Children's Reasoning About Structural Inequality

Matthew Tsui
June 2, 2023
PSYC 349
Introduction

Structural inequality occurs when social systems or institutions produce unequal outcomes between different groups of people (Rizzo & Killen, 2020). When faced with structural inequality, take for example poverty, there are two primary lines of reasoning one can take to explain this phenomenon. One way of reasoning is through utilizing a structural approach, which considers how social structures shape why certain people are more likely to be disadvantaged (Yang et al., 2021). Attributing poverty to societal factors such as high unemployment rates or discrimination are examples of using structural reasoning. In contrast, an internalist approach attributes intergroup differences to people’s individualistic traits (Peretz-Lange & Muentener, 2021). Attributing poverty to intrinsic factors such as laziness or lacking self control are examples of utilizing internalist reasoning. In adults, research largely finds observers tend to reason more internally than structurally when making attributions about poverty outcomes (Piff et al., 2020). However, less is known about the developmental trajectory of how we conceptualize inequality. Understanding this trajectory can give us insight into the origins of why adults tend to use internalist reasoning about inequalities and develop potential interventions to promote structural reasoning in childhood. As structural inequalities directly create and perpetuate harm to minority groups, the present research will examine how children begin to reason about and explain social disparities they encounter.

Internalist Defaults

Beyond reasoning about poverty, adults tend to focus on internal rather than structural factors when explaining various social inequalities (Bastian & Haslam, 2006; Rangel & Keller, 2011). This internalist default is largely based on the notion of psychological essentialism, which posits that an underlying “essence” of group members determines their observable similar
features (Gelman, 2004). Hence, holding essentialist beliefs can make existing social hierarchies (e.g., socioeconomic disparities) appear legitimate and natural (Hussak & Cimpian, 2015).

Notably, a preference for using internalist over structural reasoning about social groups is also found in childhood. As early as the preschool years, research predominantly shows children use internalist reasoning to make sense of the world around them (Rhodes & Mandalaywala, 2017). In an experimental study, Peretz-Lange and Muentener (2021) presented social status disparities between two groups playing a novel game that could be explained equally by internal or structural factors. Results found that children formed internal attributions and viewed the losing group as intrinsically inferior rather than being structurally disadvantaged when the structural inequality was not verbally highlighted to them. Crucially, children also apply these essentialist beliefs about intergroup differences to other personal characteristics like gender (Taylor et al., 2009), race (Mandalaywala et al., 2019), and language (Kinzler & Dautel, 2012).

The major concern of viewing social groups through an essentialist lens is the development of prejudice (see Rhodes & Mandalaywala, 2017 for a review). For example, Mandalaywala et al. (2017) found that greater essentialist beliefs about race were linked to greater prejudice towards Black people among both White and Black adult participants. Essential beliefs have also been directly linked to prejudice in children. Diesendruck and Menahem (2015) found that after showing 6-year-old Israeli children a story that emphasized ethnic essentialism (compared to a neutral control story), children displayed more ingroup preference and racial prejudice. Specifically, children drew an in-group Jewish character with more positive affect (assessed by their facial expression) compared to an out-group Arab character and drew more physical distance between the two characters. These findings suggest essentialism promoted prejudice in children as they had an affective bias towards their ingroup and saw out-group
members as more socially distant. Therefore, one avenue towards mitigating prejudice could be moving children away from essentialist thinking by promoting alternative forms of reasoning about how structural inequalities impact social groups. One alternative form of reasoning is structural reasoning.

**The Emergence of Structural Reasoning**

Structural reasoning involves recognizing how larger structural constraints, such as group treatment or opportunities, place certain social groups at a greater risk of unequal outcomes. Returning to the poverty example, Piff et al. (2020) found American adult participants who used structural reasoning believed their country’s income differences should be more equal (indicated higher support for economic equality) and gave more donations to an organization aimed to reduce economic inequality than participants who used internalist reasoning to explain poverty. Given structural reasoning is an important alternative to explaining inequalities, which may help in reducing prejudice about certain social groups, it is imperative to study the emergence and usage of it in childhood.

A wealth of related findings provides mixed evidence as to whether children are able to utilize structural reasoning. Some evidence points to the possibility that young children simply cannot reason structurally about inequality because they have yet to develop the cognitive ability to recognize the relationship between an individual (or social group) and their social context. For example, Richland et al. (2006) found young children consistently made errors on an analogical reasoning task when having to process multiple relations at once, suggesting they lack the relevant capacity needed to engage in more complex relational reasoning. Additionally, Cimpian and Steinberg (2014) found that even 4- to 5-year-old children prefer inherent over extrinsic explanations when reasoning about social patterns. These findings, alongside prior work finding
children hold essentialist views towards some social categories (e.g., social status; Rhodes & Mandalaywala, 2017), suggest structural reasoning may not emerge until much later in development.

On the other hand, past studies on resource allocation behaviour have demonstrated evidence for the alternative possibility that young children *can* in fact reason structurally when encountering social inequalities. For example, Rizzo et al. (2018) assessed children’s perceptions towards individually-based inequalities (i.e. a peer worked harder) or structurally-based inequalities (i.e. the one distributing the prizes had a gender bias) that caused a peer to receive more prizes than the other. Researchers found that children judged the structural inequality as more unfair and distributed more resources to the peer who was disadvantaged because of their gender in comparison to the individually-based inequality. While this finding suggests that children considered the structural source of the inequality when distributing resources, this study did not provide an explicit test of their reasoning. Additionally, this study operationalized structural inequality through the use of an identity group (i.e., gender) – an approach common amongst resource allocation paradigms seeking to examine how group characteristics impact children’s distribution decisions (Elenbaas et al., 2020). However, the use of identity as a representation of inequality makes disentangling structural and internalist reasoning difficult – for while identities such as race or gender can be constrained by structural oppression, they are also internal characteristics that individuals identify with.

Peretz-Lange & Muentener (2021) sought to directly test for structural reasoning about social inequalities by introducing 5- to 6-year-olds to two novel groups who had different intrinsic abilities while playing a novel game, resulting in a social status disparity. Through verbally highlighting the structural inequality to the children as to why the advantaged group
won the game (scoring points required a specific physical ability only one group possessed), children made more structural than internalist attributions. Hence, these findings seem to provide more concrete evidence children can use structural reasoning to disrupt their internalist defaults. However, given that the specific physical ability differences between the two groups are also internal characteristics, children again may have entangled structural and internalist reasoning. It thus remains unclear whether children can reason structurally about structural inequality in the absence of observable differences in group identities.

Vasilyeva et al. (2018) provided additional evidence that young children are capable of engaging in structural reasoning. In their experimental study, 3- to 6-year-olds were introduced to a vignette involving gender segregated classrooms. Participants were told each child tossed a pebble towards a yellow and green bucket placed side by side, and would play a game associated with whichever coloured bucket the pebble landed in. However, the crucial manipulation used was the bucket sizes – whereas the buckets were the same size in the non-structural condition, the girls’ yellow bucket was much larger than the green bucket (inverse for the boys’ classroom) in the structural condition. Structural reasoning was then assessed by first asking participants an open-ended question about why girls predominantly played the yellow bucket’s game, and then by having them evaluate a structural (“easier to get pebble into yellow bucket”), internal (“girls prefer the yellow bucket game”), and incidental (“girls got sprinkled with water”) explanation. Specifically, participants evaluated each explanation as either a “good” or “bad” explanation. Across all age groups, children were more likely to generate structural over internalist explanations for the open-ended question task in the structural condition. Additionally, 5- to 6-year-olds endorsed the structural explanation as a good idea more so than children in the non-
structural condition. This suggests children can engage in structural reasoning, even overriding intuitions to endorse an internal explanation for social disparities to do so.

It is important to note that Vasilyeva et al.’s results also revealed children highly endorsed all three explanations, including the incidental explanation, with significant differences only showing in the 5- to 6-year-olds structural ratings in the structural condition. On one hand, it is possible that children reasoned that all of these explanations are good explanations for why girls tend to play the yellow bucket’s game. However, given that children were also endorsing the incidental explanation (“girls got sprinkled with water”), it is also possible that measuring reasoning through verbal evaluations of ideas as either “good” or “bad” may have led them to respond in this way. This may be due to young children’s tendency to display a yes bias towards yes-no questions (Fritzley et al., 2012) causing children to endorse the “good” choice over the “bad” for all explanations.

The Present Research

To date, most work on structural reasoning about inequality has either entangled structural and internalist reasoning by using identity-based structural inequality, or had results possibly confounded by “yes biases” in participant responding. Therefore, further research investigating if utilizing structural reasoning can supersede the internalist default is needed to (a) shed light on children’s ability to reason structurally about inequality and (b) inform future work on how structural reasoning might be utilized to reduce social biases. In the current study, we first conducted a methodological investigation to identify a response paradigm which captures children’s causal reasoning, while overcoming “yes biases.” We then utilized this method in order to build upon Vasilyeva et al.’s (2018) work and further explore in what contexts children can favour structural over internalist reasoning in a structurally unfair situation.
Study 1 sought to test how two different framing response paradigms impact children’s responses when evaluating explanations. Specifically, a causal, non-causal, and illogical explanation were presented for the outcome of an event. As preschool children have the ability to make causal inferences (Schulz & Gopnik, 2004), the logic is that children should endorse the causal over the non-causal and illogical explanation. However, Vasilyeva et al.’s (2018) results demonstrated that young children were endorsing all proposed explanations, including illogical explanations, above chance. The first framing response paradigm directly replicated Vasilyeva et al. (2018)’s response language by asking children if each explanation is a good or bad idea. The second framing response paradigm utilized a negative question structure, asking if each explanation is a silly idea. We chose this negative question framing because children will logically have to say “no” to endorse the explanation, hopefully circumventing the yes bias. A forced choice measure was used to further probe which out of the explanations presented they endorse the strongest. We hypothesized that 3- to 6-year-old children using the replication framing method would endorse all the explanations above chance, whereas children using the negative question framing method would reliably endorse the causal explanation and reject the non-causal and illogical explanations.

Study 2 sought to investigate children’s ability to utilize structural reasoning to explain a purely contextual, structurally inequitable outcome between two groups with the same intrinsic characteristics. Two identical groups were compared in an attempt to isolate the cognitive development of structural reasoning by removing identity as a representation of inequality. Children viewed either a structurally unfair game (structural condition), in which a structural constraint placed one group at a disadvantage, or a fair game (control condition), in which both groups had an equal chance of winning a desired prize. In the absence of a structural constraint,
the logic is that children should default to internalist reasoning. However, if children are able to reason structurally about inequality, structural reasoning should supersede the internalist default when observing one group is structurally disadvantaged compared to the other. We then evaluated children’s endorsement of a structural, internalist, and illogical explanation for the disparity utilizing the negative question framing method with the additional forced choice question. Therefore, given the past positive evidence for structural reasoning in childhood, alongside a novel response paradigm, we hypothesized that 3- to 8-year-old children in the structural condition will endorse the structural explanation over the internalist explanation for the inequality when utilizing a negative question or forced choice response structure. However, we predicted that children in the control condition will endorse the internalist explanation over the structural explanation for the inequality due to their default to internalist reasoning in the absence of a structural inequality. Finally, we predict this response pattern will emerge at age 3 and strengthen with age.

Study 1

Method

Participants

Sixty 3- to 4-year-olds (mean age = 3.99, range = 3.00-4.94; 31 boys) and seventy-one 5- to 6-year-olds (mean age = 5.82, range = 5.00 - 6.94; 30 boys and 39 girls, 2 unreported) participated in this study. Given one of the conditions directly replicate Vasilyeva et al.’s (2018) response paradigm, the target sample size was based on their 20 participants per condition per age group. Participants’ ethnicity was approximately representative of the sampling population: 34% White/Caucasian, 23% East Asian, 17% Multiracial, 8% South Asian, 7% Southeast Asian, 2% Indigenous, 2% Black/African Canadian, 1% Latinx/Hispanic, 1% self-reported Somali
participant, and 5% who preferred not to answer. Fifteen additional participants were tested but excluded from analyses due to either parent or sibling interference (n = 5), incorrect age (n = 3), English comprehension issues (n = 3), or experimenter/technological errors (n = 3).

All participants were recruited by trained research assistants who approached parents visiting a local science museum located in a Canadian city. While socioeconomic information was not collected, demographic information of the surrounding area suggests our sample likely came from a middle-to-upper class socioeconomic background. Participants were randomly assigned to either the replication condition or the negative question framing condition. Children verbally provided assent and all written parental consent forms were obtained prior to beginning the procedures. Following the completion of the study, all participants received a sticker as a compensation for their time. The research methods and experimental protocols for both studies conducted in this paper were approved by the University of British Columbia’s Behavioural Research Ethics Board.

**Materials, Design, Procedure**

Trained research assistants individually tested participants in a quiet space nearby the museum exhibit from which they were recruited. This arrangement often allowed parents to sit nearby or with the participant, but parents were politely asked to refrain from interfering until the study was complete.

All participants were read a short illustrated vignette (presented on an 9.7" iPad, created on Microsoft PowerPoint) that involved two girl characters matched in race, gender and physical size and differentiated by shirt colour, hair style, and name (Camilla and Ava). Participants were told these two characters raced against each other and that Ava won. Then, participants evaluated three different explanations offered by three different animal puppets (differentiated by kind and
colour) as to why Ava won. The order of the explanations were counterbalanced across participants. Crucially, children responded using either the replication framing condition (n = 55) or the negative question framing condition (n = 74). A subset of the participants in the negative question framing condition additionally responded to a forced choice measure (n = 52).

**Replication framing condition.** This condition replicated the good or bad idea framing language used in Vasilyeva et al.’s (2018) study. In this condition, participants evaluated each explanation by choosing one of two thumbs representing a “good idea” (up) or a “bad idea” (down) – a methodology similarly following Vasilyeva et al.’s (2018) study.

**Negative question framing condition.** This condition utilized a negative question in hopes of overcoming potential “yes biases” children hold (Fritzley et al., 2012). Participants evaluated if each explanation was a “silly idea” by choosing one of two thumbs representing “yes” (up) or “no” (down).

**Forced choice measure.** Following the individual endorsements, a subset of participants (n = 52) in the negative question condition were verbally reminded and saw visual symbology for each of the explanations presented. Then, participants were asked to pick the best explanation out of the three for why Ava won the race.

**Results**

To analyze the data collected, we used a chi-square goodness of fit test, a statistical test that determines if the observed data are significantly different from chance. This test was appropriate because our main goal is to determine the extent to which participants' responses in each condition are significantly different from chance. A binary coding scheme was used where responses were coded as either the presented explanation was endorsed (1) or not (0). Hence, 0 coded for “bad idea” and 1 coded for “good idea” responses in the replication framing condition,
and 0 coded for “yes, the idea is silly” and 1 coded for “no, the idea is not silly” responses in the negative question framing condition. In the subsequent forced choice question in the latter condition, responses were coded (1) for the selected explanation and (0) for the other two unselected explanations. An analysis of variance (ANOVA) was also used to determine if the mean endorsement response patterns for each condition differed significantly from each other. All the analyses in both studies were performed using R statistical software. ANOVA and post hoc analyses were conducted with the afex and emmeans R packages.

**Explanations**

As hypothesized, a 2 x 2 x 3 mixed ANOVA on children’s evaluations as a function of age group (3- to 4-years old, 5- to 6-years old), framing condition (replication, negative question), and endorsement type (causal, non-causal, illogical) revealed a main effect of framing condition, $F(1, 125) = 23.64, p < .001$. Post hoc analysis revealed children who heard the silly framing language, compared to the good or bad idea framing language, were significantly less likely to endorse the non-causal, $t(125) = 4.71, p < .001$, and illogical explanation, $t(125) = 2.96, p = .004$. Interestingly, regardless of condition, there was an interaction between age and endorsement response type, $F(1, 88) = 4.60, p = .013$, such that as children grew older, they were more likely to endorse the causal explanation, $t(125) = -2.23, p = .024$.

To investigate if endorsement responses for the explanations presented were influenced by “yes biases”, a chi-square goodness of fit test was utilized to examine these patterns for both conditions. Endorsement response patterns to the good or bad idea framing language was at ($ps > 0.423$) or above chance ($X^2(1, N = 56) = 12.07, p < .001$) for all explanation types (see Figure 1). Specifically, older children (5- to 6-year-olds) endorsed the causal explanation significantly above chance, $X^2(1, N = 29) = 7.76, p = .005$, but did not endorse the non-causal and illogical
explanations significantly different from chance ($ps > .35$). Similarly, younger children (3- to 4-year-olds) endorsed both the causal, $X^2(1, N = 26) = 4.48, p = .03$, and non-causal explanation, $X^2(1, N = 26) = 4.48, p = .03$, significantly above chance, and did not endorse the illogical explanation significantly different from chance ($p = .847$). These patterns suggest children demonstrated a “yes bias” by endorsing every explanation as a good idea when being asked if the explanation was a good or bad, mirroring Vasilyeva et al.’s (2018) results.

Contrastingly, endorsement response patterns to the silly idea framing language were below chance ($X^2 s > 17.41, ps < .001$) for the non-causal and illogical explanations (see Figure 1). Specifically, older children endorsed the causal explanation significantly above chance, $X^2(1, N = 42) = 11.54, p = < .001$, and endorsed the non-causal and illogical explanation significantly below chance ($ps < .002$). These findings indicate asking older children if the idea was silly helped them reliably reject the non-causal and illogical explanations compared to being asked if the idea was good or bad. Similarly, younger children rejected the non-causal, $X^2(1, N = 32) =12.50, p = < .001$, and illogical explanation, $X^2(1, N = 26) =4.48, p = .004$, rating both significantly below chance. Although younger children interestingly endorsed the causal explanation at chance ($p = .48$), this may not necessarily reflect an inability to engage in causal reasoning. Rather, given they reliably rejected the alternative non-causal and illogical explanations, this demonstrates the causal explanation was the only idea not significantly endorsed as silly. This response pattern also suggests that using a negative question helped children on average overcome their “yes bias”.
Figure 1. Distribution of explanation endorsements as a function of the framing condition and age group. Error bars represent standard error.

**Forced Choice**

When children in the negative question framing condition (n = 52) were then asked to select the best explanation, a 2 x 3 mixed ANOVA as a function of age group (3- to 4-year-olds, 5- to 6-year-olds) and endorsement type (causal, non-causal, illogical) revealed a main effect of endorsement type, F(1, 78) = 17.73, p < .001. Specifically, the causal explanation (68%) was chosen significantly more than the non-causal explanation (15%) for both the younger, t(50) = 3.16, p = .008, and older children, t(50) = 3.80, p = .001. The causal explanation was also chosen significantly more than the illogical explanation (17%) for both the younger t(50) = 3.16, p = .008, and older children, t(50) = 3.80, p = .001. In sum, as this measure forced children to overcome their “yes-biases” by having them endorse a single explanation, all children reliably demonstrated causal reasoning and rejected incorrect explanations. Notably, there were also no
significant differences between the children’s age and the best idea chosen ($p = .60$), indicating similar comprehension of the forced question measure across all ages.

**Discussion**

To explore children’s ability to display causal reasoning and overcome their “yes biases” when endorsing explanations, Study 1 compared Vasilyeva et al. (2018)’s good or bad idea framing paradigm to a novel negative question framing paradigm. In line with our hypothesis, we found 3- to 8-year-old children in the replication condition endorsed all three explanations at or above chance, but reliably rejected poor or illogical answers in the negative question framing condition. Additionally, both younger and older children selected the causal explanation as the best idea when forced to choose between the three options. Together, these findings provide evidence that children overcame their “yes biases” by asking them to evaluate explanations using a negative question framing paradigm.

Given the past positive evidence that preschool children are able to make causal inferences (Schultz & Gopnik, 2004), children should logically be able to endorse the causal explanation while rejecting the alternative poorer explanations regardless of condition. However, our finding that children tended to endorse everything using the good or bad idea framing language may suggest “yes biases” masked the true results of previous work exploring children’s structural reasoning using this paradigm. Hence, the current study may be the first to show children are capable of rejecting explanations they truly believe to be poor by helping them reason with a negative question.

We also note some developmental differences between children’s endorsement of the causal explanation though when using this negative framing paradigm, as only the older children endorsed the causal explanation significantly above chance. While this finding may seem
puzzling, it may not necessarily reflect younger children’s inability to reason causally. Rather, one potential explanation is that preschool children’s less developed inhibitory control, compared to their older counterparts, render them more susceptible to “yes biases” (Moriguchi et al., 2008). As inhibitory control is important for helping children avoid saying the first thing that comes to their mind (Scullin & Bonner, 2006), perhaps younger children failed to inhibit their tendency to respond yes despite understanding the causal explanation was not silly. However, our finding that younger children still reliably rejected the other two incorrect explanations, alongside selecting the causal explanation as the best idea, suggests this paradigm was still able to circumvent some of their biases to respond yes. Thus, our novel design may be more effective in capturing children’s reasoning about explanations. We use this negative question framing paradigm in Study 2 to help capture children’s ability to engage in structural reasoning.

Study 2

Method

Participants

Twenty-four 3- to 5-year-olds (mean age = 3.95, range = 3.00-5.59; 13 boys) and twenty-nine 6- to 8-year-olds (mean age = 7.95, range = 7.00 - 8.95; 12 boys) participated in this study. The target sample size follows the sample size rationale as in Study 1. Participants’ ethnicity was approximately representative of the sampling population: 60% White/Caucasian, 20% East Asian, 8% South Asian, 4% Black/African Canadian, 4% Latinx/Hispanic, 2% Southeast Asian, 2% Indigenous. Five additional participants were tested but excluded from analyses due to parental interference (n = 2), incorrect age (n = 1), experimenter/technological error (n = 1) or failing to understand the logistics of the study (determined by failing the two comprehension checks; n = 1).
Participants were recruited in the same manner as in Study 1. While socioeconomic information was not collected, demographic information of the surrounding area suggests our sample likely came from a middle-to-upper class socioeconomic background. Participants in each age group were randomly assigned to either the structural condition (n = 26) or the control condition (n = 27) in a counterbalanced order. Children verbally provided assent and all written parental consent forms were obtained prior to beginning the procedures. Regardless of if the study was completed, all participants received a sticker as a compensation for their time.

**Materials, Design, Procedure**

Trained research assistants individually tested participants in a closed, soundproof room away from the museum exhibits. In the testing room, participants sat in a short chair (designed for child use) that was in front of a low table, with a computer monitor mounted to a wall approximately 12” away.

All participants read an illustrated vignette (presented on a 24” screen, created on Microsoft PowerPoint) in which two groups of students played a novel game to win prizes for their respective classrooms. Notably, the characteristics of both group members were identical in terms of race (2 Latinx/Hispanic, 2 White/Caucasian, 1 Asian, and 1 Black character per group) and gender (3 boys and 3 girls per group). The only observable differences between the group members were the colour of their shirts, which represented their classroom (either orange or green) and hair styles. In this novel game, participants were told students from each group took turns throwing their balls towards their classroom’s bucket. The student would win a highly desirable prize (a basket of toys and treats) if the ball landed in the bucket, but would receive a less desirable prize (a toothbrush) if the ball missed (Figure 2a). After all the students had thrown their balls, the prize distribution for each class was shown, with the orange classroom winning
more desired prizes (4 baskets and 2 toothbrushes) than the green classroom (2 baskets and 4 toothbrushes). Participants were asked to match each prize with the toss outcome, to check for comprehension.

Between subjects, the critical manipulation was the size of the buckets. In the structural condition, the orange classroom’s bucket was larger than the green classroom’s bucket (Figure 2b). In the control condition, both buckets were the same size (Figure 2c). Following the vignette, all participants were introduced to three different animal puppets (differentiated by kind and colour) who offered three different explanations about why students in the orange classroom won more baskets of toys and treats than students in the green classroom. The puppets gave either a structural explanation (“because it is easier to get the ball into the orange bucket”), an internalist explanation (“because students in the orange classroom are taller”) or an illogical explanation (“because frogs are orange”). All participants evaluated each explanation using the negative question framing (“do you think that’s a silly idea?”) and responded by choosing one of two thumbs representing “yes” (up) or “no” (down). The order of the animal puppets and three explanations presented were counterbalanced across participants. After all the explanations were evaluated, participants were verbally reminded and saw visual symbology for each of the three explanations presented. Then, a forced choice measure asked participants which of the three explanations was the best idea for why students in the orange classroom won more baskets of toys and treats.
Results

We conducted analyses using the same binary coding scheme as Study 1, where responses were coded according to whether the presented explanation was endorsed (1) or not (0). An ANOVA was used to investigate if there were significant differences in mean endorsement responses between the conditions and age groups. Then, a chi-square goodness of fit test was used to determine if response patterns within each condition differed significantly from chance.

Explanations

As hypothesized, endorsement response patterns were affected by whether or not a structural inequality was presented. A 2 x 2 x 3 mixed ANOVA on children’s evaluations as a function of age group (3- to 5-year-olds, 6- to 8-year-olds), condition (structural, control) and explanation type (structural, internalist, illogical), revealed a significant interaction between

Figure 2. Illustration of the ball throwing procedure to determine what prize each student won (a) and how the bucket game is set up for each classroom in the structural condition (b) and the control condition (c).
condition and explanation type, $F(1, 96) = 8.26, \ p = < .001$. Post hoc analysis revealed children who viewed the unequal bucket sizes in the unfair game endorsed the structural explanation significantly more than children who viewed the equal bucket sizes in the fair game, $t(49) = -3.57, p < .001$. There was also a significant condition by type by age group interaction, $F(1,96) = 3.73, p = .03$ (see Figure 3). Notably, when viewing the fair game, there were no significant differences in endorsement levels between the explanation types and age groups ($ps > .06$). However, when viewing the unfair game, older children (6- to 8-year-olds) only endorsed the structural explanation, $t(27) = -4.41, p < .001$, while rejecting the internalist and illogical explanations ($ps > .24$), whereas younger children’s (3- to 5-year-olds) endorsement of the structural and internalist explanation did not differ from each other ($p = 1.00$). This suggests a developmental difference where older, but not younger, children reliably used structural reasoning to overcome their internalist defaults when observing discrepant outcomes between the groups in the unfair game.

Given children responded using a binary forced choice measure to endorse explanations, it is also important to confirm children’s mean endorsements significantly differed from chance. Regardless of age, children who viewed the fair game endorsed all explanations at chance ($ps > .07$). However, when viewing the unfair game, younger children endorsed both the internalist and structural explanation significantly above chance ($X^2 s > 9.00, ps = .003$), whereas older children endorsed the structural explanation significantly above chance, $X^2(1, N = 17) = 13.24, p = < .001$, and the internalist explanation significantly below chance, $X^2(1, N = 17) = 9.94, p = .002$. This provides additional evidence that the ability to use structural reasoning to overcome internalist defaults develops with age. Notably, the illogical explanation was endorsed
significantly below chance for children who viewed the structurally unfair game, \(X^2s > 9.00 = 18.62, ps = < .002\), demonstrating that response patterns were not influenced by “yes biases.”

**Figure 3.** Distribution of explanation endorsements as a function of age group and framing condition. Error bars represent standard error.

**Forced Choice**

The forced choice measure revealed strikingly different patterns of endorsement between each condition, \(F(1, 66) = 26.40, p = < .001\). Interestingly, there were no significant differences between the children’s age and their forced endorsement choice \((p = .73)\), indicating similar comprehension of the forced question measure across all ages. Post hoc analysis revealed children who viewed the unfair bucket game selected the structural explanation (88%) significantly more than those who viewed the fair game (11%), \(t(49) = -6.05, p = < .001\), suggesting children were able to correctly attribute the discrepant outcome to structural factors. However, children who viewed the fair game selected the internalist explanation (67%)
significantly more than those who did not (8%), $t(49) = 5.27, p < .001$, suggesting a default to internalist reasoning in the absence of a structural inequality.

**Discussion**

Study 2 showed a structurally unfair game to explore if children could engage in structural over internalist reasoning to explain discrepant outcomes between two groups of people. As predicted, we found that children in the structural condition, compared to the control, endorsed the structural explanation significantly above chance. Crucially, regardless of age, children in the structural condition were also the most likely to select the structural explanation, whereas those in the control condition were the most likely to select the internalist explanation. Together, these findings provide evidence that children have the ability to reason structurally and overcome their tendencies to reason internally when presented with inequality.

The findings that structural reasoning may emerge in children as young as 3-years-old extends the work of Vasilyeva et al. (2018) who found children by age 5 were capable of structural reasoning about inequalities. There were, however, developmental differences in endorsement patterns of the internalist explanation in the structural condition. Whereas older children reliably rejected the internalist explanation for the inequality, which was in line with our prediction, younger children endorsed the internalist explanation above chance alongside the structural explanation. These findings highlight that both older and younger children are capable of engaging in structural reasoning, but suggest older children have a stronger ability to overcome their defaults to internalist reasoning. Additionally, regardless of age, children reliably rejected the illogical explanation, demonstrating an aversion from “yes biases” as they did not endorse every explanation.
Notably, in the control condition, we expected children would endorse the internalist over the structural explanation in the absence of a structural inequality. However, there were no differences in endorsements responses between all three explanations. Given research predominantly shows children default to internist reasoning to explain the social world (Rhodes & Mandalaywala, 2017), alongside our finding that they selected the internalist explanation the most in the forced choice measure, these endorsement result patterns may not imply children could not reason internally. Children simply may have had a harder time reasoning about the fair game in contrast to the unfair game, and therefore our sample size may have not been large enough to detect meaningful differences in their endorsement levels between each explanation.

**General Discussion**

Across two studies, we developed a novel response paradigm that captured children’s causal reasoning while overcoming yes-biases (Study 1), and utilized that measure to explore children’s ability to use structural over internalist reasoning when viewing a structurally unfair situation (Study 2). When being shown a structural constraint that created discrepant outcomes between two identical groups, all children made structural attributions to explain the inequality. However, a key developmental difference was that 6- to 8-year-olds successfully overcame their internalist defaults by rejecting the internalist explanation, whereas 3- to 5-year-olds endorsed both the structural and internalist explanation. In sum, this provides evidence that children as young as 3 have the ability to utilize structural reasoning, although this ability strengthens over development.

Novel to the current work, using a negative question framing paradigm forced children to overcome their tendency to endorse every explanation. Hence, this paradigm may be effective for future research seeking to assess children’s reasoning via endorsement of presented
explanations. Specifically, yes-no questions have been posed as the most problematic for preschoolers (Fritzley & Kang, 2003), with Okanda and Itakura (2010) finding they tend to inappropriately respond “yes” despite knowing the correct answer. More broadly, this measure may also have important implications for the field developmental psychology which often utilizes these types of questions in experimental tasks to understand and draw conclusions about children’s cognitive and social development.

Our work also dovetails previous work (Vasilyeva et al., 2018; Yang et al., 2021) that demonstrated older children have a greater ability to utilize structural over internalist reasoning compared to younger children. Whereas younger children in the structural condition endorsed the structural and internalist explanation equally, older children only endorsed the structural explanation for the inequality. While this may indeed be evidence of a developmental shift in which older children are stronger at overcoming their internalist defaults, an alternative explanation may be that younger children actually viewed the internalist explanation as plausible when reasoning about the unequal outcomes. As young children (4- to 5-year-olds) have a stronger ability to detect causal relations in their environments (Seiver et al., 2013; Lucas et al., 2014), perhaps their reduced cognitive abilities may have aided their learning of causal relations that their older counterparts miss (Gualtieri & Finn, 2022). Hence, despite there being no evidence to support being tall was an advantage in the game, perhaps younger children were able to reason being taller also improved accuracy when throwing the ball because of a better view of the bucket target. This is further supported by Gopnik et al. (2015)’s findings that younger children are (a) more likely to infer hypotheses based on limited evidence and (b) inherently have more plastic and exploratory minds (that comes at the cost of being less efficient) compared to older children. Thus, given the core of Study 2’s paradigm required causal cognition to
endorse different explanations, younger children simply may have been more open when considering the presented explanations instead of being overtaken by internalist defaults. Moreover, both younger and older children selected the structural explanation as the best idea when viewing the structural inequality, further suggesting these differences in abilities may not be as large as it seems.

Additionally, it is important to note that children across both studies were likely from families of middle-to-upper SES, as we recruited from a science museum that required a high admission fee. Thus, our sample may not be generalizable to children from families of lower SES, who are more likely to directly experience a number of socioeconomic inequalities (e.g. Weyers et al., 2009; Senn et al., 2015). Especially in the context of Study 2 where children viewed a group that was structurally disadvantaged, it can be questioned if children’s experiences with inequality impacts their ability to engage in structural reasoning. One hypothesis is that it will be harder for lower SES children to engage in structural reasoning because low-SES has been associated with impaired cognitive development (McLoyd, 1998) and executive functioning skills (Moriguchi & Shinohara, 2019; Noble et al., 2007) in childhood. In other words, perhaps tasks such as being forced to choose the best explanation for the inequality when being asked to consider multiple explanations at once, or remembering the logistics of the game throughout the study, may pose more difficulties. An alternate hypothesis is that lower SES children’s direct experiences with inequality will help them better recognize the inequality present in the game, thus enhancing their ability to engage in structural reasoning. This suggestion can be supported by findings that 3- to 8-year-old children who were assigned to receive fewer resources than another group were more likely to view inequalities as unfair and wrong (Rizzo & Killen, 2020). Thus, while experiences of inequality in experimental settings
may not generalize to real life experiences of inequality, future research should explore if lower SES children have a stronger ability to overcome their internalist defaults.

To conclude, across two studies we used a novel response paradigm to show children as young as 3 are capable of making structural attributions for observed social disparities. We provide initial evidence that structural reasoning can be used to supersede children’s internalist defaults, with this ability strengthening with age. These findings may help shed light on future interventions attempting to promote structural reasoning to combat the development of social biases and prejudice about social groups.
References


Seiver, E., Gopnik, A., & Goodman, N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? causal inference and the development of social attribution. *Child Development, 84*(2), 443–454. https://doi.org/10.1111/j.1467-8624.2012.01865.x


