.e., "Who should this person [point to target] be friends with?"). Question order was counterbalanced. For both questions, participants received a score of 0 for an inference centered on gender uncertainty (chose gender uncertain character) and 1 for a trait- or preference-based inference (chose gender specified character).

2.5 | Exploratory measures

2.5.1 | Liking and affiliation

Participants reported how much they liked ("How much do you like this person: a lot, a little, or in the middle?") and desired to affiliate with each character ("How much would you like to be friends with this person: a lot, a little, or in the middle?") in randomized order. Responses were scored as follows: 0 = a little, 1 = in the middle, 2 = a lot.

TABLE 1 Study 1 and Study 2: Means and standard deviations for the novel activity and friendship questions by similarity type

		Trait				Preference			
		Novel activity		Friendship		Novel activity		Friendship	
Study	Age	*n*	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*
1	5–6	53	.47 (.50)	53	.60 (.49)	53	.62 (.49)	53	.62 (.49)
	7–8	53	.70 (.46)**	53	.74 (.45)***	53	.59 (.50)	53	.64 (.48)*
2	5–6	36	.50 (.51)	36	.58 (.50)	36	.69 (.47)*	36	.39 (.49)

Note. Scoring: 0 = prediction centered on gender uncertainty (target and gender uncertain character), 1 = trait/preference-based prediction. * indicates significance against chance, $p < .05$, ** indicates significance against chance, $p < .01$, *** indicates significance against chance, $p < .001$.

2.5.2 | Trait attributions

Participants made trait attributions ("Do you think this person is nice, mean, or not nice or mean?") for each character in randomized order (Boseovski & Lee, 2006). Answers were coded as follows: 0 = mean, 1 = not nice or mean, 2 = nice.

2.5.3 | Gender uncertainty check

Participants answered the following about the targets: (1) "Do you think this person is a boy, a girl, or you're not sure?" (2) "Do you think this person looks like a boy, a girl, or you're not sure?" and (3) "Is this person friends with boys, girls, or both?" The targets were presented in randomized order. Question order and gender presentation order for each question (e.g., whether boy or girl was listed first for each question) was randomized. For each question, responses were coded as 1 if the participant categorized the target to a binary gender category and 0 if the participant reported uncertainty.

3 | RESULTS

The analyses below examined what information children prioritized to make novel activity and friendship predictions about gender uncertain targets. Our a priori power analysis was for a mixed ANOVA. After data collection, a statistical consultant instead suggested binary repeated measures logistic regression analyses that are detailed below. Our mixed ANOVA analyses are available in the Supporting Information section for transparency. These were consistent with our logistic regression analyses. Further, a sensitivity power analysis for a logistic regression on G*Power indicated that the sample of 106 children was enough to detect a medium effect size ($OR = .32$) with 80% power and alpha = .05.

See Table 1 for descriptive statistics by age group and similarity type for the Novel Activity Question and Friendship Question.

3.1 | Novel activity question

A 2 (age group: 5- and 6-year-olds vs. 7- and 8-year-olds) x 2 (labeling condition: labeled vs. unlabeled) x 2 (similarity type: trait vs. preference similarities) binary repeated measures logistic regression analysis was used with the GENMOD procedure on SAS (Spiegelman & Hertzmark, 2005). This assessed between- and within-subject differences

TABLE 2 Study 1: Means and standard deviations for the novel activity and friendship questions by labeling condition

	Labeled				Unlabeled			
	Novel activity		Friendship		Novel activity		Friendship	
Age	*n*	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*
5-6	53	1.04 (.81)	53	1.19 (.79)	53	1.15 (.73)	53	1.23 (1.19)
7-8	53	1.19 (.75)	53	1.30 (.79)	53	1.37 (.63)**	53	1.37 (.69)**

Note. Scoring: 1 for each trait- or preference-based prediction, summed across similarity types. ** indicates significance against chance, $p < .01$.

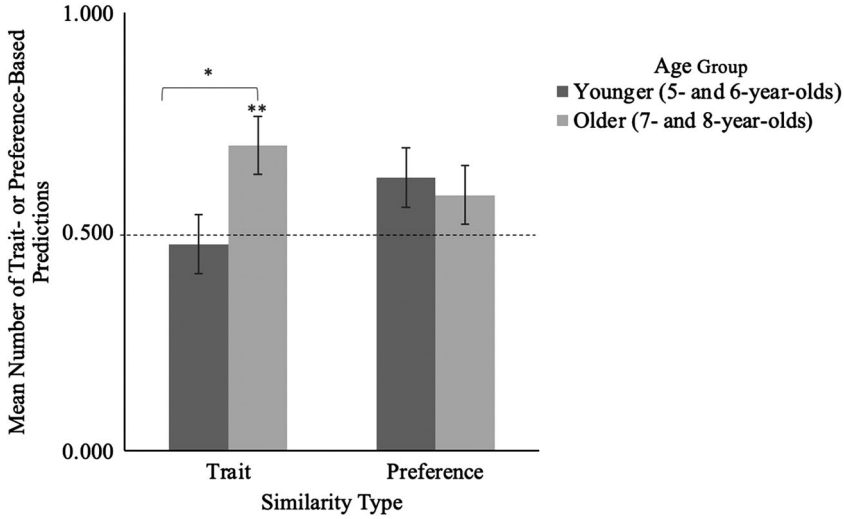

FIGURE 2 Mean number of trait- or preference-based predictions for the Novel Activity Question by age group for Study 1. Scoring: 0 = prediction centered on gender uncertainty (target and gender uncertain character), 1 = trait- or preference-based prediction (target and gender-specified character). Error bars indicate standard errors. * indicates $p < .05$. ** indicates significance against chance, $p < .01$

for the dichotomous responses (i.e., inference centered on gender uncertainty vs. trait- or preference-based inference). The same procedure was used for the Friendship Question.

Contrary to our hypotheses, there was no interaction between labeling condition and age, $\beta = .06$, $SE = .32$, $p = .84$. See Table 2 for descriptive statistics by labeling condition.

An interaction between age and similarity type emerged, $\beta = .62$, $SE = .30$, $OR = 1.85$, $p = .04$; see Figure 2. To further interpret this interaction, follow-up tests were conducted. For trait information, younger ($M = .47$, $SD = .50$) and older children ($M = .70$, $SD = .46$) differed significantly in their predictions, $\chi^2(1, N = 106) = 5.60$, $p = .02$. Older children were more likely than expected by chance to predict that the target favored the same novel activity as the gender-specified character with a similar trait rather than the gender uncertain character, $t(52) = 3.11$, $p = .003$. Younger children's predictions were no different from chance, $t(52) = -.41$, $p = .68$. Children's responses to trait similarities did not fully support our initial hypothesis, as we expected both older and younger children to systematically make trait-based predictions. For preference information, younger ($M = .62$, $SD = .49$) and older children's ($M = .59$, $SD = .50$) predictions did not differ significantly, $\chi^2(1, N = 106) = .16$, $p = .69$. Contrary to expectation, responses from both age groups did not differ from chance (*ps*>.05). Children's responses to preference similarities did not support our

initial hypothesis, as we expected younger children to systematically assume property similarities between the two gender uncertain characters (i.e., predictions centered on gender uncertainty) and older children to systematically make preference-based predictions.

3.2 | Friendship question

Contrary to our hypotheses, there was no significant interaction between age and labeling condition, $\beta = -.10$, $SE = .30$, $p = .73$. See Table 2 for descriptive statistics by labeling condition.

There was no significant interaction between age and similarity type, $p > .05$. Follow-up tests deciphered children's response patterns. For trait information, younger ($M = .60$, $SD = .49$) and older children's ($M = .74$, $SD = .45$) responses did not differ significantly, $\chi^2(1, N = 106) = 2.09$, $p = .15$. Older children were more likely than expected by chance to infer that the target should befriend the character with a similar trait rather than the gender uncertain character, $t(52) = 3.86$, $p < .001$. Younger children's responses did not differ from chance, $t(52) = 1.53$, $p = .13$. Children's responses to trait similarities did not fully support our initial hypothesis, as we anticipated that older and younger children would systematically make trait-based predictions. For preference similarities, younger ($M = .62$, $SD = .49$) and older children's ($M = .64$, $SD = .48$) responses did not differ significantly, $\chi^2(1, N = 106) = .04$, $p = .84$. Older children were more likely than expected by chance to infer that the target should befriend the character with a similar preference rather than the gender uncertain character, $t(52) = 2.13$, $p = .04$, and younger children's predictions did not differ from chance, $t(52) = 1.82$, $p = .07$. Children's responses to preference similarities did not fully support our hypothesis. Older children systematically made preference-based predictions, in support of our hypothesis. However, younger children did not systematically make predictions centered on gender uncertainty, which did not support our hypothesis.

3.3 | Exploratory measures

3.3.1 | Liking and affiliation

For the targets and gender-specified characters, there were no significant differences in liking and affiliation scores across similarity type ($ps > .05$). Scores were summed across similarity types to create Total Liking (range: 0–4) and Total Affiliation (range: 0–4) scores. Exploratory analyses indicated that responses to the Total Liking and Total Affiliation questions differed for the targets and gender-specified characters ($ps < .001$). For Total Liking, participant responses were neutral for the targets ($M = 1.96$, $SD = 1.13$), $t(105) = -.34$, $p = .73$, but greater than neutral for the gender-specified characters ($M = 2.59$, $SD = 1.18$), $t(105) = 5.20$, $p < .001$. For Total Affiliation, participant responses were greater than neutral for the targets ($M = 2.24$, $SD = 1.16$), $t(105) = 2.10$, $p = .04$, and gender-specified characters ($M = 2.73$, $SD = 1.15$), $t(105) = 6.50$, $p < .001$.

3.3.2 | Trait attributions

For the targets and gender-specified characters, trait attributions did not differ by similarity type ($ps > .05$). Scores were summed across similarity types to create Total Trait Attribution scores (range: 0–4). Exploratory analyses indicated that trait attributions for the targets and gender-specified characters did not differ significantly, $t(104) = -1.14$, $p = .26$. Children rated the targets ($M = 3.32$, $SD = 1.02$) and gender-specified characters ($M = 3.42$, $SD = .99$) positively, as responses were above neutral (targets: $t(104) = 13.25$, $p < .001$; gender-specified characters: $t(105) = 14.89$, $p < .001$).

TABLE 3 Study 1: Means and standard deviations for gender uncertainty check

	Labeled				Unlabeled			
	Trait		Preference		Trait		Preference	
Question	*n*	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*
1	53	.35 (.48)	53	.42 (.50)	53	.68 (.47)	53	.81 (.40)
2	52	.38 (.49)	53	.38 (.49)	53	.75 (.43)	53	.77 (.42)
3	52	.23 (.43)	53	.21 (.41)	53	.36 (.48)*	53	.36 (.48)

Note. Question 1: "Do you think this person is a boy, a girl, or you're not sure?" Question 2: "Do you think this person looks like a boy, a girl, or you're not sure?" Question 3: "Is this person friends with boys, girls, or both?" Note that for the third question, "both" equates to "not sure" on the table above. Scoring: 1 = boy or girl (i.e., categorizing target into a binary gender category), 0 = uncertainty (i.e., unsure/both).

3.4 | Gender uncertainty check

Chi-square analyses revealed that children in the unlabeled condition were more likely to categorize the targets into a binary gender category than children in the labeled condition (*p*s < .001). However, differences between labeling conditions did not arise when children were asked whether each target was friends with boys, girls, or both (trait: $p = .15$; preference: $p = .10$).

Binomial tests were performed to determine whether children's responses in the labeled and unlabeled conditions were no different from our expected distribution (50% chose *uncertain* if boy or girl, 50% chose a binary gender category, boy or girl) for each gender check question. For trait similarities in the unlabeled condition, 17 participants (32%) reported that they were unsure whether the target was a boy or a girl, $p = .01$, compared to 34 participants (65%) in the labeled condition, $p = .04$; 13 participants (25%) in the unlabeled condition reported that they were unsure whether the target looked like a boy or a girl, $p < .001$, compared to 32 participants (62%) in the labeled condition, $p = .13$; and 34 participants (64%) in the unlabeled condition reported that the target was friends with both boys and girls, $p = .05$, compared to 40 participants (77%) in the labeled condition, $p < .001$.

For preference similarities in the unlabeled condition, 10 participants (19%) reported that they were unsure whether the target was a boy or a girl, $p < .001$, compared to 30 participants (58%) in the labeled condition, $p = .33$; 12 participants (23%) in the unlabeled condition reported that they were unsure whether the target looked like a boy or a girl, $p < .001$, compared to 32 participants (62%) in the labeled condition, $p = .13$; and 34 participants (64%) in the unlabeled condition reported that the target was friends with both boys and girls, $p = .05$, compared to 41 participants (79%) in the labeled condition, $p < .001$.

See Table 3 for means and standard deviations for the Gender Uncertainty Check by labeling condition and similarity type.

4 | DISCUSSION

As anticipated, older children consistently made trait-based inferences about the targets for the Novel Activity and Friendship Questions. Surprisingly, younger children viewed appearance similarities (and/or gender uncertainty label similarities in the labeled condition) between the target and gender uncertain character as comparably relevant to their inferences as trait similarities between the target and gender-specified character. Unexpectedly, children made unsystematic novel activity predictions across age groups for preference information, likely due to limited knowledge about what occupations entail about people. Some children inferred that the target favored the same novel activity as the gender-specified character with a similar occupation preference, while others inferred that the target favored the same novel activity as the gender uncertain character. Moreover, only older children consistently inferred that

the target should befriend the gender-specified character with a similar occupation preference. Perhaps children perceived occupation preferences as more indicative of friendship choices than favored novel activities. Lastly, gender labels did not change children's tendencies to make trait- or preference-based predictions across age. These results are further examined in the General Discussion.

It is particularly puzzling that younger children failed to make systematic trait-based predictions about the target. Perhaps the unfamiliar gender uncertainty context dampened their abilities to overlook appearance similarities in favor of trait similarities (Heyman & Gelman, 2000). Alternatively, younger children's inconsistent predictions possibly arose due to confusion about the gender uncertainty label (i.e., "We're not sure about this person. This person doesn't look like a boy or a girl.") that implied experimenter speculation, at least in the labeled condition. A clearer label could instead note that the characters *were not* members of either binary gender category (i.e., "This person *is not* a boy or a girl") and prove more informative. To determine whether a clearer gender uncertainty label led younger children to make consistent trait-based inferences about the target, we ran a follow-up study.

4.1 | Study 2

Study 2 only included 5- and 6-year-olds to elucidate why they did not make trait-based predictions consistently during Study 1. Rather than describing the gender uncertain characters as "we're not sure about this person…this person doesn't look like a boy or a girl," Study 2 depicted gender uncertainty through lack of identification with either boys or girls. The experimenter told children that "this person is not a boy or a girl…we don't call this person a boy or a girl." Given that these label changes center on gender uncertainty through identification, Study 2 did not include a labeling manipulation.

For the Novel Activity and Friendship Questions, we expected a main effect of similarity type. We expected 5- and 6-year-olds to systematically make trait-based predictions. However, for preference similarities, we expected 5- and 6-year-olds to make predictions centered on mutual gender uncertainty. We also included the exploratory measures from Study 1, but we did not make predictions about children's responses.

5 | METHOD

5.1 | Participants

We initially sought a sample of 53 children to match our sample of 5- and 6-year-olds from Study 1 and due to the within-subject design of Study 2 (the between-subject variables from Study 1, labeling condition and age group, were not included in Study 2). We did not achieve this sample size due to COVID-19, which severely limited access to typical recruitment locations. Further, the study topic is sensitive, and some parents did not want to participate in a study about children's perceptions of gender beyond the binary categories. Given this limitation, the findings from Study 2 described below should be interpreted with caution. We also report a sensitivity analysis in the Study 2 Results section.

Thirty-six 5- and 6-year-olds (18 girls, 18 boys) participated on Zoom. Recruitment occurred through a database of interested families and ChildrenHelpingScience.com. The sample was 61.1% White, 8.3% Black, 13.9% Asian, 8.4% mixed race, and 8.3% unreported. Approximately 19.4% of the sample did not report household income and 61.2% reported over $90,000.

5.2 | Materials

All materials from Study 1 were screenshared on Zoom. Each character was emphasized with animation and a cursor circled over the character. See Appendix A and B.

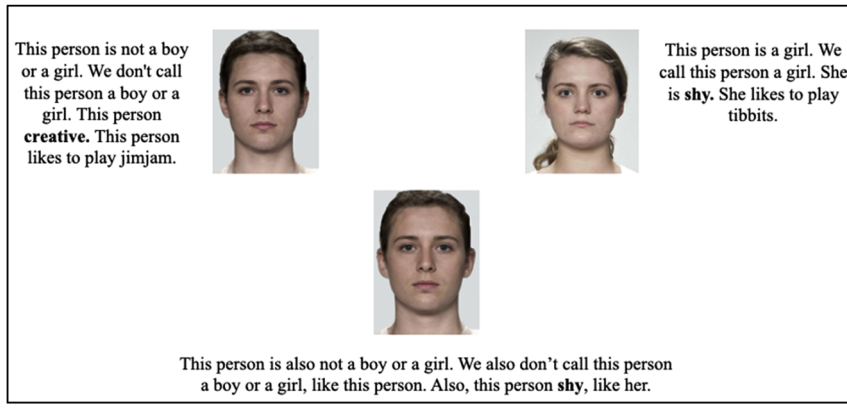

FIGURE 3 Example stimuli for Study 2. This is *trait* information that a *girl* participant would receive. Left to right: gender uncertain character, gender uncertain target, gender-specified character (woman)

5.3 | Design

A within-subject design (similarity type: trait vs. preference similarities) was used. Participants heard one vignette that established trait similarities between the target and gender-specified character and another that established preference (i.e., occupation) similarities between the target and gender-specified character. Similarity types were presented in counterbalanced order. Gender-specified characters matched participants' reported gender. Unlike Study 1, there was no labeling manipulation.

5.4 | Procedure

The procedure was nearly identical to Study 1 except that everything was administered virtually. Parents remained in the vicinity in case of internet connectivity issues. Gender uncertainty was conveyed as follows: "This person is not a boy or a girl. We don't call this person a boy or a girl." See Figure 3 for a full example and Appendix C for the full set of information presented to participants. The same measures from Study 1 were administered. Measures were scored identically to Study 1.

6 | RESULTS

Similar to Study 1, the analyses below examined what information children prioritized to make novel activity and friendship predictions about gender uncertain targets. A sensitivity power analysis for a logistic regression on G*Power suggested that a sample of 36 would detect a large effect size ($OR = .12$) with 80% power and alpha = .05. Thus, our sample size could not detect a medium or small effect size and results should be interpreted with caution.

See Table 1 for descriptive statistics for the Novel Activity Question and the Friendship Question by similarity type.

6.1 | Novel activity question

GEE analysis showed that predictions for the target did not differ by similarity type, $Wald\,\chi^2 = 2.36$, $SE = .53$, $OR = 2.27$, $p = .13$. Children ($M = .50$, $SD = .51$) were no more likely than expected by chance to provide trait-based predictions,

$t(35) = .00$, $p = 1.00$. By contrast, children ($M = .69$, $SD = .47$) were more likely than expected by chance to predict that the target favored the same novel activity as the gender-specified character with a similar preference rather than the gender uncertain character, $t(35) = 2.50$, $p = .02$. These results did not support our hypothesis for Study 2, as we anticipated that children would systematically make trait-based predictions for trait similarities, but predictions centered on gender uncertainty for preference similarities.

6.2 | Friendship question

GEE analysis demonstrated that predictions for the target did not differ by similarity type, $Wald\,\chi^2 = 2.64$, $OR = 2.20$, $SE = .49$, $p = .10$. Children were no more likely than expected by chance to make trait-based predictions ($M = .58$, $SD = .50$), $t(35) = 1.00$, $p = .32$. Similarly, children were no more likely than expected by chance to make preference-based predictions ($M = .39$, $SD = .49$), $t(35) = -1.35$, $p = .19$. Similar to the Novel Activity Question, the Friendship Question results did not support our hypothesis for Study 2.

6.3 | Exploratory measures

6.3.1 | Liking and affiliation

For the targets and gender-specified characters, liking and affiliation scores did not differ by similarity type ($ps > .05$). Scores were summed across similarity types to create Total Liking (range: 0–4) and Total Affiliation (range: 0–4) scores. Exploratory analyses indicated that responses to Total Liking did not differ for the targets and gender-specified characters, $t(35) = -1.59$, $p = .12$. Participant responses were neutral for the targets ($M = 1.78$, $SD = 1.02$), $t(35) = -1.31$, $p = .20$, and for the gender-specified characters ($M = 2.22$, $SD = 1.33$), $t(35) = 1.00$, $p = .32$. However, responses to Total Affiliation differed significantly between the targets and gender-specified characters, $t(35) = -2.99$, $p = .005$. Participants responses were less than neutral for the targets ($M = 1.56$, $SD = 1.16$), $t(35) = -2.30$, $p = .03$, but greater than neutral for the gender-specified characters ($M = 2.50$, $SD = 1.36$), $t(35) = 2.20$, $p = .03$.

6.3.2 | Trait attributions

For the targets and gender-specified characters, trait attributions did not differ by similarity type ($ps > .05$). Scores were summed across similarity types to create Total Trait Attribution scores (range: 0–4). Exploratory analyses indicated that trait attributions for the targets and gender-specified characters did not differ significantly, $t(35) = -.45$, $p = .65$. Children rated the targets ($M = 3.19$, $SD = .89$) and gender-specified characters ($M = 3.28$, $SD = .88$) positively, as responses were above neutral (targets: $t(35) = 8.07$, $p < .001$; gender-specified characters: $t(35) = 8.69$, $p < .001$).

6.3.3 | Gender uncertainty check

For each similarity type, binomial tests indicated that children's responses were no different from our expected distribution (50% chose *uncertain* if boy or girl, 50% chose a binary gender category, boy or girl) for each gender check question ($ps > .05$). There was one exception: for traits, 29 children (83%) reported that the target was friends with both boys and girls, $p < .001$. See Table 4 for means and standard deviations for the Gender Uncertainty Check.

TABLE 4 Study 2: Means and standard deviations for gender uncertainty check

	Trait		Preference	
Question	*n*	*M (SD)*	*n*	*M (SD)*
1	36	.39 (.49)	36	.58 (.50)
2	36	.56 (.50)	36	.50 (.51)
3	35	.17 (.38)	35	.37 (.49)

Note. Question 1: "Do you think this person is a boy, a girl, or you're not sure?" Question 2: "Do you think this person looks like a boy, a girl, or you're not sure?" Question 3: "Is this person friends with boys, girls, or both?" Note that for the third question, "both" equates to "not sure" on the table above. Scoring: 1 = boy or girl (i.e., categorizing target into a binary gender category), 0 = uncertainty (i.e., unsure/both).

7 | DISCUSSION

Despite a clearer gender uncertainty label (i.e., lack of identification with a gender category), children did not consistently make trait-based predictions about the target. Similar to Study 1, children valued gender category information (i.e., label, appearance) as much as trait information. This implies that children's inconsistent predictions in Study 1 did not arise due confusion about the gender uncertainty label used (i.e., Study 1: "We're not sure…"), as the inconsistency persisted with a changed label in Study 2. However, and unexpectedly, children in Study 2 predicted that the target favored the same novel activities as the gender-specified character with a similar occupation preference. For the Friendship Question, children did not systematically choose the gender-specified character with a similar trait or occupation preference as the target.

7.1 | General discussion

The present work explored the type of information that children use to make predictions about gender uncertain targets. In Study 1, children's predictions varied by similarity type and age, but predictions did not vary by labeling condition. As expected, 7- and 8-year-olds consistently made trait-based predictions about the targets for the Novel Activity and Friendship Questions, along with preference-based predictions for the Friendship Question. Children hold increasingly flexible gender attitudes with age (Blakemore, 2003; Conry-Murray & Turiel, 2012; Ruble et al., 2006) and perhaps this includes the willingness to discount a lack of clarity about an individual's gender. Conversely, younger children did not consistently make trait- or preference-based predictions for the Novel Activity or Friendship Questions in Study 1. Therefore, some younger children relied on gender category information (i.e., labeled condition: labels, appearance; unlabeled condition: appearance) and others relied on trait or preference information. In turn, younger children as a group used both kinds of information to guide their inductive inferences about the gender uncertain target. Alternatively, it is possible that younger children engaged in random responding, although this is unlikely. As mentioned previously, past findings reveal that children use traits, preferences, and gender category information (label, appearance), to guide their predictions about other people (e.g., Diesendruck & haLevi, 2006; Heyman & Gelman, 2000; Martin, 1989). If the present results arose from random responding, it would imply that children disregarded critical information that they used consistently in past research.

To clarify why younger children did not make the consistent trait-based predictions anticipated for Study 1, we conducted a follow-up study (Study 2) with a new label that replaced experimenter speculation with a lack of identification with a binary gender category. Younger children still valued gender category information (i.e., labels, appearance) as much as trait information for the Novel Activity and Friendship Questions. Unexpectedly, children in Study 2 privileged preferences over gender category information for their predictions about the target for the Novel Activity Question.

Across studies, children reported neutral to positive liking and desire for affiliation with the targets. Further, children reported above neutral (i.e., nice) ratings for the targets and gender-specified characters for the trait attribution question. This suggests flexible gender attitudes. This contrasts with research that indicates negativity against people who deviate from gender norms (Blakemore, 2003), but follows research that documents neutral to positive views about counter-stereotypical people (Boseovski et al., 2016). It is critical to further investigate gender diverse contexts, as adhering to *neither* gender category is perhaps perceived more positively than belonging to one gender category but failing to conform to gender category norms.

7.2 | Are gender uncertainty labels relevant to children?

Results from the labeled condition (Study 1) reveal that a mutual gender uncertainty label did not lead children to assume additional similarities between people. We expected only older children to overlook and therefore *not* use a mutual gender uncertainty label to make predictions about the target. However, the present findings suggest that both older *and* younger children chose *not* to use a mutual gender uncertainty label to make predictions about the target. Perhaps uncertainty labels are not as inductively powerful as familiar, binary gender category labels. Further, past research indicates that preschoolers engage in trait- over appearance-based reasoning (Heyman & Gelman, 2000). However, results from the unlabeled condition (Study 1) suggest that appearance information becomes more relevant to children's predictions when category information that is often distinctive through appearance (i.e., gender) is unclear. Indeed, children in the unlabeled condition did not systematically make trait- or preference-based predictions, such that that some made predictions about the target based on similar appearance characteristics between the target and gender uncertain character.

Although Study 2 included a clearer gender uncertainty label, young children did not make systematic predictions centered on gender uncertainty information. By contrast, in binary gender contexts, 5- and 6-year-olds use gender labels to make inductive inferences about people (Gelman et al., 1986; Ruble et al., 2006; Taylor et al., 2009). As mentioned previously, it is possible that gender uncertainty labels simply do not have the same inductive value as traditional gender category labels. It is also possible that some children perceived the target as a girl or a boy, as gender encoding occurs automatically around preschool age (Weisman et al., 2015). Since children recognize multiple identities for single entities by 4 years of age (Doherty & Perner, 2020; Perner et al., 2011), they perhaps recognized the experimenter's gender uncertainty label for the target when provided (labeled condition, Study 1; Study 2), yet maintained their own label for the target (e.g., this person is a boy).

Without gender uncertainty labels (unlabeled condition, Study 1), automatic categorization likely occurred across age groups and paralleled what occurs in the real world: spontaneous categorization, despite an appearance that does not necessitate a specific gender category. The Gender Uncertainty Check from Study 1 supports this interpretation: responses from children in the labeled condition were no greater than chance and therefore unsystematic about the target, yet those in the unlabeled condition systematically categorized the target into a binary gender category (i.e., boy or girl). Notably, the trait- vs. preference-based findings detailed below hold, as labeling conditions did not systematically change children's predictions about the target across age.

7.3 | Distinctive trait- vs. preference-based predictions

The finding that 7- and 8-year-olds made trait-based predictions for the target indicates improved trait understanding (Gonzalez et al., 2010). Sophisticated trait knowledge entails a better understanding of trait implications, which allowed older children to overlook gender uncertainty information. Conversely, 5- and 6-year-olds made unsystematic predictions for the target across both studies, rather than trait-based predictions, which entails less reliance on trait-based information when gender categorization is unclear. Perhaps this arose because young children's attention

centered on gender category differences (i.e., specified vs. uncertain in the present work). Thus, gender category information influenced some young children's predictions, even with other predictive (i.e., trait) information provided. This is consistent with findings that established that 5-year-olds use traits *and* gender category labels to make predictions about others (Diesendruck & haLevi, 2006). Given that gender is relevant early in ontogeny (Bigler & Liben, 2006; Weisman et al., 2015), it follows that gender labels persist as a basis for induction (Gelman et al., 1986).

Children made unsystematic novel activity predictions in response to preference information in Study 1, but systematically made preference-based novel activity predictions in Study 2. These results suggest that children's beliefs about gender uncertainty are less susceptible to the biases demonstrated in binary gender contexts. In past work, young children prioritized gender categories over preferences, while older children integrated both information sources (Martin, 1989). This indicates reliance on gender categories among younger children (Gelman et al., 1986; Taylor et al., 2009), but more reliance on preferences among older children (Biernat, 1991; Martin, 1989). However, this did not occur in the present data. In contrast to contexts with specified gender categories and familiar activities, perhaps participants in the present studies could not use their existing gender knowledge to guide their decisions. Indeed, children have schemas for the traditional boy and girl categories (see Ruble et al., 2006 for review), but it is unclear if they have a schema for gender uncertainty. Lack of gender uncertainty knowledge persisted in Study 2, where identification ("This person *is not* a boy or a girl.") replaced experimenter uncertainty from Study 1 ("We're not sure about this person…"). The identification label perhaps clarified to children that they did not know much about gender uncertainty and could therefore not use it inductively, at least not as much as a known and familiar occupation preference.

Both studies used occupation preferences that children perceive as gender neutral (Liben & Bigler, 2002). This might explain why 5- and 6-year-olds considered preferences to the same extent as gender category information (Study 1: labeled condition - label, appearance; unlabeled condition - appearance) or over gender category information (Study 2: label, appearance). Perhaps neutral occupations *and* a lack of pre-existing knowledge about people who do not *identify* with either gender category led 5- and 6-year-olds to make preference-based novel activity predictions in Study 2. This might not occur if children make judgments about individuals in the binary gender categories with gender-stereotypic or counter-stereotypic preferences. Instead, children might predict that a girl with counter-stereotypic preferences shares properties with other girls rather than boys with the same preferences as the girl. Indeed, it will be important for future work to incorporate gender-typed preferences, akin to past literature (e.g., Conry-Murray, 2015; Conry-Murray et al., 2020).

In Study 1, older children inferred that the target should befriend the gender-specified character with a similar occupation preference, likely because the present work incorporated a *friendship* context. When older children heard that the target wanted to pursue a specific occupation, perhaps they understood that the target partakes in activities common for that occupation. In turn, this suggests common behaviors between people with similar occupation preferences (e.g., artists paint). Consistent with the present findings, past research suggests that children view similar activities as a basis for friendship (Maccoby, 1990).

In Study 1 and Study 2, some younger children claimed the target should befriend the gender-specified character, while others instead chose the gender uncertain character. Thus, children did not systematically rely on similar gender uncertainty information to make friendship predictions about the target, despite the dominance of gender category-based reasoning at 5 to 6 years of age (e.g., Taylor et al., 2009). As mentioned previously, it is likely that gender uncertainty labels do not hold the same inductive potential as the binary gender category labels, at least among young children.

7.4 | Limitations

Although the gender-specified characters were described as short and no height information accompanied the gender uncertain characters (i.e., "We're not sure if this person is short or tall.") in the unlabeled condition for Study 1, perhaps

children adhered to gender stereotypes about height (e.g., women are shorter than men). Also, descriptions in the unlabeled condition included adjectives, whereas the labeled condition included nouns. Compared to adjectives, nouns provide a stronger basis of inference about others (Gelman et al., 2000). Still, labeling conditions did not systematically change children's predictions about the targets.

Additionally, future research should assess the potential association between cognitive flexibility and children's predictions and attitudes toward gender uncertain individuals, including comparisons among children with high vs. low endorsements of gender stereotypes. In the present work, performance did not differ from chance for some measures in Study 1 and 2, suggesting different prediction strategies across participants. Although memory check questions ensured that all included children remembered the information provided, it is possible that children with better working memory ability better remembered character details. Additionally, the ability to categorize individuals on multiple dimensions entails the use of cognitive flexibility, which is related to less gender stereotyping (Bigler, 1995; Bigler & Liben, 1992) and therefore less reliance on categories for inferring an individual's properties. Consequently, cognitive flexibility skills might have facilitated children's abilities to shift between the target's reported gender uncertainty and their own beliefs that the target perhaps belonged to a binary gender category.

Future studies should also account for the impact of participants' own traits and preferences. There is a possibility that participants selected characters that aligned with their own traits and preferences, particularly among young children who made unsystematic predictions about the targets. Information about each participant's own traits and preferences could help determine how much those details influenced their decision making. Still, randomization of the traits and preferences assigned to each character helped to rule out any effects related to the specific traits and preferences included in the present work.

The present research only included cisgender participants according to parent reports. It is possible that non-binary or transgender youth, or children frequently exposed to gender diversity, would make different predictions about gender uncertain targets. Additional demographic data, such as parents' perceived gender identity or gender role expectations in the home, could also be relevant. In fact, it is critical to note the effect of participant gender on the present data. This is in the Supporting Information section, but we did not have sufficient power to account for participant gender effects. The results should be taken with caution.

Sensitivity analyses for Study 1 indicated enough power to detect medium effects, while Study 2 had enough power to detect large effects. Thus, results from Study 2 should be taken with caution, as additional, small effects of age or similarity type could potentially be found with a larger sample. However, we did not design the present work to detect small effects of age (Study 1 and 2), similarity type (Study 1 and 2), or labeling condition (Study 1). Importantly, we detected significant results (discussed above) in both studies.

8 | CONCLUSION

Although 5- and 6-year-olds used gender category information (Study 1: labeled condition - label, appearance; unlabeled condition – appearance; Study 2: label, appearance) to the same extent as traits to guide their inferences about gender uncertain people, this dissipated with age. Indeed, 7- and 8-year-olds valued traits over gender category information to guide their inferences about gender uncertain people. Moreover, 5- and 6-year-olds valued preferences to the same extent as gender category information, but preferences gained relevance when identification (Study 2), rather than experimenter speculation (Study 1), conveyed gender uncertainty. It is important to note that null findings could indicate random responding. However, this interpretation is questionable, as it is unlikely that children disregarded the character information presented in the current study. Notably, these findings showcase the need to further investigate how children reason about gender diversity. Conceptualizing social characteristics on a continuum, instead of strict categories, may minimize the use of stereotypes (Master et al., 2012). This developmental period may be ideal for educating children about gender diversity to avoid the putative onset of prejudice and negative stereotyping of gender uncertain people.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest. This study was approved by the University of North Carolina at Greensboro Institutional Review Board.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request.

ORCID

Andrea C. Yuly-Youngblood https://orcid.org/0000-0003-1525-4604
Janet J. Boseovski https://orcid.org/0000-0002-9712-1401

REFERENCES

Biernat, M. (1991). Gender stereotypes and the relationship between masculinity and femininity: A developmental analysis. *Journal of Personality and Social Psychology*, *61*(3), 351–365. https://doi.org/10.1037/0022-3514.61.3.351

Bigler, R. S. (1995). The role of classification skill in moderating environmental influences on children's gender stereotyping: A study of the functional use of gender in the classroom. *Child Development*, *66*(4), 1072–1087. http://doi.org/10.1111/j.1467-8624.1995.tb00923.x

Bigler, R. S., & Liben, L. S. (1992). Cognitive mechanisms in children's gender stereotyping: Theoretical and educational implications of a cognitive-based intervention. *Child Development*, *63*(6), 1351–1363. http://doi.org/10.1111/j.1467-8624.1992.tb01700.x

Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. In R.V. Kail (Ed.), *Advances in child development and behavior* (pp. 39–89). Elsevier. https://doi.org/10.1016/S0065-2407(06)80004-2

Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex Roles*, *48*(9-10), 411–419. https://doi.org/10.1023/A:1023574427720

Boseovski, J. J., Hughes, C., & Miller, S. E. (2016). Expertise in unexpected places: Children's acceptance of information from gender counter-stereotypical experts. *Journal of Experimental Child Psychology*, *141*, 161–176. https://doi.org/10.1016/j.jecp.2015.09.002

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology*, *42*(3), 500–513. https://doi.org/10.1037/0012-1649.42.3.500

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*(4), 1686–1706. http://doi.org/10.1111/j.1467-8624.1996.tb01821.x

Conry-Murray, C. (2013). Children's reasoning about gender-atypical preferences in different settings. *Journal of Experimental Child Psychology*, *115*(1), 210–217. http://doi.org/10.1016/j.jecp.2012.09.007

Conry-Murray, C. (2015). Children's judgments of inequitable distributions that conform to gender norms. *Merrill-Palmer Quarterly*, *61*(3), 319–344. http://doi.org/10.13110/merrpalmquar1982.61.3.0319

Conry-Murray, C. (2017). Children's distributive justice:Tthe role of gender norms in different settings. *European Journal of Developmental Psychology*, *14*(1), 62–74. http://doi.org/10.1080/17405629.2016.1152176

Conry-Murray, C., Kim, J. M., & Turiel, E. (2020). Culture and children's reasoning about preferences and gender norms. *Journal of Experimental Child Psychology*, *196*, 104861. http://doi.org/10.1016/j.jecp.2020.104861

Conry-Murray, C., & Turiel, E. (2012). Jimmy's baby doll and Jenny's truck: Young children's reasoning about gender norms. *Child Development*, *83*(1), 146–158. https://doi.org/10.1111/j.1467-8624.2011.01696.x

Diesendruck, G., & haLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development*, *77*(3), 539–553. https://doi.org/10.1111/j.1467-8624.2006.00889.x

Doey, L., Coplan, R. J., & Kingsbury, M. (2014). Bashful boys and coy girls: A review of gender differences in childhood shyness. *Sex Roles*, *70*(7-8), 255–266. https://doi.org/10.1007/s11199-013-0317-9

Doherty, M. J., & Perner, J. (2020). Mental files: Developmental integration of dual naming and theory of mind. *Developmental Review*, *56*, 100909. http://doi.org/10.1016/j.dr.2020.100909

Dunham, Y., & Olson, K. R. (2016). Beyond discrete categories: Studying multiracial, intersex, and transgender children will strengthen basic developmental science. *Journal of Cognition and Development*, *17*(4), 642–665. https://doi.org/10.1080/15248372.2016.1195388

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: a meta-analysis. *Psychological Bulletin*, *132*(1), 33–72. https://doi.org/10.1037/0033-2909.132.1.33

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development*, *57*(2), 396–404. https://doi.org/10.2307/1130595

Gelman, S. A., Hollander, M., Star, J., & Heyman, G. D. (2000). The role of language in the construction of kinds. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 39, pp. 201–263). Academic Press. https://doi.org/10.1016/S0079-7421(00)80035-3

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*(3), 183–209. http://doi.org/10.1016/0010-0277(86)90034-X

Gonzalez, C. M., Zosuls, K. M., & Ruble, D. N. (2010). Traits as dimensions or categories? Developmental change in the understanding of trait terms. *Developmental Psychology*, *46*(5), 1078–1088. https://doi.org/10.1037/a0020207

Halim, M. L. D. (2016). Princesses and Superheroes: Social-Cognitive Influences on Early Gender Rigidity. *Child Development Perspectives*, *10*(3), 155–160. https://doi.org/10.1111/cdep.12176

Heyman, G. D., & Gelman, S. A. (1999). The use of trait labels in making psychological inferences. *Child Development*, *70*(3), 604–619. https://doi.org/10.1111/1467-8624.00044

Heyman, G. D., & Gelman, S. A. (2000). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology*, *77*(1), 1–19. https://doi.org/10.1006/jecp.1999.2555

Liben, L. S., & Bigler, R. S. (2002). Monographs of the society for research in child development. *Society for Research in Child Development*, *67*(2), 1–6. https://doi.org/10.1111/1540-5834.t01-1-00188

Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development*, *78*(5), 1543–1558. https://doi.org/10.1111/j.1467-8624.2007.01082.x

Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist*, *45*(4), 513–520. http://doi.org/10.1037/0003-066X.45.4.513

Markman, E. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.

Martin, C. L. (1989). Children's use of gender-related information in making social judgments. *Developmental Psychology*, *25*(1), 80–88. https://doi.org/10.1037/0012-1649.25.1.80

Master, A., Markman, E. M., & Dweck, C. S. (2012). Thinking in categories or along a continuum: Consequences for children's social judgments. *Child Development*, *83*(4), 1145–1163. http://doi.org/10.1111/j.1467-8624.2012.01774.x

Perner, J., Mauer, M. C., & Hildenbrand, M. (2011). Identity: Key to children's understanding of belief. *Science*, *333*(6041), 474–477. http://doi.org/10.1126/science.1201216

Pillow, B. H., Allen, C., Low, N., & Vilma, T. (2019). Young children's use of gender for inductive generalizations about biological and behavioral characteristics: The influence of gender categories and gender stereotypes. *Journal of Educational and Developmental Psychology*, *9*(2), 37–44. http://doi.org/10.5539/jedp.v9n2p37

Rubin, K. H., Chen, X., & Hymel, S. (1993). Socioemotional characteristics of withdrawn and aggressive children. *Merrill-Palmer Quarterly*, *39*(4), 518–534. http://jstor.org/stable/23087247

Rubin, K. H., Wojslawowicz, J. C., Rose-Krasnor, L., Booth-LaForce, C., & Burgess, K. B. (2006). The best friendships of shy/withdrawn children: Prevalence, stability, and relationship quality. *Journal of Abnormal Child Psychology*, *34*(2), 139–153. http://doi.org/10.1007/s10802-005-9017-4

Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In W. Damon, R. M. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology* (6th ed., pp. 858–932). John Wiley. https://doi.org/10.1002/9780470147658.chpsy0314

Spiegelman, D., & Hertzmark, E. (2005). Easy SAS calculations for risk or prevalence ratios and differences. *American Journal of Epidemiology*, *162*(3), 199–200. https://doi.org/10.1093/aje/kwi188

Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development*, *80*(2), 461–481. http://doi.org/10.1111/j.1467-8624.2009.01272.x

Weisman, K., Johnson, M. V., & Shutts, K. (2015). Young children's automatic encoding of social categories. *Developmental Science*, *18*(6), 1036–1043. https://doi.org/10.1111/desc.12269

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Yuly-Youngblood, A. C., & Boseovski, J. J. (2022). Children's inductive inferences about individuals with gender category uncertainty. *Social Development*, 1–25. https://doi.org/10.1111/sode.12609

APPENDIX A

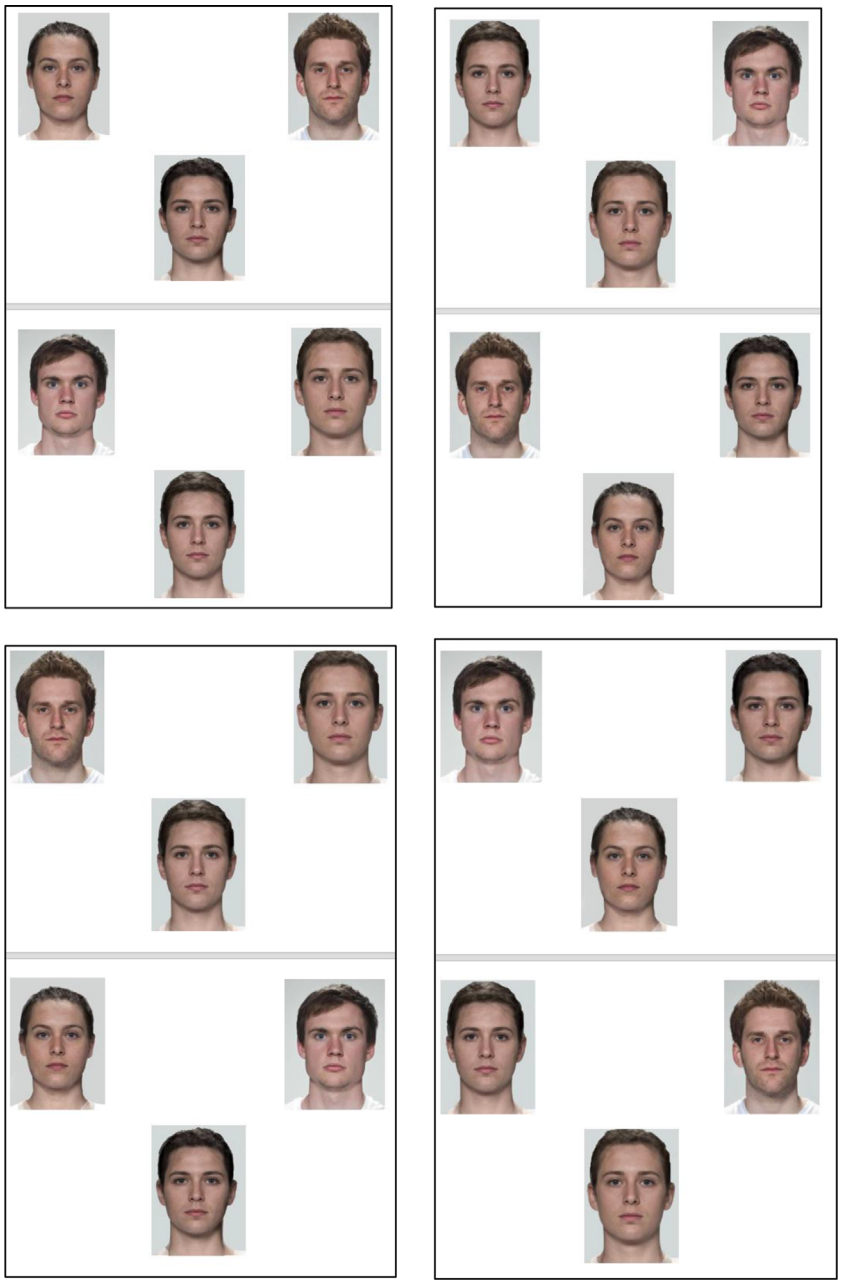

FIGURE A1 Stimuli for boys (pilot study, Study 1, and Study 2)

APPENDIX B

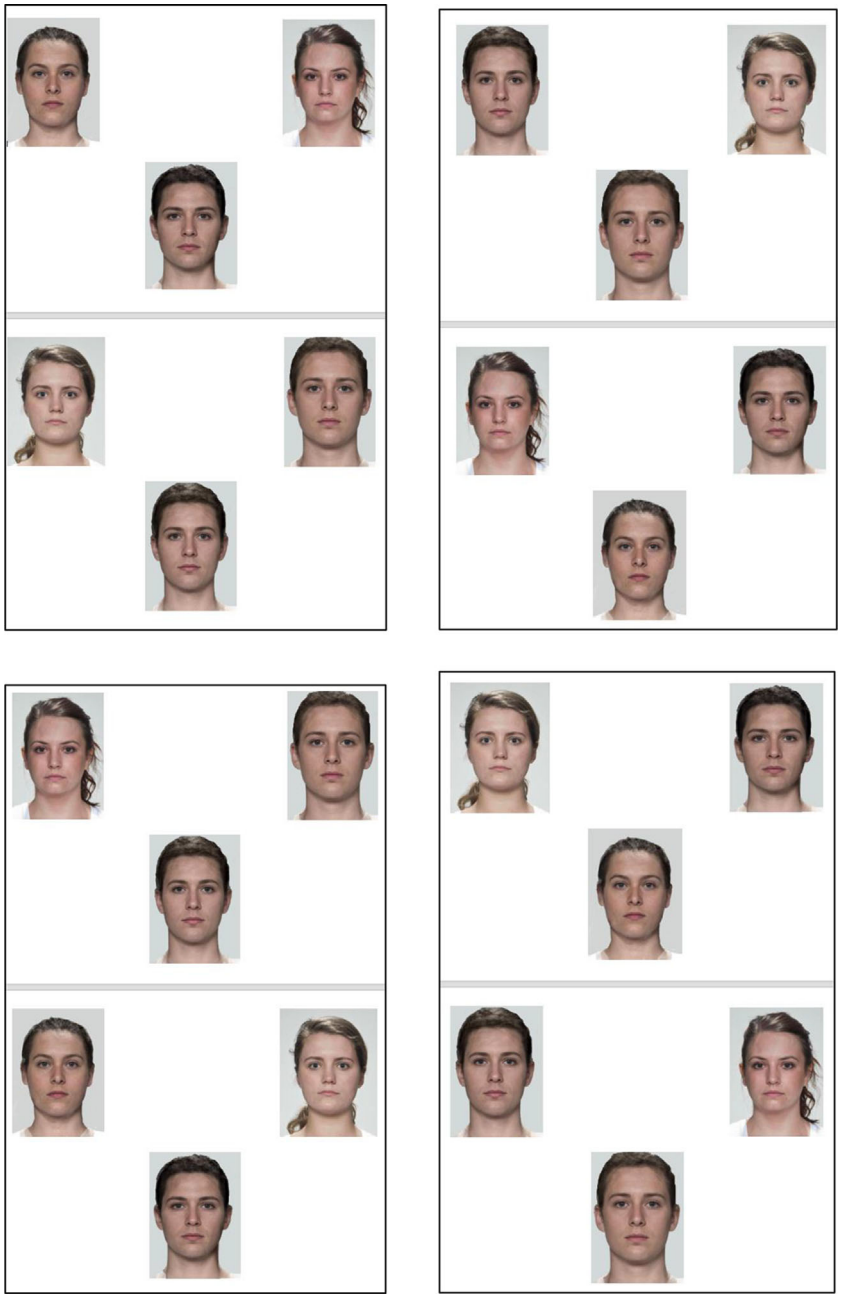

FIGURE B1 Stimuli for girls (pilot study, Study 1, and Study 2)

APPENDIX C

Information presented to participants in Study 1 and Study 2 (Figure C1, C2 and C3)

Girls, labeled condition, trait (order for 2 and 3 varies due to randomization)
1) Here are three people.

2) This person *[point to gender specified character]* is a girl. She looks like a girl. She is **SHY / CREATIVE**. She likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) We're not sure about this person *[point to gender uncertain character*]. This person doesn't look like a boy or a girl. This person *[continue pointing]* is **SHY / CREATIVE**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) We're not sure about this person *[point to target]*. This person doesn't look like a boy or a girl, like this person *[point to gender uncertain character]*. Also, this person *[continue pointing to target]* is **SHY / CREATIVE**, like her *[point to gender specified character]*.

Girls, labeled condition, occupation (order for 2 and 3 varies due to randomization)
1) Here are three people.

2) This person *[point to gender specified character]* is a girl. She looks like a girl. She *[continue pointing]* wants to be a **COOK / ARTIST**. She likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) We're not sure about this person *[point to gender uncertain character*]. This person doesn't look like a boy or a girl. This person *[continue pointing]* wants to be a **COOK / ARTIST**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) We're not sure about this person *[point to target]*. This person doesn't look like a boy or a girl, like this person *[point to gender uncertain character]*. Also, this person *[continue pointing to target]* wants to be a **COOK / ARTIST**, like her *[point to gender specified character]*.

Girls, unlabeled condition, trait (order for 2 and 3 varies due to randomization)
1) Here are three people.

2) This person *[point to gender specified character]* is short. This person wears small clothes. This person is **SHY / CREATIVE**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) We're not sure if this person *[point to gender uncertain character]* is tall or short. We're not sure if they wear big or small clothes. This person *[continue pointing]* is **SHY / CREATIVE**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) We're not sure if this person *[point to target]* is tall or short. We're not sure if they wear big or small clothes, like this person [*point to gender uncertain character]*. This person *[continue pointing to target]* is **SHY / CREATIVE**, like this person *[point to gender specified character]*.

Girls, unlabeled condition, occupation (order for 2 and 3 varies due to randomization)
1) Here are three people.

2) This person *[point to gender specified character]* is short. This person wears small clothes. This person *[continue pointing]* wants to be a **COOK / ARTIST**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) We're not sure if this person *[point to gender uncertain character]* is tall or short. We're not sure if they wear big or small clothes. This person *[continue pointing]* wants to be a **COOK / ARTIST**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) We're not sure if this person *[point to target]* is tall or short. We're not sure if they wear big or small clothes, like this person [*point to gender uncertain character]*. This person *[continue pointing to target]* wants to be a **COOK / ARTIST**, like this person *[point to gender specified character]*.

FIGURE C1 Study 1, Girls

Boys, labeled condition, trait (order for 2 and 3 varies due to randomization)
1) Here are three people.

2) This person *[point to gender specified character]* is a boy. He looks like a boy. He is **SHY / CREATIVE**. He likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) We're not sure about this person *[point to gender uncertain character]*. This person doesn't look like a boy or a girl. This person *[continue pointing]* is **SHY / CREATIVE.** This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) We're not sure about this person *[point to target]*. This person doesn't look like a boy or a girl, like this person *[point to gender uncertain character]*. Also, this person *[continue pointing to target]* is **SHY / CREATIVE**, like him *[point to gender specified character]*.

Boys, labeled condition, occupation (order for 2 and 3 varies due to randomization)
1) Here are three people.

2) This person *[point to gender specified character]* is a boy. He looks like a boy. He wants *[continue pointing]* to be a **COOK / ARTIST**. He likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) We're not sure about this person *[point to gender uncertain character]*. This person doesn't look like a boy or a girl. This person *[continue pointing]* wants to be **COOK / ARTIST**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) We're not sure about this person *[point to target]*. This person doesn't look like a boy or a girl, like this person *[point to gender uncertain character]*. Also, this person *[continue pointing to target]* wants to be **COOK / ARTIST**, like him *[point to gender specified character]*.

Boys, unlabeled condition, trait (order for 2 and 3 varies due to randomization)
1) Here are three people.

2) This person *[point to gender specified character]* is short. This person wears small clothes. This person is **SHY / CREATIVE**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) We're not sure if this person *[point to gender uncertain character]* is tall or short. We're not sure if they wear big or small clothes. This person *[continue pointing]* is **SHY / CREATIVE**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) We're not sure if this person *[point to target]* is tall or short. We're not sure if they wear big or small clothes, like this person *[point to gender uncertain character]*. This person *[continue pointing to target]* is **SHY / CREATIVE**, like this person *[point to gender specified character]*.

Boys, unlabeled condition, occupation (order for 2 and 3 varies due to randomization)
1) Here are three people.

2) This person *[point to gender specified character]* is short. This person wears small clothes. This person *[continue pointing]* wants to be a **COOK / ARTIST**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) We're not sure if this person *[point to gender uncertain character]* is tall or short. We're not sure if they wear big or small clothes. This person *[continue pointing]* wants to be a **COOK / ARTIST**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) We're not sure if this person *[point to target]* is tall or short. We're not sure if they wear big or small clothes, like this person [*point to gender uncertain character]*. This person *[continue pointing to target]* wants to be a **COOK / ARTIST**, like this person *[point to gender specified character]*.

FIGURE C2 Study 1, Boys

Girls, trait (order for 2 and 3 varies due to randomization):
1) Here are three people.

2) This person *[point to gender specified character]* is a girl. We call this person a girl. She is **SHY / CREATIVE**. She likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) This person *[point to gender uncertain character]* is not a boy or a girl. We don't call this person a boy or a girl. This person *[continue pointing]* is **SHY / CREATIVE**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) This person is also not a boy or a girl *[point to target]*. We also don't call this person a boy or a girl, like this person *[gender uncertain character]*. Also, this person *[continue pointing to target]* is **SHY / CREATIVE**, like her *[point to gender specified character]*.

Girls, occupation (order for 2 and 3 varies due to randomization):
1) Here are three people.

2) This person *[point to gender specified character]* is a girl. We call this person a girl. She *[continue pointing]* wants to be a **COOK / ARTIST**. She likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) This person *[point to gender uncertain character]* is not a boy or a girl. We don't call this person a boy or a girl. This person *[continue pointing]* wants to be a **COOK / ARTIST**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) This person is also not a boy or a girl *[point to target]*. We also don't call this person a boy or a girl, like this person *[point to gender uncertain character]*. Also, this person *[continue pointing to target]* wants to be a **COOK / ARTIST**, like her *[point to gender specified character]*.

Boys, trait (order for 2 and 3 varies due to randomization):
1) Here are three people.

2) This person *[point to gender specified character]* is a boy. We call this person a boy. He is **SHY / CREATIVE**. He likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) This person *[point to gender uncertain character]* is not a boy or a girl. We don't call this person a boy or a girl. This person *[continue pointing]* is **SHY / CREATIVE.** This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) This person is also not a boy or a girl *[point to target]*. We also don't call this person a boy or a girl, like this person *[point to gender uncertain character]*. Also, this person *[continue pointing to target]* is **SHY / CREATIVE**, like him *[point to gender specified character]*.

Boys, occupation (order for 2 and 3 varies due to randomization):
1) Here are three people.

2) This person *[point to gender specified character]* is a boy. We call this person a boy. He wants *[continue pointing]* to be a **COOK / ARTIST**. He likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

3) This person *[point to gender uncertain character]* is not a boy or a girl. We don't call this person a boy or a girl. This person *[continue pointing]* wants to be **COOK / ARTIST**. This person likes to [play **TIBBITS / JIMJAM OR** spend time at **VILLING / KRANOOT**].

4) This person is also not a boy or a girl *[point to target]*. We also don't call this person a boy or a girl, like this person *[point to gender uncertain character]*. Also, this person *[continue pointing to target]* wants to be **COOK / ARTIST**, like him *[point to gender specified character]*.

FIGURE C3 Study 2