THE DEVELOPMENT OF STRUCTURAL REASONING

ABOUT SOCIAL INEQUALITY

by

Cameron E. Hall

B.A. (Hons.), The University of Texas at Austin, 2019

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENT OF THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2023

© Cameron E. Hall, 2023
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

The Development of Structural Reasoning about Social Inequality

Submitted by Cameron E. Hall in partial fulfillment of the requirements for the degree of Master of Arts in Psychology

Examining Committee:

Dr. Andrew Scott Baron, Professor, Psychology, UBC
Supervisor

Dr. Lauren Emberson, Professor, Psychology, UBC
Supervisory Committee Member

Dr. Darko Odic, Professor, Psychology, UBC
Supervisory Committee Member
Abstract

Structural inequalities create and perpetuate discrepant outcomes between social groups. However, both adults and children more often attribute discrepancies between social groups to internal factors of individuals, rather than to structural causes—a tendency which has been linked to increased prejudice toward disadvantaged group members across development. Therefore, promoting structural reasoning in childhood may serve as a means to mitigate the development of social biases. Across three studies we investigated children’s (3-to-8-years-old) ability to selectively make structural attributions for social inequality. Study 1 examined how equitable or inequitable outcomes between groups influenced children’s perception of structural factors and their ability to endorse structural attributions. Study 2 aimed to address alternative explanations for previous findings. In Study 3 we explored children’s ability to accurately map structural and internalist attributions to appropriate scenarios. Together, these three studies shed light on the developmental trajectory of structural reasoning. At 3-to-4-years-old, children demonstrated the capacity to recognize structural constraints and appropriately attribute outcomes to their respective internal and structural causes. By age five, children can also overcome internalist defaults to evaluate structural attributions for social inequality as superior. These findings suggest that children as young as three years old are capable of selectively applying structural reasoning and provide insight into how this type of reasoning may influence the development of intergroup biases.
Lay summary

When attempting to explain social inequalities, such as wealth gaps or unequal access to opportunities, people tend to attribute these disparities to individual characteristics or traits, while overlooking the influence of social structures and systems. Unfortunately, this tendency can contribute to increased prejudice and discrimination towards disadvantaged groups, including women, people of color, and individuals from low-income backgrounds. However, acknowledging and endorsing the role of social structures in creating unequal outcomes between groups can help diminish biased attitudes. Given that social biases begin to develop in childhood, this study aimed to investigate if, when, and how children consider structural factors when evaluating inequality between groups of people. The findings revealed that children as young as three-years-old can recognize and endorse structural causes for inequality. We discuss the implications of our findings for the development of prejudice in childhood.
Preface

This thesis is an original, unpublished intellectual product of the author, C. Hall. The research was conducted by the author at the University of British Columbia, Social Cognitive Development Lab, under the supervision of A. S. Baron, who was involved in research design. All work and associated methods were approved by the University of British Columbia’s Research Ethics Board (H10-0047: The Development of Social Cognition).
# Table of Contents

Abstract.......................................................................................................................... iii
Lay summary..................................................................................................................... iv
Preface............................................................................................................................... v
Table of Contents ............................................................................................................ vi
List of Figures .................................................................................................................. vii
Acknowledgements .......................................................................................................... viii

## Introduction

- Internalist reasoning about social categories ................................................................ 1
- Structural reasoning about social categories ................................................................. 6
- The present study ........................................................................................................... 11

## Study 1

- Method ......................................................................................................................... 12
  - Participants .................................................................................................................. 12
  - Materials and Procedure ............................................................................................ 12
- Results ........................................................................................................................... 15
- Discussion ...................................................................................................................... 19

## Study 2

- Method ......................................................................................................................... 22
  - Participants .................................................................................................................. 22
  - Materials and Procedure ............................................................................................ 23
- Results ........................................................................................................................... 24
- Discussion ...................................................................................................................... 28

## Study 3

- Method ......................................................................................................................... 30
  - Participants .................................................................................................................. 30
  - Materials and Procedure ............................................................................................ 31
- Results ........................................................................................................................... 33
- Discussion ...................................................................................................................... 36

## General Discussion

- ........................................................................................................................................... 38

## References

- ........................................................................................................................................... 43
List of Figures

Figure 1. Schematic of Study 1 design .......................................................................................... 14

Figure 2. Study 1 mean endorsement levels for the internalist and structural attributions for resulting prize distributions as a function of condition................................................................. 17

Figure 3. Study 1 mean endorsement levels for the internalist and structural attributions for resulting prize distributions as a function of age and condition. .................................................. 18

Figure 4. Study 2 mean endorsement levels for the explanations for why the character won the race.................................................................................................................................................. 27

Figure 5. Schematic of Study 3 design .......................................................................................... 33

Figure 6. Study 3 mean endorsement levels for three attributions for why the orange classroom won more prizes........................................................................................................................................... 35
Acknowledgements

First and foremost, I would like to express my gratitude and appreciation to my advisor Andy Baron, for his endless guidance, insights, and kindness. I am eternally thankful for your unwavering support and patience. Thank you for giving me the opportunity to learn from you.

Thank you to my committee members, Lauren Emberson and Darko Odic, for your encouragement, support, and thought-provoking conversations. Our initial meeting inspired me to come to UBC in the first place. Thank you for enriching my journey here with your presence and feedback.

This project would not have been possible without the help of the wonderful research assistants at the Social Cognitive Development Lab—Amanda Yang, Christiane Peretz, Esra Yarar, Fiona Guan, Golzar Ejadi, Ivy Jiang, Jocelyn Chun, Kayla Vo, Nicki Antone-Diamond, Nina Griggs, Victoria Cheah— I am so grateful for your work and dedication. A special thank you to our honors students, Matthew Tsui and Valerie Teng, for their invaluable insights and efforts.

In addition, I would like to thank all our participating children and families – we could not do this research without you!

To my cohort and dearest friends – Jessica Lee, Miranda Long, Nicole Stuart, Melanie Butt – thank you for your friendship and your endless support, I could not do this without you. Jess, I am forever thankful that I have you to share this journey with.

To the Developmental area, thank you, I don’t know what I would do without your advice, support, and friendship.

To my loved ones, Cypress Hall, Melissa Hall, Jason Hall, Cole Whitmore, and Tatiana Lau – thank you, I love you.
Introduction

Graduating from university is the leading determinant of upward social mobility in the United States (de Alva, 2019). Over three hundred years after the first American college opened its doors, Black Americans were given access to this opportunity too (*Brown v. Board of Education*, 1954). In response to the historical and ongoing structural barriers that have denied marginalized groups access to higher education, *affirmative action* has emerged as an integral—but highly debated—approach to address this discrimination. Critics of affirmative action argue that such policies are discriminatory, promote inequality, and undermine the merit-based college admissions process. Proponents of affirmative action point to societal structures that hinder academic attainment for Black Americans and other marginalized racial groups, including the legacy of slavery and segregation, redlining, underfunded schools, and wealth gaps resulting from housing and hiring discrimination (Herring & Henderson, 2016).

These opposing viewpoints are characterized by two distinct types of attributions: *internalist* and *structural*. Internalist attributions ascribe certain outcomes or discrepancies between people to internal traits of the individuals or groups. Take for example claims that academic achievement is an accurate reflection of a meritocratic system where an individual’s success is determined solely by their intelligence and efforts. In contrast, structural attributions emphasize how broader social structures limit individuals’ behavior and outcomes (Haslanger, 2006; Ross, 2023). For instance, contextualizing how academic achievement is shaped by disparities in educational resources, such as outdated textbooks, limited technology access, inadequate materials, and insufficient teacher training— and how these disparities align with racial and class divisions.
The type of attribution made for social inequality is consequential. Research indicates that attributing inequalities to internalist factors is associated with increased prejudice towards disadvantaged group members in both adults and children (Mandalaywla et al., 2018; Rhodes & Mandalaywala, 2017). Unfortunately, both adults and children are more inclined to attribute social inequalities to internalist, rather than structural, causes (Gelman, 2003; Gilbert & Malone, 1995). Given that beliefs about social inequalities form early in childhood (Shutts, 2015), it is crucial to understand how children reason about advantaged and disadvantaged groups. Exploring whether children can make structural attributions during the early stages of social categorization may help prevent or reduce the formation of intergroup bias. Therefore, this project investigated children's ability to make structural attributions when reasoning about unequal outcomes between social groups.

Internalist reasoning about social categories

When contemplating social categories or considering social inequalities, a general tendency has been observed for individuals to primarily focus on internal or inherent features when explaining perceived differences. This predisposition has been extensively studied and explained by various theoretical frameworks that converge on a similar understanding of an internalist bias (Cimpian & Salomon, 2014; Gelman, 2003; Gilbert & Malone, 1995; Hirschfeld, 1996; Vasilyeva & Lombrozo, 2019). One prominent framework is psychological essentialism, which describes the inclination of individuals to perceive certain categories or concepts as having underlying, unchangeable "essences" or fundamental features that define their inherent nature (Gelman, 2003). To illustrate, imagine a housecat which undergoes cosmetic surgery to look like a bunny and begins to behave like a bunny. Is the animal a cat or a bunny? An essentialist approach would claim that this animal, although it looks and behaves like a bunny, is
nevertheless still a cat due to a “true cat nature” that is beyond observable features or traits and gives the animal its identity.

One possible mechanism underlying essentialism is the inheritance heuristic, an implicit cognitive process that primarily attributes observed patterns in the world to the inherent features of their components (Cimpian & Salomon, 2014). For example, when considering the pattern that girls often wear pink, the inheritance heuristic suggests that people assume girls like pink because both "girls" and "pink" share overlapping inherent characteristics, such as pink being perceived as possessing feminine traits like softness or delicacy. Cimpian & Salomon (2014) propose that individuals tend to make these assumptions as a mental shortcut, seeking the most readily available information to construct a satisfactory explanation for the observed pattern. Since observable, surface-level traits are more accessible than other factors (such as historical precedent, marketing campaigns, or social norms), they are prioritized in our efforts to comprehend our environment. Once these traits become affixed to a certain category, they can be seen as contributing to the “essence” of the category, thus creating essentialist views.

Essentialist views often become problematic when they are applied to social kinds or used to explain social inequalities. Social essentialism is the belief that social categories (gender, race, socioeconomic status, sexuality, etc.) reflect meaningful, fundamental, and fixed differences between people (Hirschfeld, 1996; Rhodes & Moty, 2020). Consequently, social essentialism implies that members within a social category share fundamental similarities (which can contribute to stereotypes), as well as fundamentally differ from other groups (which can fuel prejudice and discrimination). Take for example the social essentialist view that women are inherently weaker and less dominant than men (Hentschel et al., 2019). This viewpoint perpetuates gender stereotypes and discrimination by assuming that certain innate qualities or
characteristics make men better leaders (Eagly & Karau, 2002), and contributes to gender bias in hiring and promotion decisions. In fact, in adults, essentializing social groups has been linked to higher endorsement of stereotypes (Bastian & Haslam, 2006) and increased racial prejudice (Mandalaywla et al., 2017).

Given the centrality of internalist defaults in human cognition and the role social essentialism plays in fostering prejudice and inequality, it is critically important that we understand the origins of essentialism itself. Essentialism has been proposed as an adaptive strategy which aids children in learning and reasoning about categories. Believing that categories are defined by “essences” allow children to make inductive inferences about members of a category without having to understand what those differences are (Gelman, 2004; Gelman & Markman, 1986; Medin & Ortony, 1989). In this way, a child could abstract that all bunnies hop and all cats pounce without having to understanding the differential physiological and biological drivers of those actions. This category-based inductive reasoning develops early, with infants as young as thirteen months old able to generalize information about one object to an entire category based on shape similarity or verbal labels (Graham et al., 2004). By the age of three, children can also extend new facts about one member of a category to all members in the case of animals, natural kinds, and social categories (Gelman & Markman, 1986; Heyman & Gelman, 2000).

However, inductive generalizations alone do not necessitate essentialist beliefs. Nevertheless, when children’s tendency to generalize over categories is partnered with strong beliefs in innate influence, essentialist views can arise. This inclination is best illustrated by “switched-at-birth” paradigms, which show that children tend to prioritize the influence of innate, over environmental, factors. In fact, four-year-olds expect kangaroos raised by goats to be
good at hopping over climbing (Gelman & Wellman, 1991) and by 5-years-old, expect that human children will speak the language of their birth, rather than adoptive, parents (Hirschfeld & Gelman, 1997). Critically, children view these underlying “essences” as causal forces and tend to attribute actions and outcomes to these internal traits (Gelman, 2002; Gelman, 2004).

Evidence of social essentialism emerges as early as infancy. At seven months of age, preverbal infants expect individual members of a novel social group to exhibit similar behaviors, indicating their belief in fundamental similarities among category members (Powell & Spelke, 2013). Throughout infancy, children develop the ability to socially categorize individuals based on factors such as race (Bar-Haim et al., 2006; Kelly et al., 2007), language (Pun et al., 2018), and gender (Quinn et al., 2002). By early childhood, they also demonstrate preferences for familiar and high-status groups (Baron & Banaji, 2006; Kelly et al., 2005; Qian et al., 2016; Shutts, 2015).

As children learn about social groups, social hierarchies, and begin to form social biases, holding social essentialist beliefs can contribute to increased bias in children (Rhodes & Mandalaywala, 2017). For instance, a study by Pauker et al. (2016) revealed that children’s essentialist beliefs about race predicted their endorsement of negative racial stereotypes. Furthermore, social essentialist views have been linked to own-race biases, wherein white children who engage in racial essentialist thinking demonstrate better memory for white faces compared to racially ambiguous and Black faces, whereas children who did not essentialize race exhibited better memory for racially ambiguous faces (Gaither et al., 2014). Moreover, research has shown that when children essentialize outgroup members, they tend to share fewer resources (Rhodes et al., 2018), display a greater desire for social distance, and ascribe less positive affect.
to the outgroup members compared to children who do not essentialize group membership (Diesendruck & Menahem, 2015).

**Structural reasoning about social categories**

While many adults and children value equality, internalist defaults and essentialized views cause them to overlook the structural factors that shape social and economic inequalities. The structural factors include, for example, the historic and current discrimination, marginalization, and exclusion of individuals from educational, occupational, residential, and social opportunities on the basis of gender, race, socioeconomic status (SES), sexuality, or other group memberships (Bonilla-Silva, 1997; Crenshaw, 1989; Massey & Denton, 1993; De la Roca et al., 2014; Corliss et al., 2011; for review, Pakulski, 2022). Consequently, these excluded social groups face limited prospects for upward mobility and are more susceptible to adverse emotional, educational, and health consequences (Crenshaw, 2017; Frost et al., 2015; Homan, 2019; Owens, 2018; Wilkinson & Pickett, 2017). Understanding and acknowledging how these structural factors contribute to social inequality may promote motivation for social change and decrease bias towards disadvantaged group members (Elenbaas et al., 2020, Piff et al., 2020; Roberts & Rizzo, 2021).

Structural attributions, specifically for social phenomena, acknowledge that we all exist within social structures (i.e., networks of social and institutional relations) that both shape and constrain individual and group level behaviors and outcomes (Haslanger, 2016). In other words, (social) structural attributions are simply explanations for individual or group outcomes which account for the complex social structures individuals are nested in and their position in that network. In this way, structural attributions have similarities with causal attributions to environmental, extrinsic, or external factors but refer specifically to social structures as the
causal motivating or constraining variables which serve to systematically advantage some social
groups and disadvantage others (Amemiya et al., 2022; Ross, 2023).

To effectively attribute social inequality to structural factors, three key elements are
necessary: a) awareness of the structural factors involved, b) understanding of how these factors
contribute to observed inequality, and c) motivation to make structural attributions (Amemiya et
al., 2022; Elenbass et al., 2020). From early in development children demonstrate sensitivity to
social inequality, particularly regarding wealth disparities and differential access to resources
(Hazelbacker et al., 2018; Sommerville & Ziv, 2018). By adolescence, American children begin
associating wealth disparities with race, perceiving Asian and white Americans as having higher
income and wealth compared to Black and Latinx Americans (Ghavami & Mistry, 2019).
Children additionally possess the capacity to recognize and differentiate between the sources of
the inequitable resource distributions they observe (Elenbaas, 2019; Rizzo & Killen, 2020). For
example, in a study by Rizzo et al. (2020), three-year-old children successfully distinguished
between an individual merit-based inequality (where one child worked harder and earned more
prizes) and a structural bias-based inequality (where the individual distributing prizes showed
preferential treatment toward their gender ingroup). The children judged the former as fair and
the latter as unfair, which in turn prompted a desire for rectification.

However, despite children's ability to recognize inequality and differentiate between
different sources of inequality, it remains unclear whether they can make structural attributions
for social inequality. For instance, in a study where 4-to-8-year-old children were asked to
explain social inequities between two unfamiliar groups, they showed a preference for intrinsic
factors (such as greater intelligence or work ethic) over extrinsic factors (such as previous
victory in a conflict or windfall resources; Hussak & Cimpian, 2015). However, considering that
even young children possess advanced causal reasoning abilities (Koslowski & Masnick, 2010; Muentener & Bonawitz, 2017) and can consider and apply multiple potential causes simultaneously (Goddu et al., 2021), it seems likely that children do have the capability to make structural attributions, but fail to do so given a lack of practice, limited exposure to social learning models, and restricted access to information about the complex historical and social factors that underlie social inequalities.

Recent research conducted by Vasilyeva et al. (2018) provides compelling evidence in support of this claim. In their study, 3-to-6-year-old participants were first introduced to a school where boys and girls had separate classrooms and given the information that during recess the girls predominantly played “Yellow-Ball” while the boys predominantly played “Green-Ball”. Next, participants were told that the game each student played was determined by tossing a pebble toward two buckets set up in each classroom: if the pebble fell into a green bucket, the child played Green-Ball, if the pebble fell into a yellow bucket, the child played Yellow-Ball. Critically, participants were either shown a bucket set up which favored either internalist or structural attributions. In the internalist framing the green and yellow buckets were identical sizes in each classroom, thus suggesting the observed distribution of boys and girls playing the games was due to preference. However, in the structural framing, the boys’ classroom had a comparatively larger green bucket to their yellow bucket, with the reverse in the girls’ classroom. The structural set up thus introduced structural constraints and sought to motivate participants to formulate and endorse structural attributions for the observed distribution.

Vasilyeva et al. (2018) found that children as young as 3 years old were able to recognize the structural constraint and self-generate some structural attributions for the unequal distribution, and by 5 years of age reliably endorsed provided structural attributions for the
inequality. Furthermore, work by Peretz-Lange & Muentener (2021) found that while 5-to-6-year-old children preferentially generated internalist attributions over structural attributions for an observed disparity between two novel groups, verbally highlighting structural constraints led to increased structural attributions. Taken together, these findings suggest children do have the ability to make structural attributions, albeit not necessarily exclusively or consistently.

However, despite being aware of and considering structural factors, children may still prioritize internalist attributions for inequality (Amemiya et al., 2022; Yang et al., 2020). It can be argued that making structural attributions is cognitively more challenging (Cimpian & Salomon, 2014), and thus, children would require additional motivation to prioritize structural factors and selectively make structural attributions. According to a framework proposed by Amemiya et al. (2022), individuals must perceive that structural factors have a discernible impact on observed inequality above-and-beyond what can be attributed to internal factors alone. They suggest that individuals can evaluate this criterion by engaging in counterfactual reasoning—for instance, considering whether the inequality would change if the social structure were different. Since children are capable of counterfactual reasoning (German & Nichols, 2003; Harris et al., 1996) and often motivated to promote equality (Burns & Sommerville, 2014; Geraci & Surian, 2011; Rizzo & Killen, 2020), there is evidence to suggest that with appropriate scaffolding and modeling, even young children can develop the ability and motivation to make structural attributions for inequality. Supporting this claim, studies have shown that discussing discrimination and encouraging civic engagement with adolescents can increase their structural attributions for racism (Bañales et al., 2019) and poverty (Flanagan et al., 2014).

Understanding children’s capacity for structural reasoning and encouraging its development is critically important when considering structural attributions role in bias
mitigation. The logical connection between attribution and bias can easily be illustrated with an example: First, consider an individual who lost their job due to repeated tardiness and missed shifts. Imagine this tardiness is due to low work ethic and lack of responsibility. In this case, one might think ill of this individual’s character, view the job loss as a natural and fair consequence of their actions, and not be motivated to change this outcome. Now imagine this tardiness is due to the termination of the only bus route that would allow this individual to make it to work on time. In addition, this cut to transit funding is the result of discriminatory redlining practices, which have historically denied public services to our protagonist’s neighborhood. Under these circumstances, one might place no evaluative judgments on this individual’s character, view the job loss as an unjust consequence, and be motivated to change the factors which resulted in this outcome. When these lines of thinking are scaled up to group level identities and outcomes, it is evident how internalist attributions can result in prejudice, whereas structural attributions can motivate individuals to pursue social change.

Indeed, in adults, attributing economic inequality to structural, rather than internalist, factors increased egalitarian views and promoted inequality-reducing behavior (Piff et al., 2020). In children, attributing race-based economic inequality to structural factors (i.e., “because of things that happen in the world”) over internalist factors (i.e., “because of who they are on the inside”) was associated with no increase in pro-white/anti-Black bias over six-months, while favoring the internalist attribution was connected to a marked increase in racial bias (Rizzo et al., 2022). As social biases have been found to be more malleable in childhood compared to adulthood (Block et al., 2022; Forscher et al., 2019; Gonzalez et al., 2021; Lai et al., 2014; Qian et al., 2019), it is likely that the earlier structural reasoning can be fostered and utilized the more successful structural attributions will be at mitigating bias.
The present study

Given the potential benefits of promoting structural reasoning during childhood, the present studies investigated children’s ability to correctly make structural attributions for social inequalities, as well as the factors that motivate children to prioritize structural causes. Study 1 is a modified conceptual replication of Vasilyeva et al. (2018) which utilized the developmentally friendly “Bucket Game” example of structural inequality to assess children’s structural reasoning. The goals of Study 1 were three-fold. First, we aimed to conceptually replicate the findings of Vasilyeva et al. (2018) that children can make structural attributions for inequality in a familiar and more naturalistic context (i.e., winning prizes in a game). Second, we explored if this more familiar set-up would aid younger children in recognizing structural factors and reasoning about their effects on outcomes. Third, in line with the logic of the difference-making framework put forward by Amemiya et al. (2022), we aimed to see if the presence of a structural constraint creating inequality for a desirable resource (prizes) would motivate children to make structural attributions in comparison to a scenario in which merely a structural factor creates an equitable outcome.

We predicted that by age three children would be able to endorse structural attributions for both scenarios as good ideas. However, we predicted children would default to internalist reasoning and preferentially endorse the internalist attribution for the equitable outcome. In contrast, we predicted the inequitable outcome would motivate children to preferentially attend to the structural factors at play and prefer the structural attribution. Studies 2 and 3 address potential confounds and limitations present in Study 1 (and previous work), while further examining children’s ability to selectively endorse structural attributions.
Study 1

Method

Participants

Participants included one hundred forty-six 3-to-4-year-olds (mean age 4.17 years, range 3.00-4.99; 68 girls and 78 boys), one hundred twenty-one 5-to-6-year-olds (mean age 6.02 years, range 5.00-6.99; 52 girls, 69 boys), and one hundred eighteen 7-to-8-year-olds (mean age 7.88 years, range 7.00-8.95; 59 girls, 59 boys), who were recruited at a science center in a large North American city in the Pacific Northwest. An additional 65 participants were recruited but then excluded from analysis due to: parental/sibling interference (12), not completing the study (10), distraction (10), experimenter error/technical issue (11), falling outside the age range (7), English comprehension (4), and failing study comprehension checks (15). Participant race/ethnicity was provided by parents for 355 participants (108 East Asian, 51 South Asian, 4 Southeast Asian, 35 Biracial/Multiracial, 10 Black, 1 Caribbean, 1 Indigenous, 8 Middle Eastern/Arab, 137 White).

Research assistants recruited participants as they walked around the science center exhibits or entered the lab space on their own and asked to participate. Parents provided written consent for their child’s study participation; children also provided verbal assent before beginning study procedures. Participants were randomly assigned to either view the equitable outcome (equality condition) or the inequitable outcome (inequality condition) in a counterbalanced order. All participants received a sticker for their participation. Our study protocol received ethics approval from the University of British Columbia, H10-0047, “The Development of Social Cognition.”

Materials and Procedure

Participants were tested individually in a designated lab space. All participants were read an illustrated vignette (created on Microsoft PowerPoint and presented on desktop computer)
which introduced two gender segregated classrooms which were both to play a fictitious game (the “Bucket Game”) to win prizes. Each classroom contained 5 students, matched for race/ethnicity (2 East Asian, 2 White/Caucasian, and 1 Black character). The “Bucket Game” was adapted from Vasilyeva et al. (2018). Participants were told that in each classroom a teacher had set out two buckets beside each other, one yellow and one green, and to play the game the students would take turns trying to throw a pebble into one of the buckets. In the rules of the game, if the pebble fell into the green bucket, the student won a desirable prize (basket of toys and treats), but if the pebble fell into the yellow bucket, the student won a less desirable prize (a bowl of carrots; See Figure 1).

The critical manipulation was the relative size of the buckets between the classrooms. In the equality condition, both classrooms’ buckets were of equal size. In the inequality condition, one classroom had a relatively larger green bucket than their yellow bucket (advantaged classroom), whereas the other classroom had a relatively larger yellow bucket compared to their green bucket (disadvantaged classroom). The unequal bucket sizes created a structural inequality between the classrooms, placing one group of students at an advantage for winning the desirable prize. The gender of the advantaged classroom was counterbalanced across participants.

Participants were then shown the resulting prize distribution in each classroom. In the equality condition, each classroom received the same prize distribution (7:8 desirable-to-undesirable). In the inequality condition, the advantage classroom had more baskets of toys and treats (12:3 desirable-to-undesirable) and the disadvantaged classroom had more bowls of carrots (3:12 desirable-to-undesirable).
Attribution endorsement. Next, participants were introduced to two animal puppets who had both also seen the students play the Bucket Game and who each had different ideas about the game. We had animal puppets put forth the different attributions instead of the experimenter or characters from the story to avoid influence of authority, familiarity, or ingroup preference on children’s answers (Aldan & Soley, 2019; Chen et al., 2013; Lucas et al., 2017; Ma & Woolley, 2013; Meltzoff, 2007; Woolley et al., 2021). Participants were asked to evaluate each idea using a two-step, 4-point scale: they first chose one of two thumbs representing either a “good idea” (thumbs up) or a “bad idea” (thumbs down), and then they chose between two subsequent options based on their previous choice: “a little good/bad” (small thumb) or “very good/bad” (big thumb)—a scale previously used in similar work (Vasilyeva et al., 2018; Yang et al., 2021).
In the equality condition, the puppets proposed that the equal prize distribution resulted because either: “Boys and girls like bowls of carrots and baskets of toys and treats equally” (internalist attribution) or “It was equally easy to throw the pebble into each bucket” (structural attribution). In the inequality condition, the puppets proposed that unequal prize distribution (“boys/girls’ classroom has more bowls of carrots”) resulted because either: “Girls/boys like carrots more than boys/girls do” (internalist attribution) or “It was easier to throw the pebble into the bigger bucket” (structural attribution). The internalist attribution was selected as young children have been found to use statistical prevalence to infer the preferences of others (Kushnir et al., 2010; Ma & Xu, 2011) and also infer preferences of others based on a singular action (Pesowski et al., 2016). Attributions were presented in a counterbalanced order across participants.

**Structural constraint recognition.** Finally, participants were asked two questions to ensure the child perceived the structural inequality and understood how a larger green bucket advantaged one classroom. Participants were asked if they were to play the bucket game: 1) which prize they would rather win (a basket of toys and treats or a bowl of carrots) and 2) which classroom they would prefer to play in (advantaged or disadvantaged classroom). For the second question all participants, regardless of condition, saw an unequal game set up (one bucket larger than the other). If children recognized the structural factor and were aware how the different sized buckets created unequal odds of success, children should pick the classroom in which the bucket for their desired prize was larger.

**Results**

Unless otherwise noted, preliminary analyses revealed that there was no effect of order (which explanation was presented first), participant gender, participants own preferred prize, or
if the participant’s gender ingroup was advantaged or disadvantaged for the participants own preferred prize, and thus these factors were not discussed further.

**Attributions.** A 2 (Condition: Equality or Inequality) X 2 (Attributions Type: Internalist or Structural) X 3 (Age Group: Younger (3-to-4-year-olds), Middle (5-to-6-year-olds), and Older (7-to-8-year-olds)), three-way mixed analysis of variance (ANOVA) revealed a significant interaction between condition and attribution type, $F(1, 384)= 11.87, p < .001$, such that between conditions, participants who viewed the unfair game endorsed the internalist attribution significantly less than the participants who viewed the fair game, $t(384) = 3.73, p < .001$. Contrary to our hypotheses, there were no differences between conditions in endorsement of the structural attribution, $t(384) = -1.29, p = .20$. Additionally, within conditions, participants who viewed the fair game defaulted to internalist reasoning and endorsed the internalist attribution significantly more than the structural attribution, $t(384) = 3.19, p < .01$. There were no significant differences in endorsement of the structural and internalist attribution for children who viewed the unfair game, $t(384) = -1.64, p = .10$. See *Figure 2.*

Results also revealed an interaction between age and attribution type, such that, regardless of if children saw a structurally fair or unfair game, participants were less likely to endorse the internalist attribution as they aged, $F(2, 384)= 10.42, p < .001$. Specifically, 7-to-8-year old children were less likely then both 3-to-4-year-old children, $t(384) = 4.84, p < .001$, and 5-to-6-year-old children, $t(384) = 4.44, p < .001$, to endorse the internalist attribution.
Figure 2. Study 1 mean endorsement levels for the internalist and structural attributions for resulting prize distributions as a function of condition. Responses were scored on a scale of “very bad” (-2) to “very good” (2). The dashed line indicates chance level responding. Error bars indicate 95% confidence intervals.

Exploratory analyses examining endorsement patterns by age group reveal that both 3-to-4-year-old children, $t(146) = 2.24, p < .05$, and 5-to-6-year-old children, $t(122) = 3.15, p < .01$, in the equality condition endorsed the internalist attribution over the structural. However, both younger age groups in the inequality condition endorsed both attributions to similar degrees, $ts < 1.51, ps > .13$. Older children differed in their response patterns, in that 7-to-8-year-olds who viewed the equal outcome did not prefer the internalist attribution to the structural attribution, $t(114) = 0.30, p = .76$, but endorsed the structural over the internalist attribution for the outcome of the unfair game, $t(116) = -4.67, p < .001$.

Additional follow-up testing utilizing a Chi-square goodness-of-fit test was conducted to ascertain if participants endorsed or rejected the attributions in proportions different from chance.
on a binary measure. Children of all ages, in both conditions, significantly endorsed both the internalist and structural attributions, \( \chi^2 > 6.56, ps < .01 \), save for 5-to-6-year-old children in the equality condition who endorsed the structural attribution at chance levels, \( \chi^2(1, N = 58) = 2.48, p = .12 \), and 7-to-8-year-old children in the inequality condition who came close to rejecting the internalist attribution, \( \chi^2(1, N = 63) = 3.57, p = .059 \). See Figure 3.

**Figure 3.** Study 1 mean endorsement levels for the internalist and structural attributions for resulting prize distributions as a function of age and condition. Responses were scored on a scale of “very bad” (-2) to “very good” (2). The dashed line indicates chance level responding. Error bars indicate 95% confidence intervals.

**Recognition of structural constraints.** Next, we evaluated our hypothesis that children would be able to recognize structural constraints and make personal choices to structurally advantage themselves. To do so, we first conducted separate exact binomial tests for each condition. Responses were coded as either choosing the classroom which advantaged the participant for their desired prize (“1”) or the classroom which disadvantaged them (“0”). Results

18
reveal that children in both the control and structural condition successfully selected the advantageous classroom, demonstrating both recognition and consideration of structural constraints. Children in the control condition selected the advantageous classroom (N=181, K=114) 64% of the time and children who viewed the structurally unfair game were slightly more successful in selecting the advantageous classroom (N=209, K=148) at 71% of choices, \( ps < .001 \).

To examine possible influences on participant’s ability to recognize the structural constraint and select the advantageous classroom we next conducted a logistic regression with condition (Equality vs. Inequality), age group (3-to-4, 5-to-6, 7-to-8), own gender advantage (participant’s gender ingroup depicted as structurally advanced vs disadvantaged for participant’s desired prize), and participant gender (girl vs boy) as moderators. The model revealed a significant main effect of advantage, \( \chi^2(388) = 78.72, p < .001 \), with participants whose own gender was advantaged more likely select the advantageous classroom than children whose gender ingroup was at a structural disadvantage for their desired prize (OR=6.18, 95% CI [3.55, 10.76]). This result is likely due to a conflict between gender in-group loyalty and structural advantage for participants who viewed a disadvantaged set-up. Follow up binomial tests show that when children’s own gender is depicted as being disadvantaged for the prize they want, their responses are equally characterized by selecting the structurally advantageous room and the classroom of their gender in-group, \( p = .298, \) CI 95% [0.47, 0.62]. The age, gender, or condition of the participant did not impact their ability to select the structurally advantageous classroom, \( ps > .053 \).

**Discussion**
Results provide evidence that by age three, children can recognize structural constraints and reason about how these constraints affect outcomes. Nearly all participants endorsed structural attributions as good ideas and were able to select the advantageous classroom for the prize they desired, demonstrating structural reasoning. In the absence of a structural constraint, 3-to-6-year-old children defaulted to internalist reasoning, endorsing the internalist attribution over the structural attribution. The presence of a structural constraint encouraged older children (7-to-8-year-olds) to overcome internalist reasoning and endorse the structural attribution over the internalist attribution. These findings both confirm and contradict previous work (Vasilyeva et al., 2018; Yang et al., 2021) in that there is evidence for structural reasoning by age three but failure to replicate the previous finding that children can overcome internalist defaults by age five in similar paradigms.

Contrary to our initial hypotheses, which predicted that the presence of a structural constraint creating an unequal opportunity to win a desired prize would motivate children to think structurally and increase their likelihood of endorsing structural attributions, we found no significant difference in the endorsement of structural attributions across conditions. While it is possible that inequality and structural constraints do not lead children to prioritize structural factors, the consistently high levels of endorsement (at or above chance levels) for all types of attributions, across all children and in both conditions, raise the possibility of an artificial ceiling effect masking their effect.

Previous research has indicated that children between the ages of 2 and 5 can exhibit a propensity for "yes-biases," meaning they have a tendency to disproportionately or incorrectly respond "yes" to yes-or-no questions (Fritzley et al., 2012; Okanda & Itakura, 2007; Peterson et al., 1999). This same issue may be influencing the results observed in our study, with children
demonstrating a bias to answer in the affirmative and rate attributions as "good ideas." This pattern becomes evident upon closer examination of the two previous studies that employed a similar paradigm. In Vasilyeva et al. (2018), participants evaluated internalist, structural, and incidental (non-causal) explanations for an unequal distribution from an identical bucket game set up. 3-to-6-year-old children endorsed all three attributions as good ideas at or above chance levels. To stress, this means that children rated the non-causal explanation that the unequal outcome occurred because "[the children] got sprinkled with water" as a good idea to a similar degree as they endorsed the internal and structural attributions. Similarly, in Yang et al. (2021), which closely replicated Vasilyeva et al. (2018), 4-to-5-year-old children endorsed all three explanations (internalist, structural, and incidental/non-causal) above chance levels, and it was only by the age of seven that they were able to reject the non-causal attribution. In Study 2, we aim to investigate whether the current measurement of children's internalist and structural reasoning, which asks participants to rate attributions as either good or bad ideas, may be creating an artificial ceiling effect driven by young children's "yes biases."
Study 2

To investigate the potential influence of "yes-biases" on the findings of Study 1 and previous research, Study 2 compared the endorsement of causal explanations by 3-to-6-year-olds using two different types of response language. Recognizing that "yes biases" may arise due to children defaulting to affirmative responses, a negative question response language was introduced as a means to overcome this default. A negative question is grammatically structured in a way that requires a "no" response to indicate agreement and a "yes" response to indicate disagreement. A common example is a statement like "Would you mind helping me tomorrow?", for in order to express willingness to help (affirmative response), one would say "No, I don't mind," while indicating unwillingness to help (negative response) would be expressed as "Yes, I do mind."

If the previous findings were indeed influenced by "yes-biases," we would expect to see children endorse all proposed explanations, even non-causal or illogical explanations, for a simple cause-and-effect scenario when asked to rate explanations as “good” or “bad ideas”. In contrast, if “yes-biases” are influencing children’s endorsements, we would expect to see relatively lower levels of endorsement and successful rejection of incorrect explanations when asked to rate ideas as either “silly” or “not silly”. A simple causal scenario was utilized instead of one governed by structural factors to avoid any uncertainties regarding children's reasoning abilities, as while the ability of 3-to-6-year-olds to reason structurally remains a topic of interest, a substantial and longstanding body of evidence supports the notion that children of this age should possess the capacity to reason about direct cause-and-effect scenarios (Sobel and Kirkham, 2006; for review, Muentener & Bonawitz, 2017).

Method
Participants

Participants included fifty-three 3-to-4-year-olds (mean age 3.99 years, range 3.00-4.94; 27 girls and 26 boys), and fifty-seven 5-to-6-year-olds (mean age 5.82 years, range 5.00-6.94; 32 girls, 23 boys, and 2 not specified), who were recruited at a science center in a large North American city in the Pacific Northwest. An additional 15 participants were recruited but then excluded from analysis due to: parental/sibling interference (4), not completing the study (4), experimenter error/technical issue (3), falling outside the age range (3), and English comprehension (1). Participant race/ethnicity was provided by parents for 103 participants (25 East Asian, 8 South Asian, 8 Southeast Asian, 19 Biracial/Multiracial, 1 Black, 3 Indigenous, 1 Latinx/Hispanic, 38 White). Recruitment, consent, and participant compensation procedures were identical to those of Study 1. Participants were sequentially assigned to conditions, with the first 57 participants assigned to use the previous “good vs bad idea” response language (replication condition) and the latter 53 participants to use the negative question response language (negative question condition). Data collection for both conditions was completed over a span of three months and preliminary analyses show no significant differences between the conditions in regard to mean age per age group, gender ratio, or ethnicity demographics.

Materials and Procedure

Participants were tested individually in quiet sections of the science center exhibits. All participants were read an illustrated vignette (presented on an iPad using the online survey software Qualtrics), which introduced two girl characters who were to race against each other. The characters were of the same depicted race and differed only in shirt color, hair style, and name. Participants were told that one character won the race and then were either asked an open-ended question on why they believed the character had won over the other then rate three
provided explanations (*negative question condition*), or only rate the explanations (*replication condition*). Open-ended responses were manually recorded verbatim by the experimenter. Then, regardless of their own explanation, we asked participants to assess three proposed explanations offered by three different animal puppets, differentiated by kind and color. The puppets stated either that the character won the race “because she is faster” (*causal explanation*), “because she is wearing an orange shirt” (*non-causal explanation*), or “because the sky is green” (*illogical explanation*) in a counterbalanced order.

Critically, half of the participants evaluated the explanations using the response language of Study 1 and the other half utilized a negative question response language. For the replication response language, participants evaluated each puppet’s explanation as either a good idea or a bad idea—a methodology mirroring Study 1 and Vasilyeva et al. (2018). Children could answer verbally or choose one of two thumbs representing a “good idea” (thumbs up) or a “bad idea” (thumbs down). For the negative question response language, participants evaluated if each explanation was or was not a “silly idea.” Children could answer verbally or choose one of two thumbs representing “silly” (thumbs up) or “not silly” (thumbs down). This response language necessities children respond in the negative (“not silly”) if they wished to endorse the explanation, thus circumventing possible yes biases (Fritzley et al., 2012). In addition, participants in the negative question condition were asked to select which puppet’s explanation they thought best after the individual evaluations. Children were verbally reminded, and saw representative symbols, of each explanation before responding. The representative symbols for each explanation were: a running figure (causal), an orange shirt (non-causal), and a green cloud (illogical).

**Results**
**Open-Ended Responses.** Of the subset of 53 participants who were prompted to provide a verbal explanation for why they believed the character won the race, only 20 children provided codable responses (remaining 33 stated they did not know or remained silent). Of the 20 participants, 65% (13) generated causal explanations (e.g., “Faster”, “Taller”, “Worked really hard”), 30% (6) generated non-causal, circular explanations (e.g., “Crossed the finish line first”, “Because she won the race”, “Went in front of everyone”), and 5% (1) generated an illogical explanation (“Ate someone’s hands”).

**Explanations.** A Chi-square goodness-of-fit test was performed to ascertain if children endorsed or rejected the three explanations at proportions different than chance on a binary measure. For analysis, responses were coded as either endorsed (“1”) or rejected (“0”). Thus, in the replication condition responses of “good idea” counted as endorsement, and “bad idea” as rejection, and negative question responses of “not silly” were scored as endorsement and “silly” as rejection. For younger children in replication condition, children endorsed both the causal ($\chi^2(1, N = 27) = 4.5, p < .05$) and non-causal ($\chi^2(1, N = 27) = 6.3, p = .01$) explanations greater than chance levels, but did not differ from chance in their endorsement of the illogical explanation $\chi^2(1, N = 27) = 0.04, p = .85$. Older children in the replication condition differed from chance only in their endorsement of the causal explanation ($\chi^2(1, N = 28) = 7.0, p < .01$), responding at chance levels for both the non-causal ($\chi^2(1, N = 28) = 0.6, p = .45$) and illogical explanations ($\chi^2(1, N = 28) = 0, p = 1$). When responding using the negative question response language, younger children rejected both the non-causal ($\chi^2(1, N = 25) = 11.56, p < .001$) and illogical explanations ($\chi^2(1, N = 25) = 4.8, p < .05$), but were only at chance levels for endorsement of the causal explanation ($\chi^2(1, N = 25) = 2.0, p = .16$). However, older children were able to successfully endorse the causal explanation ($\chi^2(1, N = 25) = 4.8, p < .05$), while
rejecting the incorrect non-causal ($\chi^2(1, N = 25) = 14.4, p < .001$) and illogical ($\chi^2(1, N = 25) = 9, p < .01$) explanations, demonstrating that by age five, children are able to successfully reason about cause-and-effect scenarios, and that utilizing a negative question response language paradigm better captures this reasoning ability than previous methods. See Figure 4.

A 2 (Condition: Replication or Negative Question) X 3 (Explanation Type: Casual, Non-Causal, Illogical) X 2 (Age Group: Younger (3-to-4-year-olds), Older (5-to-6-year-olds)) three-way mixed analysis of variance (ANOVA) revealed a main effect of condition, such that children in the negative question condition were less likely to endorse explanations overall than children in the replication condition, $F(1, 101)= 29.79, p < .001$. Across conditions, children were more likely to endorse the causal explanation than both the non-causal and illogical explanations, $F(2, 190)= 12.88, p < .001$, and this effect strengthened with age, $t(101) = -2.24, p < .05$. Collapsed across ages, children’s endorsement of each explanation was affected by the response language used, $F(2, 190)= 2.38, p = .1$. In the replication condition, children were more likely to endorse the causal explanation than the illogical explanation, $t(101) = 2.41, p < .05$. In the negative question condition, children were more likely to endorse the causal explanation than both the non-causal, $t(101) = 4.44, p < .001$, and illogical explanations, $t(101) = 3.16, p < .01$. 
Forced choice election. A Chi-square goodness-of-fit test was performed to ascertain if the subset of participants (n=52) tasked with selecting the best explanation of the three after individual endorsements, differed from chance in their selection. Results reveal that both younger, $\chi^2(2, N = 25) = 13.76, p < .01$, and older children, $\chi^2(2, N = 27) = 14, p < .001$, significantly differed in chance in their selection. To determine which specific choices contributed to this significant difference, a closer examination of residuals was undertaken. Based on a predetermined significance level of $\alpha = 0.05$ and two degrees of freedom, the critical value for significance was determined to be ±1.96. Choices with residuals exceeding this critical value in either a positive or negative direction were considered significantly different from chance levels. Among the explanations, only the causal explanation exhibited a substantial

![Figure 4](image)

*Figure 4.* Study 2 mean endorsement levels for the explanations for why the character won the race. For graphical depiction, responses were scored as either endorsed (“1”) or rejected (“-1”). The dashed line indicates chance level responding. Negative scores indicate rejection of the explanation while positive scores indicate endorsement. Error bars indicate 95% confidence intervals.
positive residual ($r = 3.67$ for both), revealing that irrespective of age, participants preferred the causal explanation, selecting it as the best explanation $68\%$ (younger children) and $67\%$ (older children) of the time respectively.

**Discussion**

The way we ask children to evaluate causal attributions can influence their ability to demonstrate their reasoning. The findings of Study 2 shed light on the potential impact of "yes-biases" on previously employed methods (as seen in Study 1, Vasilyeva et al., 2018, and Yang et al., 2021) to assess children's causal reasoning about internalist and structural factors. When participants were asked to rate causal attributions as either "good ideas" or "bad ideas", children responded at or above chance levels for all attributions, including illogical attributions such as a character winning a race because the sky is green. However, when utilizing negative question response language and asking children to rate each attribution as either “silly” or “not silly”, 3-to-4-year-olds successfully rejected incorrect attributions, while 5-to-6-year-olds both rejected incorrect attributions and significantly endorsed correct causal attributions.

Moreover, although younger children responded at chance levels when evaluating causal attributions individually, 3-to-4-year-olds were equally as successful at selecting the causal attribution as the best explanation as 5-to-6-year-old participants. These findings suggest that while individually rating causal attributions may present challenges for young children, ranking attributions or identifying the superior attribution may be more developmentally accessible.

Together the results of Study 2 offer evidence for a novel measure of children's causal attributions that addresses previous concerns related to "yes-biases" and better captures children's causal reasoning. Study 3 utilizes this approach, incorporating negative question response language for rating attributions along with an additional task of selecting the preferred
attribution, to re-examine children's ability to endorse internalist and structural factors as causal attributions.
Study 3

Study 3 is a close conceptual replication of Vasilyeva et al. (2018) with four critical differences: 1) children rated attributions using negative question response language; 2) in the story depicting structural inequality, classrooms were divided by color, not gender; 3) the classrooms each only had one bucket, not two; and 4) as in Study 1, the outcome of the game was winning different prizes, not assignment to another activity. We decided to divide the classrooms by color instead of gender in light of the gender ingroup-biases observed in Study 1 which confounded our measure of recognition of structural inequality. We moved from two buckets to one bucket as it allowed for a more naturalistic event (ball either goes into the bucket or misses, as opposed to falling into one of two buckets) to further aid younger children’s understanding of the scenario. As in Study 1, we kept our distribution outcome prizes instead of activity assignment to make the scenario more naturalistic and comprehensible for young children.

With this study we investigated the development of children’s structural reasoning about inequality. We aimed to assess if children could selectively make structural attributions for an inequality resulting from structural factors, without influence of artificial inflation from “yes-biases”. We predict that by age three, children can recognize structural constraints and accurately map structural and internalist attributions to appropriate scenarios. In addition, by age five, we predict children will be able to overcome internalist defaults and preferentially rate structural attributions for inequality.

Method

Participants
Participants included fifty-two 3-to-4-year-olds (mean age 4.16 years, range 3.01-4.98; 24 girls, 27 boys, and 1 unspecified), fifty-three 5-to-6-year-olds (mean age 6.02 years, range 5.00-6.98; 26 girls, 26 boys, and 1 unspecified), and fifty 7-to-8-year-olds (mean age 7.93 years, range 7.00-8.99; 25 girls and 25 boys), who were recruited at a science center in a large North American city in the Pacific Northwest. An additional eight participants were recruited but then excluded from analysis due to: not completing the study (1), experimenter error/technical issue (2), falling outside the age range (2), and comprehension issues (3). Participant race/ethnicity was provided by parents for 149 participants (32 East Asian, 11 South Asian, 1 Southeast Asian, 17 Biracial/Multiracial, 2 Black, 1 Indigenous, 5 Latinx/Hispanic, 1 Middle Eastern/Arab, 79 White). Recruitment, consent, and participant compensation procedures were identical to those of the previous two studies. Participants were randomly assigned to either the equal condition (buckets equal sizes in each classroom) or unequal condition (buckets different sizes in each classroom) in a counterbalanced order.

Materials and Procedure

Participants were tested individually in a designated lab space. All participants were read an illustrated vignette (created on Microsoft PowerPoint and presented on desktop computer) which introduced two classrooms in a school which were both to play a fictitious game, the “Bucket Game”, to win prizes. Each classroom contained 6 students, matched for gender (3 girl and 3 boy characters) and race/ethnicity (2 Latinx/Hispanic, 2 White/Caucasian, 1 East Asian, and 1 Black character). Student characters differed between the classrooms in hairstyles and shirt colors (representative of their classrooms -- orange classroom wore orange shirts; green classroom wore green shirts).
The “Bucket Game” from Study 1 was modified for Study 3. Participants were told that to play the game, the students each took turns trying to throw a ball into a bucket. If the ball went into the bucket, the student won a desirable prize (basket of toys and treats). If the student missed and the ball did not go into the bucket, the student won a less desirable consolation prize (a toothbrush). The critical manipulation was the relative size of the buckets between the classrooms. In the unequal condition, the orange classroom’s bucket was substantially larger than the green classroom’s bucket, thus introducing a structural inequality between the classrooms. In the equal condition, both classrooms’ buckets were of equal size. All participants were shown the green and orange classrooms side by side and asked to select in which classroom they would like to play the game in if they were to play the game themselves. See Figure 5.

In both conditions, participants were shown an unequal resulting prize distribution in which the orange classroom won more desirable prizes than the green classroom. Participants were then asked to assess three proposed attributions as to why the orange classroom had more baskets of toys and treats than the green classroom. The three attributions were offered by three different animal puppets, differentiated by kind and color. The puppets stated either that the orange classroom had more “because the kids in the orange classroom are taller” (internalist attribution), “because it was easier to get the ball into the orange bucket” (structural attribution), or “because frogs are orange” (illogical attribution) in a counterbalanced order. The internalist attribution was selected as children often point to physical attributes as internal causal factors (Gelman, 2004; Inagaki & Hantao, 2004) and expect physically larger individuals to prevail in conflict and exhibit higher competence in games (Charafeddine et al., 2015; Thomsen et al., 2011). Participants evaluated each attribution using the negative question response language from Study 2 (“Do you think that’s a silly idea?”) and responded either verbally or by choosing
one of two thumbs representing “silly” (thumbs up) or “not silly” (thumbs down). Then participants were verbally reminded and shown representative symbols of all attributions and asked to select the best attribution for why the orange classroom had more baskets of toys and treats. Representative symbols for each attribution were: orange classroom characters’ faces (internalist), the orange bucket (structural), and an orange frog (illogical).

**Results**

**Recognition of structural inequality.** Chi-square goodness-of-fit tests were performed to ascertain 1) if participants in the unequal condition automatically encoded the structural inequality, and 2) if participants in the equal condition demonstrated a color preference between orange and green. Regardless of age, children in the equal condition did not differ from chance in their selection between the two classrooms, $\chi^2 < 1.38, ps > .24$. For children who viewed the structurally unequal game, younger children did not seem to understand that the structural inequality disadvantaged the green classroom (or perhaps did but was unable to import that understanding into their classroom choice), $\chi^2(1, N = 26) = 1.38, p = .239$. However, by age five children reliably selected the orange classroom over the green classroom, $\chi^2 > 7.54, ps < .01$.  

![Figure 5. Schematic of Study 3 design: (a) Depiction of classroom setup, bucket size, and resulting prize distribution in the equal condition; (b) Depiction of classroom setup, bucket size, and resulting prize distribution in the unequal condition.](image-url)
**Attributions.** A 2 (Condition: Equal or Unequal) X 3 (Explanation Type: Internalist, Structural, Illogical) X 3 (Age Group: Younger (3-to-4-year-olds), Middle (5-to-6-year-olds), older (7-to-8-year-olds)) three-way mixed analysis of variance (ANOVA) revealed a significant interaction between condition and attribution type, $F(2, 291) = 6.38, p < .01$, such that between conditions, participants who viewed the unfair game endorsed the structural attribution significantly more than the participants who viewed the fair game, $t(147) = -3.37, p = .001$. Additionally, within conditions, participants in the unequal condition endorsed the structural attribution more than the internalist attribution, $t(147) = 3.68, p = .001$, and the structural and internalist attributions more than the illogical attribution, $ts > 0.33, ps < .001$. Participants in the equal condition also endorsed the structural and internalist attributions more than the illogical attribution, $ts > 0.25, ps < .001$, but did not differentially endorse the structural and internalist attributions, $t(147) = 0.02, p = .999$.

The interaction between condition and attribution type was additionally influenced by age, $F(4, 291) = 3.31, p < .05$. This three-way interaction revealed that the previously described pattern was true of all participants save 5-to-6-year-olds who viewed the fair game and 3-to-4-year-olds who viewed the unfair game. 5-to-6-year-olds in the equal condition failed to endorse the structural and internalist attributions more than the illogical attribution, $ts < 1.00, ps > .56$, and 3-to-4-year-olds in the unequal condition failed to endorse the structural over the internalist attribution, $t(73) = -0.04, p = .94$.

**Attributions v. chance.** Additional follow-up testing utilizing a Chi-square goodness-of-fit test was conducted to ascertain if participants endorsed or rejected the attributions in proportions different from chance on a binary measure. Results indicate that the negative question response language avoided “yes-biases” and possible ceiling effects, as all children,
regardless of condition, were able to reject the illogical attribution, $\chi^2 > 3.85$, $ps < .05$, save 5-to-6-year-olds in the equal condition who endorsed the attribution no different than chance levels, $\chi^2(1, N = 27) = 0.93$, $p = .33$. For the structural and internalist attributions, children of all ages who viewed the fair game endorsed both attributions at chance levels, $\chi^2 < 3.00$, $ps > .08$, save 5-to-6-year-olds who endorsed the internalist attribution significantly less than chance, $\chi^2(1, N = 27) 4.48$, $p = .03$. Children who viewed the unfair game endorsed the internalist attribution at chance regardless of age, $\chi^2 < 2.46$, $ps > .12$, but endorsed the structural attribution above chance, $\chi^2 > 7.54$, $ps < .01$, save for 3-to-4-year-olds who did not differ from chance, $\chi^2(1, N = 26) 1.96$, $p = .16$. See Figure 6.

Figure 6. Study 3 mean endorsement levels for three attributions for why the orange classroom won more prizes. For graphical depiction, responses were scored as either endorsed (“1”) or rejected (“-1”). The dashed line indicates chance level responding. Negative scores indicate rejection of the attribution while positive scores indicate endorsement. Error bars indicate 95% confidence intervals.
**Forced choice selection.** A Chi-square goodness-of-fit test was performed to ascertain if participants’ choice of which attribution was best differed from chance. Results reveal that participants, regardless of age or condition, significantly differed from chance in their selections, $\chi^2 > 8.62$, $ps < .01$, save for 5-to-6-year-olds in the equal condition, $\chi^2(1, N = 27) = 3.56$, $p = .17$. To determine which specific choices contributed to this significant difference, a closer examination of residuals was undertaken. Based on a predetermined significance level of $\alpha = 0.05$ and two degrees of freedom, the critical value for significance was determined to be ±1.96. Choices with residuals exceeding this critical value in either positive or negative direction were considered significantly different from chance levels.

For children in the unequal condition, regardless of age, only the structural attribution exhibited a substantial positive residual ($r_{\text{younger}} = 3.47$; $r_{\text{middle}} = 4.71$; $r_{\text{older}} = 6.06$), revealing that children who viewed the structurally unequal game viewed the structural attribution as the best idea for why the orange classroom was advantaged. The ability to pick out the structural attribution as superior strengthened with age as 65% of younger children, 77% of middle children, and 92% of older children selected the structural attribution. For younger and older children in the equal condition, only the internalist attribution exhibited a substantial positive residual ($r_{\text{younger}} = 2.22$; $r_{\text{older}} = 4.71$), demonstrating that most children who viewed the fair game thought the orange classroom had more prizes because the children in that classroom were taller. This preference for the internalist attribution also strengthened with age, as 54% of 3-to-4-year-olds and 77% of 7-to-8-year-olds selected the internalist attribution as best.

**Discussion**

Building upon previous research (Vasilyeva et al., 2018; Peretz-Lange & Muentener, 2021; Yang et al., 2021), Study 3 successfully replicated findings demonstrating that 5-to-6-year-
old children can make structural attributions for social inequalities, and that 3-to-4-year-olds show evidence of structural reasoning. A novel contribution of this work is that children were able to endorse structural attributions while simultaneously rejecting incorrect and illogical attributions. Importantly, this pattern only emerged when a structural constraint was present, suggesting that children can intentionally and selectively apply structural reasoning based on structural factors.

One potential limitation of the study pertains to the strength or validity of the provided internalist attribution. Based on past research, the results of Study 1, and our hypotheses, children in the equal condition, who observed a structurally fair game, were expected to default to internalist reasoning and endorse the internalist attribution. However, children's endorsement levels did not exceed chance, and 5-to-6-year-olds even significantly rejected the internalist attribution. It is possible that it was challenging for children to perceive how physical size could translate to higher success rates in the game. Nevertheless, children of all ages in the equal condition still selected the internalist attribution as the best explanation for the inequality, indicating that they still considered it an acceptable explanation. Future studies of a similar nature should consider incorporating more stable, trait-level internalist attributes such as attributions to work ethic, talent, or ability to provide a stronger internalist attribution for comparison.
General Discussion

Taken together the three studies show the development of structural reasoning in early childhood. While 3-to-4-year-old children have difficulty evaluating individual attributions and fail to demonstrate both causal and structural reasoning when asked to do so, they excel at selecting correct causal and structural attributions from a series of choices. This finding suggests that contrasting possible explanations for outcomes and inequalities can aid young children in their causal reasoning, including about social, structural inequalities. As they age, children demonstrate the ability to selectively make casual (Study 2) and structural attributions (Study 3), as well as attribute both equality (Study 1) and inequality (Study 1 & 3) to their structural causes by age five. Asking children to rate attributions using a negative question response language strengthened the validity of children’s structural attributions and more cleanly displayed their causal reasoning. This work is the first to show that children can selectively map internalist and structural reasoning to appropriate scenarios. Future developmental work should consider adopting negative question response language when asking affirmative/negative binary questions to circumvent possible “yes-biases” (Fritzley et al., 2012; Okanda & Itakura, 2007).

Given the evidence that children can indeed overcome internalist defaults and make structural attributions, two primary questions remain: Does an early foundation of structural attributions mitigate the development of bias? And if so, what can be done to enhance structural reasoning in childhood? Regarding the first question, previous research has provided positive, albeit mixed, evidence that engaging in structural attributions can decrease bias in children. For instance, when assessing children's social biases through preferred affiliation or "friendship" questions, it was observed that children were more likely to affiliate with disadvantaged novel group members when they were aware of structural constraints and able to attribute their
disadvantage to those constraints (Peretz-Lange & Muentener, 2021). However, children were not necessarily more inclined to want to affiliate with a gender non-conforming peer, even when attributing their gender-atypical behavior to structural factors (Yang et al., 2021). Nevertheless, Yang et al. (2021) discovered that children evaluated gender non-conforming peers less negatively when they attributed their behavior to structural constraints rather than internal preferences.

Understanding the impact of how structural attributions on racial biases is of upmost importance, given the pervasive nature of systemic racism (also known as institutional or structural racism) and its role in creating and perpetuating racial inequalities. Many forms of racial discrimination, oppression, and prejudice can be directly attributed to systemic and structural practices. If we apply the previously discussed logic, it suggests that racial biases may be particularly amenable to mitigation through the use of structural attributions, given the wide range of structural factors at play. However, recent research on children’s structural attributions for racism has yielded mixed results.

Supporting the notion that structural attributions for racism can reduce bias, Rizzo et al. (2022) demonstrated that children who endorsed structural attributions rather than internalist attributions for race-based economic inequalities at the age of four exhibited no subsequent increase in anti-Black/pro-white racial biases six months later, unlike their peers who favored intrinsic explanations. However, in work examining how white parents and children discuss and reason about anti-Black racism, Sullivan et al. (2023) found that white parents self-generated structural attributions less than 25% of the time, and white 7-to-13-year-old children barely self-generated any structural attributions at all. Moreover, in this study, children rejected two
structural attributions for anti-Black police brutality while endorsing an internalist explanation (i.e., attributing police behavior to "bad apples") as a "good idea" (Sullivan et al., 2023).

These findings align with a substantial body of research on conversations about racism, indicating that white individuals, including parents, often approach discussions on race with colorblind messages instead of adopting color-conscious perspectives. Colorblind messages disregard how race differentially impacts individuals' experiences and opportunities, and instead promote notions of unity and equality (e.g., "we are all one race, the human race"); Bigler et al., 2022; Vittrup, 2018). In contrast, color-conscious messages acknowledge how race shapes individuals' lived experiences and outcomes. These findings raise important questions regarding whether structural attributions are genuinely more cognitively challenging for children to make or if their limited experience in seeking out and reasoning about structural causes is a result of lacking role models and exposure to such attributions.

Another element to consider is the impacts of structural attributions on disadvantaged individuals. While bias mitigation in advantaged individuals is a necessary pursuit, attributing your own personal disadvantage to lofty and historic social structures may make the inequality feel insurmountable and immutable. According to system justification theory, some disadvantaged individuals turn away from acknowledging how social structures create inequality and towards views that justify and perpetuate existing social structures in order to manage uncertainty and threat (Jost et al., 2002; Jost & Hunyady, 2005). While children do have preference for high status group members (a trademark of system justification theory; Baron & Banaji, 2006; Newheiser et al., 2014; Qian et al., 2016; Shutts, 2015) and there is evidence that advantaged children are more likely to justify unfair hierarchies than disadvantaged children (Rizzo et al., 2023), to our knowledge there is no existing work which provides support for the
notion that disadvantaged children rationalize social inequities and hierarchies as just. In contrast, qualitative work has shown that while structural attributions can insight concerns about immutability in young children (Nguyen, 2022), older children and adolescence from disadvantaged groups often pair mention of societal oppression with messages of resistance (Rogers et al., 2022).

Future work should examine the cognitive abilities, motivation levels, and available information that enable children to make structural attributions for social inequalities. Several cognitive factors warrant consideration in the realm of structural reasoning, but chief among them are the abilities to: 1) understand protracted, downstream causality, 2) flexibly update attributions considering counterevidence, and 3) understand commonly held beliefs or norms may be incorrect. Additionally, children must be motivated to put in the cognitive effort required to make structural attributions. While the results of Study 1 suggest that children are not more motivated to make structural attributions for structurally constrained inequality over structural-based equality, factors such as personal experiences with inequality or intrinsic desires to be unbiased can potentially serve as motivators for children to make structural attributions. Furthermore, children must have access to developmentally appropriate information about social structures and how they contribute to inequality. Future research should explore how parents, teachers, and media can serve as models and provide scaffolding to enhance children’s understanding and development of structural reasoning.

In sum, the present work contributed to a growing body of literature on the development of structural reasoning in childhood by demonstrating children’s ability to selectively apply structural reasoning as early as three-years-old. In addition, this work created a novel measure for children’s casual attributions, which can circumvent “yes-biases” and served to disentangle
previous evidence of structural reasoning from said “yes-biases”. These findings advance our understand of children’s ability and motivation to make structural attributions and provide a foundation for future work investigating how this ability may serve to reduce the bias development associated with attributing inequalities between social groups to internalist factors.
References


https://doi.org/10.1177/17456916221093593


doi:10.1111/jora.12485


https://doi.org/10.1111/j.1467-9280.2005.01664.x


Haslanger, S. (2016). What is a (social) structural explanation? *Philosophical Studies, 173*(1), 113–130. [https://doi.org/10.1007/s11098-014-0434-5](https://doi.org/10.1007/s11098-014-0434-5)


Personally Inferior. Frontiers in Psychology, 12, 716662.

https://doi.org/10.3389/fpsyg.2021.716662


https://doi.org/10.1016/j.cognition.2016.07.004


Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. Proceedings of the National Academy of Sciences, 110(41).

https://doi.org/10.1073/pnas.1304326110


https://doi.org/10.1111/desc.12586


https://doi.org/10.1037/dev0000555

https://doi.org/10.1016/j.cognition.2020.104383

https://doi.org/10.1177/0192513X16676858


