"I only want one!"
Children's evaluations of inequality by choice

Marissa Britt Drell
Charlottesville, VA

Master of Arts in Psychology, University of Virginia, 2015
Master of Arts in Psychology, Brandeis University, 2010
Bachelor of Arts, Vassar College, 2008

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Psychology

University of Virginia
March, 2018

Vikram K. Jaswal (Chair) _____

Angeline S. Lillard _____

Amrisha Vaish _____

Jamie Jirout _____

Abstract

Young children believe individuals should receive the same number of resources (Shaw & Olson, 2012). However, research has found that, by age 6, children have learned that it is acceptable to deviate from equality in the case of effort and need (Rizzo & Killen, 2016; Rizzo, Elenbaas, Cooley, & Killen, 2016). In the studies here, I asked whether 6- and 8-year-olds also consider inequality that occurs when one person *chooses* to take less than his or her fair share of the available resources as another acceptable reason to deviate from equality. Study 1 found evidence for a developmental shift in response to inequality by choice such that 6-year-olds object, 8-year-olds are ambivalent (some think it is okay while others do not), and adults think that it is permissible for one person to take less than half the resources and for a second person to take the remainder. Study 2 found that 6-year-olds' objection to inequality by choice in Study 1 is not due to a failure to understand that two people can desire different amounts but rather is driven by a normative response that everyone *should* take the same amount. Study 3 and 4 found that 6- and 8-year-olds' objection to inequality by choice is primarily driven by the violation of the equal outcome norm as opposed to other potential norm violations. Study 5 found that 6-year-olds are unable to override the equal outcome norm even when the amount of cognitive effort needed to understand inequality by choice is reduced. This research adds to a growing body of literature on the developmental timeline of the factors children integrate into their resource distribution judgments.

Keywords: fairness, inequity aversion, social cognition, desire, free will

Table of Contents

Introduction

Most children and adults agree that resources should be divided equally between equally deserving individuals (e.g., Fiske, 1992; Shaw & Knobe, 2013; Shaw & Olson, 2012; Smith, Blake, & Harris, 2013). But as adults, we know that there are situations in which it is acceptable for one person to get more of a resource than someone else (Arneson, 1993; Cohen, 1989; Dworkin, 1981; Temkin, 2011). For example, if a faculty member puts some leftover pizza in her lab for her two graduate students to split, one might expect the two students to split the pizza in half. However, if one of the students announces she only wants one of the four slices, most adults would think it perfectly acceptable for the other student to take the remaining three. After all, the first student made clear her desire, and adults recognize that people are generally satisfied when their desires are fulfilled. The first student may have wanted only one slice because she was not hungry or because she was feeling generous toward the other student. But in either case, intuitively and from an adult perspective, it seems like the resulting unequal distribution of pizza slices would be perfectly acceptable. In contrast, if the first graduate student had announced a desire for and took three pieces, leaving just one for the second student, this seems objectionable: The first student seems greedy. In this dissertation, I investigate how 6- and 8-year-old children evaluate inequality that occurs when one person *chooses* to take less than his or her fair share of available resources versus when one person ends up with less than his or her fair share.

Investigating how children evaluate inequality that occurs by choice has important theoretical implications for how we think about children's understanding of fairness. As I will describe below, a great deal of research on how children distribute and evaluate others' distribution of resources has shown that, from a very early age, children expect resources to be

distributed equally. However, there are some circumstances where a society or culture may approve of the unequal distribution of resources—as in the choice example just described, but also in cases of merit or need.

I expect that by six years of age (the earliest tested in the studies to be described here), children have experienced or observed examples of inequality that occurs by choice. For example, a parent may choose to eat less than half of a banana, leaving the rest for the child. Or children may see one child at the library checking out one book while another checks out a stack of them. Studying inequality that occurs by choice offers the opportunity to investigate how children make sense of one familiar type of inequality that is acceptable by adult standards in Western cultures (Schäfer, Haun, & Tomasello, 2015), and adds to a growing body of literature on the developmental timeline of the factors children integrate into their resource distribution judgments. Studying children's understanding of inequality that occurs by choice also sheds light on when in development children either learn certain types of inequality are acceptable or are able to think abstractly about the reason social norms exist and generate the cases under which it is acceptable to violate these norms. The studies reported in this dissertation were not designed to test which one of these mechanisms explain how children learn inequality by choice is permissible, but rather serve as a necessary first step to document when children recognize this type of inequality is permissible.

I begin by describing children's expectation that resources will be distributed equally, then turn to some previous research on two exceptions to this expectation, which include merit and need. I then provide an argument for why 6-year-olds might be expected to understand that inequality that occurs by choice is also an exception. After reviewing this literature, I will describe five experiments with 6-year-olds, 8-year-olds, and adults, which show a developmental

shift between early childhood and adulthood in response to inequality that occurs by choice. To foreshadow the results, 6-year-olds respond negatively to unequal outcomes even when the person with less chooses to take less, 8-year-olds are ambivalent (some think it is okay while others do not), and adults think that it is okay for one person to take less if that is all the person wanted.

Equality Bias

Beginning in infancy, children have a strong bias toward resource equality (Meristo, Strid, & Surian, 2016; Schmidt & Sommerville, 2011; Sloane, Baillargeon, & Premack, 2012; Sommerville, Schmidt, Yun, & Burns, 2013). For example, looking time studies have shown that infants expect resources to be distributed evenly between individuals. In Meristo et al. (2016), for example, 10-month-olds watched someone distribute two strawberries between two individuals. Infants looked longer when the distributor gave two strawberries to one recipient and none to the other than when the distributor gave each recipient one strawberry.

Looking time studies with infants suggest there is an innate equality preference. The idea that there is an innate predisposition to expect individuals to have the same number of resources corresponds with a social relational theory put forth by Fiske (1992). According to Fiske, there are four models that people use to interpret the social world and one of these models is an equality matching principle. Equality matching describes relationships in which individuals follow a tit-for-tat system; individuals are expected to contribute equally in order to receive an equal share. Interestingly, Fiske thought that equality matching emerged at age four but, as stated above, we now know that even infants expect individuals to be given the same number of resources.

Preschoolers, too, are biased toward resource equality. When given the opportunity to allocate resources between others, they tend to do so equally, even when one might expect them to favor one of the recipients (Baumard, Mascaro, & Chevallier, 2012; Olson & Spelke, 2008; Paulus, 2014). In Olson and Spelke (2008), for example, 3-year-olds were asked to help a protagonist doll allocate resources between four other dolls. The recipient dolls were either the protagonist's siblings, friends, or strangers. Children had to choose how to distribute the items between two siblings versus two friends, two siblings versus two strangers, and two friends versus two strangers. When children were given four items to distribute, most children distributed one item to each of the individuals regardless of their relationship to the protagonist—even when the comparison was between two siblings and two strangers. Of course, this does not address how children actually share resources when they are responsible for allocating resources to their own friends and family members. However, Moore (2009) showed that 4- to 6-year-olds who were given a choice either to split two resources equally between themselves and a peer (1-1) or to keep a single resource for themselves (1-0) chose the split option when the peer was a friend and even when it was stranger. As long as there were enough resources so everyone could receive the same amount and children's own allotment was not at stake, they prioritized equality.

In fact, children's preference for equality is so strong that many would prefer to throw away an extra resource rather than share resources unequally between two individuals. For example, in Shaw and Olson (2012), 6- to 8-year-olds watched as a researcher distributed two colorful erasers to each of two boys. When they were asked to decide whether a fifth and final

eraser should be given to one of the boys (which would entail favoring one of them) or thrown away (which would maintain equality between them), all 20 children in the sample said to throw it away.

The prevailing explanation for why young children have such a bias toward equality is that they consider it a social norm: A rule that has to be followed to maintain social order and cooperative interactions (Turiel, 1983). On this view, inequality conveys favoritism, which can lead to envy and conflict (Fehr, Fischbacher, & Gächter, 2002; McAuliffe, Blake, Steinbeis, & Warneken, 2017; Shaw & Knobe, 2013). One reason to believe that children see equality as a social norm is that they rate individuals who violate equality as less likable than those who do not. In Shaw, DeScioli, and Olson (2012), for example, 6-year-olds watched as two characters distributed erasers to two boys. One character gave two to each of the boys (2-2) whereas the other character gave two to one boy and six to the other (2-6). Note that the latter distribution represents the provision of more erasers overall but demonstrates favoritism towards one of the boys. When children were asked to decide which character was nicer, they chose the one who gave two erasers to both boys.

When is Inequality Acceptable?

According to Shaw (2013), inequality that occurs because of a factor a community approves of—like effort—is not a violation of a social norm. In some cultures, it is permissible for a hard worker to receive more resources than a lazy worker and is even preferable to equality (Damon, 1977; Sigelman & Waitzman, 1991). Similarly, believers of some religions give to the less fortunate rather than more fortunate even though the proximal effect is that they give more to one person (or group) than another.

A number of studies have shown that younger children are less likely to recognize the circumstances under which inequality is acceptable in their culture. For example, when it comes to effort, Rizzo, Elenbaas, Cooley, and Killen (2016) found that 3- to 5-year-olds rewarded lazy workers as much as hard workers whereas 6- to 8-year-olds rewarded hard workers with more resources than lazy workers (see also Nisan, 1989; Shaw & Olson, 2012, Study 3). In addition to allocating resources, children were also asked to evaluate the acceptability of equal allocations and allocations where the hard worker received more resources (i.e. meritorious allocations). Whereas 3- to 5-year-olds evaluated equal allocations as more acceptable than meritorious ones, 6- to 8-year-olds evaluated equal and meritorious distributions as equally acceptable. Thus, in response to merit, 3- to 5-year-olds prioritize equality whereas 6- to 8-year-olds evaluate both equal and meritorious distributions as equally acceptable.

Just as children learn that merit is a culturally acceptable reason to distribute items unequally, they also learn with age that need is an acceptable reason to distribute items unequally. For example, Rizzo and Killen (2016) found that 3- to 4-year-olds distributed an equal number of resources to poor and rich characters, but 5- to 8-year-olds distributed more to the poor character (see also Li, Spitzer, & Olson, 2014; Paulus, 2014). In addition to asking children to allocate resources, Rizzo and Killen also asked children to judge a series of distributions between poor and rich characters. Overall, children judged equal allocations less positively with age and evaluated giving more items to the poor character more positively with age. Thus, as children get older, they attempt to rectify inequality that appears to have been brought about unjustly without a reasonable explanation.

An important question to ask is whether children merely accept inequality under some conditions or whether they think there is a moral imperative to distribute resources unequally in these conditions. In Rizzo and Killen (2016), children's evaluations were compared to a neutral response that fell in the middle between objecting and endorsing. In response to equitable distributions where the poor character was given more items, 3- to 4-year-olds responded neutrally while 5- to 8-year-olds responded positively, judging them as "okay." An age difference also emerged in response to equal allocations, but the breakdowns were slightly different. Three- to 6-year-olds responded to equal allocations positively, judging them as "okay," but 7- to 8-year-olds responded neutrally. In sum, when it comes to need, 6-year-olds seem to be in a transition stage, endorsing both equal and equitable distributions as "okay." This pattern of results is similar to those described earlier in regards to merit where 6-year-olds evaluated both equal and meritorious allocations as equally acceptable.

In conclusion, by age 6, children believe merit and need are two conditions under which giving more items to one person than another is acceptable.[1] Of interest in the studies to be described here is whether children at this age also believe it is acceptable for one person to end up with less if s/he chooses to take less.

Competencies Needed to Recognize Choice Inequality as Okay

Having described young children's apparent bias toward resource equality as well as a number of conditions under which they distribute and accept unequal distributions, I turn now toward considering two specific competencies children would need in order to recognize that it is

[1] While the ultimate goal of distributing more resources to someone in need is to bring about equality, it is only with age that children learn that it is acceptable to deviate and share more with someone in need. Thus, need is a relevant domain to study when investigating children's willingness to accept unequal distributions.

okay for one person to *choose* to take less than another. First, they would need to understand that someone could want less than another person, and second, they would need to understand that people act on their desires when they can freely choose their actions.

Desire. By at least the end of the preschool period, children understand that people can have different desires. For example, even before their second birthday, children begin to use words such as "want" to inform others of their desires (e.g., "I wanna cookie"; Bartsch & Wellman, 1995, p. 75-76), and by their third birthday, children talk about having desires that differ from other people's (e.g., "I want blue. What color you want?"; Bartsch & Wellman, 1995, p. 78).

Experimentally, there is evidence that children understand other people's desires can differ from their own beginning at 18-months of age. In Repacholi and Gopnik (1997), for example, researchers first presented children with a bowl of Goldfish crackers and a bowl of broccoli to assess which one they preferred. Unsurprisingly, the majority (93%) of 18-month-olds preferred the Goldfish crackers. Next, toddlers watched an experimenter taste the two foods. In one condition, the experimenter expressed the same desire that children had just shown: She displayed happiness when she ate the Goldfish crackers (e.g., "Mmm!") and disgust when she ate the broccoli (e.g., "Eww!"). When she placed her hand between the two bowls and requested some food, 76% of children gave her crackers. In the other condition, the experimenter expressed disgust upon tasting the Goldfish crackers and happiness upon eating the broccoli. When she placed her hand between the two bowls, just 31% of children gave her crackers—most gave her broccoli. Many toddlers seemed to understand that the researcher could have a different desire from their own, and they acted on that understanding by giving her what she desired (rather than what they desired).

In addition to conveying information about preferences, desire information provides important cues about how individuals will act and feel. In Wellman and Woolley (1990), children listened to a series of stories in which a character wanted to find something, such as his dog, and were told that the dog could be in one of two locations. After going to look in the first location, the character either found the dog, did not find anything, or found something else. Children as young as 2 years of age believed that the character would stop looking for the dog once he found it but would continue looking for it in the other two conditions. In addition, they knew that the character would feel happy if the dog was found and sad if it was not. Thus, toddlers believe that individuals will behave in ways that are consistent with their desires, and that their emotions can depend on whether those desires are fulfilled.

To my knowledge, children's understanding of desire has not been examined in relation to quantity—that is, whether children recognize two people can want different amounts of a resource. But it seems reasonable to expect (and Study 2 described below demonstrates) that by 6 years of age (and probably sooner), children can recognize that two people can desire different amounts of a resource, people will take the number of items they want, and people will feel happy when they possess the number of items they want even if it is fewer than their fair share.

Free will. A second competence children need in order to understand that inequality by choice is okay is at least some rudimentary understanding of free will. Specifically, they need to understand that someone could freely choose to act in a particular way (even when it may be different from how the child would act or what the child believes is in the individual's best interest), and they need to be able to distinguish that from a situation where someone had no choice but to act in that way.

There is some evidence that, by preschool, children can reason about whether individuals in a given situation freely chose to take a particular action or whether they had to take that action because of the constraints of the situation. In Kushnir, Gopnik, Chernyak, Seiver, and Wellman (2015), for example, 4-year-olds heard a story about a character who wanted to carry out either a physically possible or physically impossible action. In the physically possible story, children heard that a girl wanted to step off of a stool onto the ground and then did so. They were asked whether she *had* to step off the stool or whether she could have stayed on the stool. Most children correctly said that she could have chosen to stay on the stool. In the physically impossible story, they heard that the girl wanted to float in the air forever. Children were told that she, too, stepped off the stool onto the ground, and they were asked if she *had* to step off the stool and go down to the ground or whether she could have floated in the air. Most children correctly said that the character had to step off the stool onto the ground. In other words, preschoolers used their knowledge about possible and impossible actions to reason about whether characters could *choose* to act otherwise or were constrained by the physical limitations of the environment. This finding suggests that 6-year-olds in the studies to be described here should be able to recognize that a character who chose to take fewer than his or her fair share could have done otherwise.

Children also make different inferences about a person's preferences, depending on whether her behavior was freely chosen or constrained. In Pesowski, Denison, and Friedman, (2016, Study 2), 5-year-olds were introduced to a character who saw two toys on a bookshelf: one was attractive and one was less attractive. When both toys were within the character's reach and she took the less attractive one, children inferred that she liked it more than the attractive toy. But when only the less attractive toy was within the character's reach and she took that toy,

children inferred that she actually liked the attractive and out-of-reach toy more. They recognized that desires can be inferred from actions only when an individual freely makes their choices. Together with the Kushnir et al. (2015) study, this work demonstrates that young children recognize that someone who chooses to carry out an action probably wants to do so, and someone who is forced to carry out that action may not.

Interestingly, there is one study that suggests 6-year-olds may not appreciate the role of free will when it comes to resource distribution. McAuliffe, Jordan, and Warneken (2015) showed that many 6-year-olds objected to scenarios in which one character generously and of her own free will gave all of her resources to another character. In this study, 6-year-olds watched as a protagonist who controlled six candies either kept them all for herself, gave them all to another peer, or split them equally between herself and the peer. Participants could reject the protagonist's allocation by pulling a handle that caused all of the candies to "disappear" (in which case neither character got any). Not surprisingly, participants rarely rejected the protagonist's allocation when she divided the resources equally (3-3), but often did so when she distributed the candies selfishly (6-0). Interestingly, children also sometimes rejected the allocation when the protagonist divided the resources generously—when she gave all of the candy to the other character (0-6). They did not reject the generous allocations as often as the selfish ones, but they did reject the generous allocations more often than the equal ones.

One explanation is that some 6-year-olds were averse to any inequality--even though in the generous case, the person with less was responsible for the inequality. It is also possible that some 6-year-olds objected to the fact that the individual who made the generous offer ended up with no candies. Perhaps they would not have objected had the generous protagonist kept at least one candy for himself (Paulus, 2016). Previous research has shown that even when children

allocate resources unequally, as in the case of merit, they do not allocate all of the resources to the more deserving person, but rather opt to give the other person fewer resources, giving everyone *something* (e.g., Baumard et al., 2012; Kenward & Dahl, 2011). In the studies described in this dissertation, the character who takes less still ends up with at least one resource so s/he has not deprived himself or herself entirely of having resources. Thus, if children object to inequality that occurs by choice, it cannot be based on an individual denying himself or herself of resources.

Summary

Six-year-olds have a tenuous understanding of the conditions under which resource inequality is okay. They think it is acceptable for people who work harder to be compensated with more resources, but many also object to when one person acts generously and offers another all of his or her own resources. Given that 6-year-olds' ability to correctly evaluate resource inequality as fair or not shifts depending on the context, it is not evident—despite having many of the necessary competencies— that they will recognize inequality is okay when one person has chosen to take less.

The Present Studies

In Study 1, I investigated how 6-year-olds, 8-year-olds, and adults evaluated scenarios in which one person chooses to take less than half the resources and the second person takes the remainder. In Study 2, I investigated whether 6-year-olds understand that two people can desire different quantities of a resource, and if their predictions about how many items character will take are in line with how many items they think the characters *should* take. In Studies 3 and 4, I investigated which of several norms children who objected to inequality by choice in Study 1 felt had been violated. Finally, in Study 5, I investigated whether 6-year-olds would continue to

object to what I'm calling "inequality by choice" when the amount of cognitive effort needed to override the "equal outcome norm" was reduced.

I focused on 6-year-olds because this is the youngest age at which children have been shown to use merit and need as an acceptable reason for unequal distributions (e.g., Nisan, 1989; Rizzo et al., 2016; Rizzo & Killen, 2016; Shaw & Olson, 2012). I expected that by six years of age, children would have learned based on experience and observation that inequality based on choice is another justifiable condition under which inequality is acceptable. I included 8-year-olds and undergraduates as comparison groups. Previous research has found that 8-year-olds are more willing to accept inequality than 6-year-olds under two conditions: Eight-year-olds are more willing to advocate for inequality by choice in a first-party context and are more willing to accept inequality when it occurs as the result of a fair procedure, such as when it occurs by chance (e.g., Shaw, Choshen-Hillel, & Caruso, 2016; Shaw & Olson, 2014).

Study 1

Method

Participants. Sixteen 6-year-olds (M = 78.00 months, SD = 3.29 months; range = 73 months to 83 months; 8 girls), sixteen 8-year-olds (M = 101.69 months, SD = 3.61 months; range = 96 months to 107 months; 8 girls), and sixteen undergraduates (M = 19 years, SD = 1 year; range = 17 to 20 years; 13 women) participated in a single 15-min session. Four additional participants (one 6-year-old; three adults) were excluded for failing the catch trials (described below). Children were primarily white and from middle-class backgrounds, and were recruited from a database of families that expressed interest in participating in research. Undergraduate participants participated in exchange for course credit.

Procedure. Children sat at a table in front of 13" laptop and next to an experimenter. Each child participated in a training phase and a test phase. Adults did not participate in the training phase, but were provided with written directions about how to complete the study independently.

Training phase. The training phase was derived from previous work in early social and moral development (e.g., Rizzo et al., 2016; Rizzo & Killen, 2016; Smetana, 1981; Tisak & Turiel, 1988). Children first evaluated whether the following four scenarios were "okay"/"not okay:" giving your mom a hug, kicking a friend, wearing pajamas to school, and waving hello to someone (in that order). All children evaluated the hugging and waving (prosocial) scenarios as "okay." All children evaluated the kicking scenario (a moral transgression) as "not okay," and all but two children (one 6-year-old; one 8-year-old) reported that wearing pajamas to school (a social-conventional transgression) was "not okay" (the children who responded incorrectly were immediately corrected).

After children indicated that kicking a friend or wearing pajamas to school was "not okay," they evaluated how bad that transgression was. Three faces with frowns that increased in intensity from left to right appeared on the laptop. The experimenter asked, "How bad is it to [kick your friend]/[wear pajamas to school]? Is it a tiny bit bad, really bad, or really, really bad?" (1 = tiny bit bad; 2 = really bad; 3 = really, really bad). Consistent with previous research showing that children believe moral transgressions are more serious than social-conventional ones (e.g., Smetana, 1981; Tisak & Turiel, 1988), 6- and 8-year-olds judged kicking a friend to be worse (M = 2.50, SD = .57) than wearing pajamas to school (M = 1.63, SD = .71). A 2 × 2

(Age × Transgression Type) mixed analysis of variance (ANOVA) yielded a main effect of transgression type, $F(1, 30) = 32.85$, $p < .0005$, $\eta_p^2 = .52$, 95% CI [.56, 1.19], and no difference across the two age groups, $F(1, 30) = .14$, $p = .714$, $\eta_p^2 = .01$.

Before proceeding to the test phase, the experimenter prompted children to point to the faces in the following order: "tiny bit bad," "really, really bad," and "really bad." All but three 6-year-olds pointed to the three faces on the severity scale correctly. The three children who initially pointed to the wrong face were corrected and prompted again at the end of the training to point to the face they had identified incorrectly; all three did so successfully.

The training phase remained the same in Studies 3 and 4 (Study 2 employed a different paradigm and did not have a training phase; Study 5 did not use the severity scale). In Study 3 and Study 4, children continued to evaluate kicking a friend as worse than wearing pajamas to school.

Test phase. The experimenter explained to children that they would next listen to some stories and decide if what happened was "okay" or "not okay." On each of eight trials in this phase, participants saw a pair of cartoon boys or girls (different pairs on each trial, gender alternating across the trials) with a table between them. On the table was a plate with four pieces of the same desirable food item (different items on each trial): Hershey's Kisses, chocolate chip cookies, animal crackers, Skittles, M&M's, Oreos, Goldfish crackers, and lollipops.

The eight test trials were comprised of four "choose-one" and four "choose-three" trials. On the choose-one trials, participants heard one character say, "I only want one;" watched as s/he took one item; heard the other character say, "Hmmm;" and watched as s/he took the remaining three items. On the choose-three trials, they heard one character say, "I want three;" watched as s/he took three items; heard the other character say, "Hmmm;" and watched as s/he

took the only remaining item. The character who spoke first was on the left on the first and fourth trials of each block for half the participants at each age, and on the right on the first and fourth trials of each block for the other half. Figure 1 shows an example of a final frame from a test trial (an example of an entire test trial animation is available at https://osf.io/wpbmj/?view_only=27c91fa131964e958c7922a93bb8ca66).

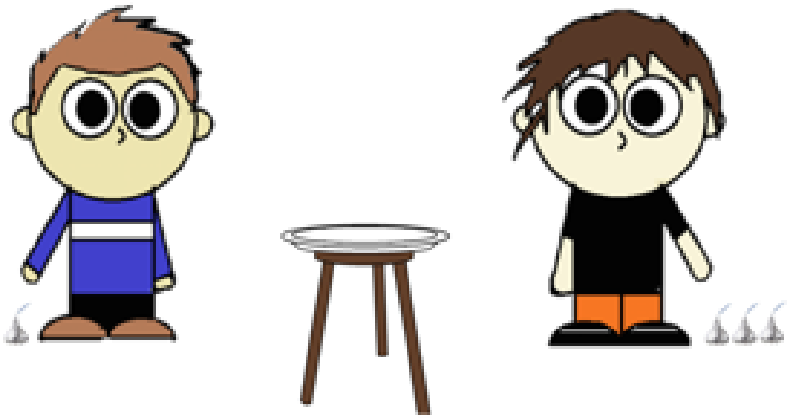

Figure 1. Example of image depicting final allocation in Study 1.

At this point, the researcher asked children about the depicted outcome: "Is it okay or not okay that this girl/boy got one and this girl/boy got three?" On trials when they indicated that an unequal outcome was "not okay," the researcher asked them to indicate how bad it was using the three-point face scale. Adult participants completed the task on their own, using a worksheet. If adults thought the outcome was "not okay," the worksheet instructed them to indicate if they thought the outcome was "a tiny bit bad," "really bad", or "really, really bad." Finally, children were asked and adults were instructed in writing to indicate which one of the characters felt happier or if they felt the same.

The four choose-one trials were always presented before the four choose-three trials because my primary interest was in understanding how participants evaluated situations where one person *chose* to take fewer than his/her fair share. Based on work described in the

Introduction showing that children attempt to rectify unjust inequality (e.g., Li et al., 2014; Paulus, 2014; Rizzo & Killen, 2016), I expected (and found) that participants at all three ages would object overwhelmingly to the choose-three trials. I did not want this response pattern to prime or influence their responses to the choose-one trials.

At the end of the session, I included two catch trials where one character expressed a desire for and took two items and the other character took the remaining two items. The items displayed on the table were four Teddy Grahams or gummy bears. Based on previous work showing that children as young as 10 months believe equally deserving individuals should receive equal resources (e.g., Meristo et al., 2016), I expected participants at all ages to judge these trials to be "okay." If participants judged either or both of these catch trials to be "not okay," it could suggest that they had not been attending to the narration and/or the animations.[2] Additionally, these catch trials ensured that participants were able to use the "okay" side of the scale and were not primed by the test question to report that the outcome was "not okay." As noted earlier, three adults and one 6-year-old were excluded from the final analyses because they said that one or both of these equal outcomes were not okay. (The pattern of results reported here was the same regardless of whether they were included.)

After the test trials and catch trials, participants watched the first choose-one trial again and were asked to explain why they had earlier described that unequal outcome as "okay" or "not okay."

Results

Preliminary analyses revealed no main effect of or interactions with gender on any of the measures; therefore, gender will not be considered in the analyses below.

[2] Three 6-year-olds participated before the catch trials were added to the procedure. These participants are included in the final sample.

"Not okay" judgments. Figure 2 shows the average number of times that participants at each of the three ages said the final distribution of items was "not okay" as a function of trial type. As the figure shows, on the choose-three trials, where the first character chose to take three items and left the second character with just one, 6-year-olds, 8-year-olds, and adults almost always indicated that the outcome was "not okay," and all did so more often than would be expected by chance of two trials, *t*s(15) > 11.21, *p*s < .0005, *d*s > 2.80, 95% CIs [1.37, 2.01], [1.80, 2.07], [no variability], respectively. At the individual level, 12/16 six-year-olds, 15/16 eight-year-olds, and 16/16 adults said that all four choose-three trials were "not okay."

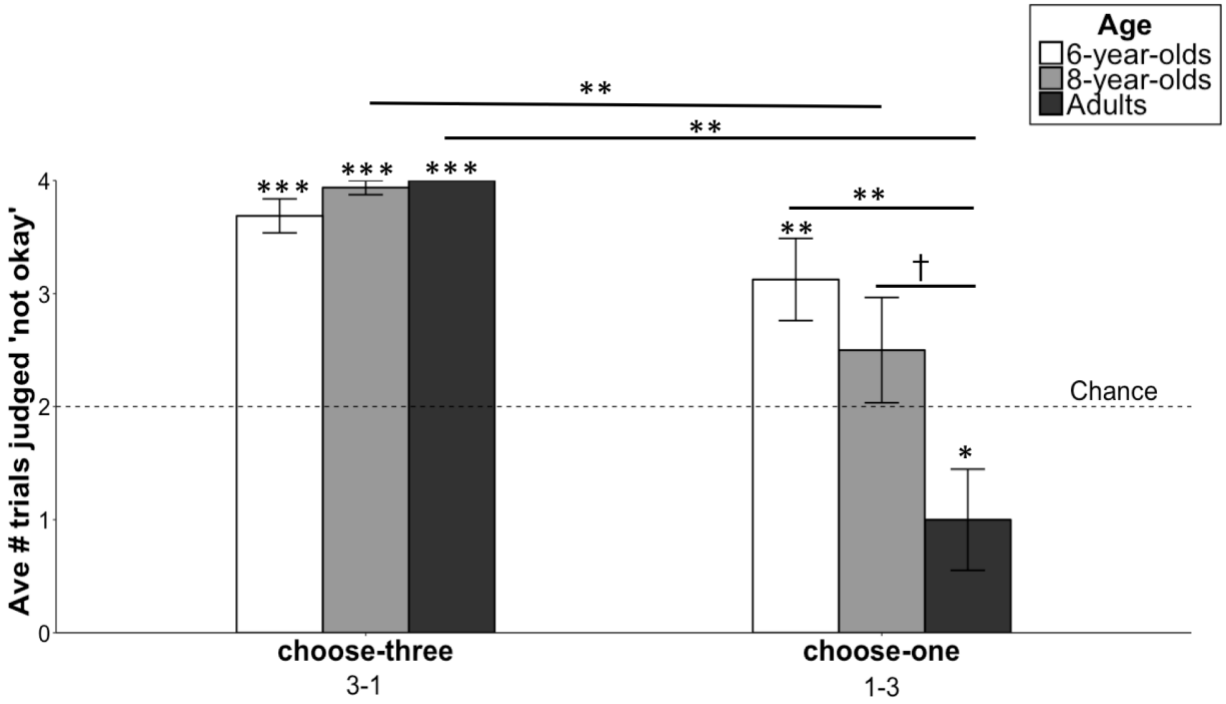

Figure 2. Average number of trials on which participants evaluated unequal outcomes as "not okay" in Study 1. Dashed line indicates chance level. Error bars show SEM. † $p < .10$ * $p < .05$ ** $p < .01$ *** $p < .0005$

In contrast, on the choose-one trials, where the first character chose to take one item (but could have taken more), and the other character took the three remaining items, age differences emerged. As Figure 2 shows, 6-year-olds said it was "not okay" on more trials than expected by

chance, $t(15) = 3.09$, $p = .007$, $d = .77$, 95% CI [.35, 1.90], 8-year-olds were ambivalent, $t(15) = 1.07$, $p = .300$, $d = .27$, 95% CI [-.49, 1.49], and adults were less likely than chance to say this was "not okay" (i.e., they said it was "okay" more than chance), $t(15) = 2.24$, $p = .041$, $d = .56$, 95% CI [-1.95, -.05]. At the individual level, 11/16 six-year-olds, 9/16 eight-year-olds, and just 4/16 adults said that all four choose-one trials were "not okay."

A 3 × 2 (Age × Trial Type) mixed ANOVA on the data in Figure 2 yielded a main effect of age, $F(2, 45) = 4.38$, $p = .018$, $\eta_p^2 = .16$, a main effect of trial type, $F(1, 45) = 47.66$, $p < .0005$, $\eta_p^2 = .51$, and an Age × Trial Type interaction, $F(2, 45) = 8.72$, $p = .001$, $\eta_p^2 = .28$. To follow up on the interaction, I conducted simple effects tests using Bonferroni correction comparing the average number of choose-three and choose-one trials evaluated as "not okay" within each age. While 6-year-olds judged the same number of choose-three and choose-one trials as "not okay," $p = .185$, 95% CI [-.28, 1.41], 8-year-olds and adults judged more choose-three than choose-one trials as "not okay," ps < .001, 95% CIs [.60, 2.28] and [2.16, 3.84], respectively. To investigate age differences within the choose-one trials, I conducted Bonferroni post-hoc comparisons on the number of choose-one trials participants at each age judged to be "not okay." Six-year-olds judged more choose-one trials to be "not okay" than did adults, $p = .003$, and the same number as 8-year-olds, $p = .921$. Eight-year-olds judged marginally more choose-one trials to be "not okay" compared to adults, $p = .051$.

Analyses at the individual level yielded the same pattern of results. Table 1 shows the number of participants at each age judging more choose-three than choose-one trials to be "not okay," the reverse, and the number judging the same number of each to be "not okay." There was no difference in the number of six-year-olds who judged more choose-three than choose-one trials as "not okay" than the reverse (Wilcoxon signed-rank: $Z = 1.63$, $p = .104$); most judged the

same number of each (i.e., four) to be "not okay." In contrast, more adults and 8-year-olds judged more choose-three than choose-one trials to be "not okay" than the reverse (Wilcoxon signed-rank: $Z = 3.46$, $p = .001$, $Z = 2.45$, $p = .014$, respectively).

Table 1

Number of Participants at Each Age in Study 1 Judging More Choose-Three or Choose-One Trials as "Not Okay," or the Number who Judged the Same Number of Each Trial Type as "Not Okay"

	More Choose-Three	More Choose-One	Same
6-year-olds	4	1	11
8-year-olds	7	0	9
Adults	12	0	4

Table 2 shows the number of participants at each age who said the choose-one trials were not okay on 0-1 trials, 2 trials, or 3-4 trials as a function of age. As the table shows, most 6- and 8-year-olds said the choose-one trials were "not okay" on three or four of the four trials, while few of the adults did so. Fisher exact tests showed that the distribution of 6-year-olds and adults who evaluated the choose-one trials as not okay 0-1 versus 3-4 times differed, $p = .003$. The distribution of 8-year-olds did not differ from that of 6-year-olds, $p = .385$, and was marginally different from that of adults, $p = .063$.

Table 2

Distribution of Participants' Choose-One "Not Okay" Evaluations in Study 1

	0-1 trials	2 trials	3-4 trials
6-year-olds	2	3	11
8-year-olds	5	2	9
Adults	12	-	4

Severity ratings. When participants evaluated a scenario as "not okay," they rated how bad it was on a 3-point scale. I was interested in whether choose-three trials judged to be "not okay" would be evaluated as worse than choose-one trials judged to be "not okay." After all, on choose-three trials, the character who expressed a desire for three and took three resources was

greedy whereas the character who took three resources on the choose-one trials was only taking resources the first character did not want. Only the 14 six-year-olds and 11 eight-year-olds who judged at least one choose-three and one choose-one outcome as "not okay" could contribute to this analysis (in fact, all of these children judged at least two of each trial type to be "not okay"). I chose to focus this analysis on data from the children because as Table 2 shows, only 4 of the 16 adult participants would have contributed data to this analysis.

Figure 3 shows the average severity rating as a function of age and trial type for participants who said both types of trials were "not okay" on at least two trials. A 2 × 2 mixed ANOVA (Age × Trial Type) on these data yielded a main effect of age, $F(1, 23) = 7.17$, $p = .013$, $\eta_p^2 = .24$, indicating that when 6-year-olds judged unequal outcomes to be "not okay," they rated them as worse than 8-year-olds. There was additionally a marginal Age × Trial Type interaction, $F(1, 23) = 3.69$, $p = .067$, $\eta_p^2 = .14$. Follow-up simple effects using Bonferroni correction showed that whereas 6-year-olds rated the two trial types as equally bad ($p = .495$, 95% CI [-.38, .19]), 8-year-olds rated choose-three trials as marginally worse than choose-one trials ($p = .063$, 95% CI [-.02, .62]). This provides a hint that 8-year-olds (but not 6-year-olds) were distinguishing between the two types of inequality, although some caution is needed in this interpretation because the interaction and follow-up tests were only marginal, the order of choose-one and choose-three trials was not counterbalanced, and the sample size included only a subset of the participants.

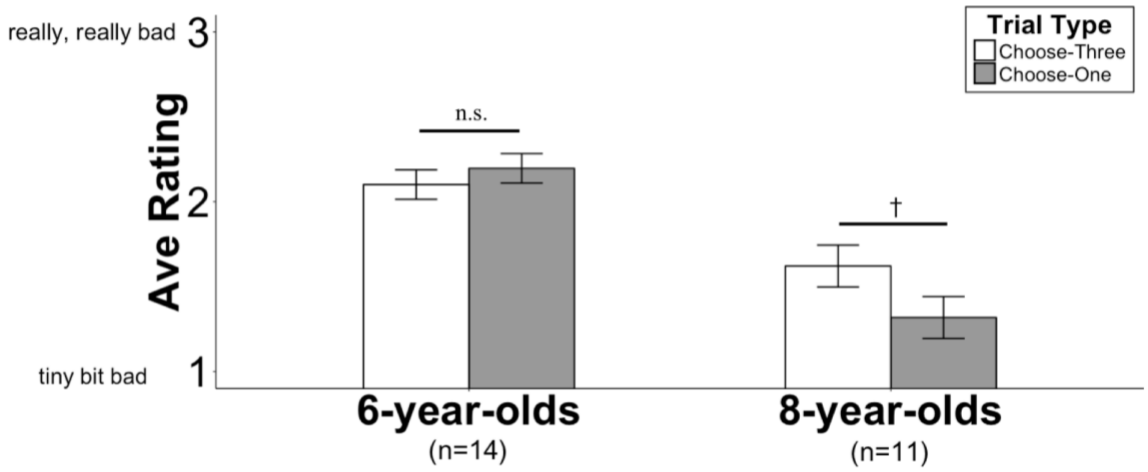

Figure 3. Average severity ratings on choose-three and choose-one trials among participants who rated choose-one trials as "not okay" on at least two out of the four trials in Study 1. Error bars show SEM. Error bars show SEM. † $p < .10$

Relative emotion judgments. At the end of each trial, participants were asked whether the character with three items was happier than the character with one item, the reverse, or whether they were equally happy. For the choose-three trials, I expected that participants would infer the first character who initially took three items felt happier than the second character who was left with just one item. In contrast, for the choose-one trials, I expected that participants might judge the two characters to be equally happy; the first character had asserted a desire for just one item and children from 2 years of age know that people are happy when their desires are fulfilled (Wellman & Woolley, 1990).

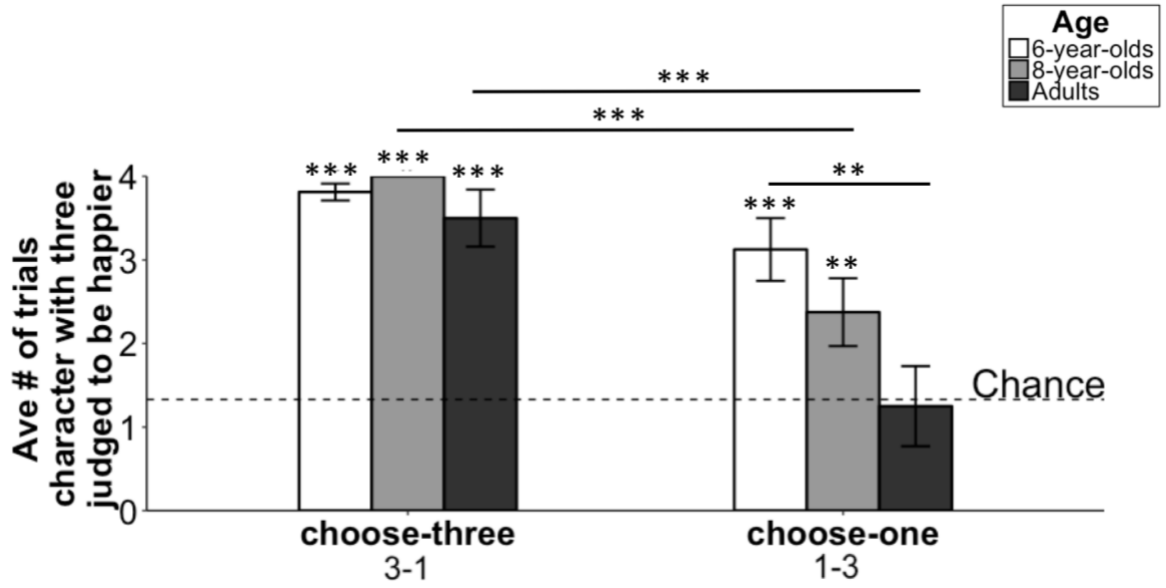

Figure 4. Average number of times participants in Study 1 evaluated the character with three items as happier as a function of trial type. Dashed line indicates chance level (1.33 trials). Error bars show SEM. ** $p < .01$ *** $p < .0005$

Figure 4 shows the average number of times that participants judged the character with three items as happier as a function of trial type. The pattern of results mirrored those from the judgment ratings in Figure 2. As Figure 4 shows, on the choose-three trials, 6-year-olds, 8-year-olds, and adults almost always said that the character who took three items was happier than the character who ended up with one, and did so more often than would be expected by chance (chance is 1.33 because the three possible response options for who was happier were character 1, character 2, or the same), *t*s(15) > 6.35, *p*s < .0005, *d*s > 1.59, 95% CIs [2.27, 2.70], [no variability], and [1.44, 2.90], respectively. On the choose-one trials, age differences emerged. Six- and 8-year-olds judged the character who ended up with three items as happier more often than chance, *t*s(15) > 2.57, *p*s < .021, *d*s > .64, 95% CIs [1.00, 2.59] and [.18, 1.91], respectively, but adults did not, $t(15) = .167$, $p = .870$, $d = .04$, 95% CI [-1.10, .94]. When participants did not

evaluate the character with three as happier, they tended to rate the two characters as equally happy. (The character who ended up with one was selected as happier, on average, on just 1.56% of choose-three trials and 3.13% of choose-one trials.)

A 3 × 2 (Age × Trial Type) mixed ANOVA on the data in Figure 4 yielded a main effect of age, $F(2, 45) = 4.95$, $p = .011$, $\eta_p^2 = .18$, a main effect of trial type, $F(1, 45) = 38.38$, $p < .0005$, $\eta_p^2 = .46$, and a significant Age × Trial Type interaction, $F(2, 45) = 3.42$, $p = .041$, $\eta_p^2 = .13$. Simple effects using Bonferroni correction revealed that 6-year-olds evaluated the character with three items as happier than the character with one item on the same number of choose-three and choose-one trials, $p = .113$, 95% CI [-.17, 1.54], but 8-year-olds and adults evaluated the character with three items as happier on more choose-three than choose-one trials, ps < .0005, 95% CIs [.77, 2.48] and [1.39, 3.11,], respectively. To investigate age differences within the choose-one trials, I conducted Bonferroni post-hoc comparisons on the number of choose-one trials that participants at each age evaluated the character with three items as happier. Six-year-olds judged the character with three items as happier on more choose-one trials than did adults, $p = .009$. The number of times that 8-year-olds judged the character with three items as happier fell in-between that of 6-year-olds and adults and did not differ from either age group, ps > .199.

Justifications. After the catch trials, participants were reminded of whether they said the first choose-one trial was "okay" or "not okay," and they were asked to explain that choice. I asked children to justify their first judgment after the eight test trials and two catch trials had been administered because I did not want their justifications to influence their evaluations on the subsequent test trials. Figure 5 shows the number of participants who said the first trial was "okay" or "not okay," as a function of age. As the figure shows, the number of participants at each age who evaluated the first choose-one trial as "not okay" follows the same pattern as the

data from the average number of choose-one trials that were evaluated as "not okay" in Figure 2: Most 6-year-olds, around half the 8-year-olds, and fewer than half the adults evaluated the first choose-one trial as "not okay."

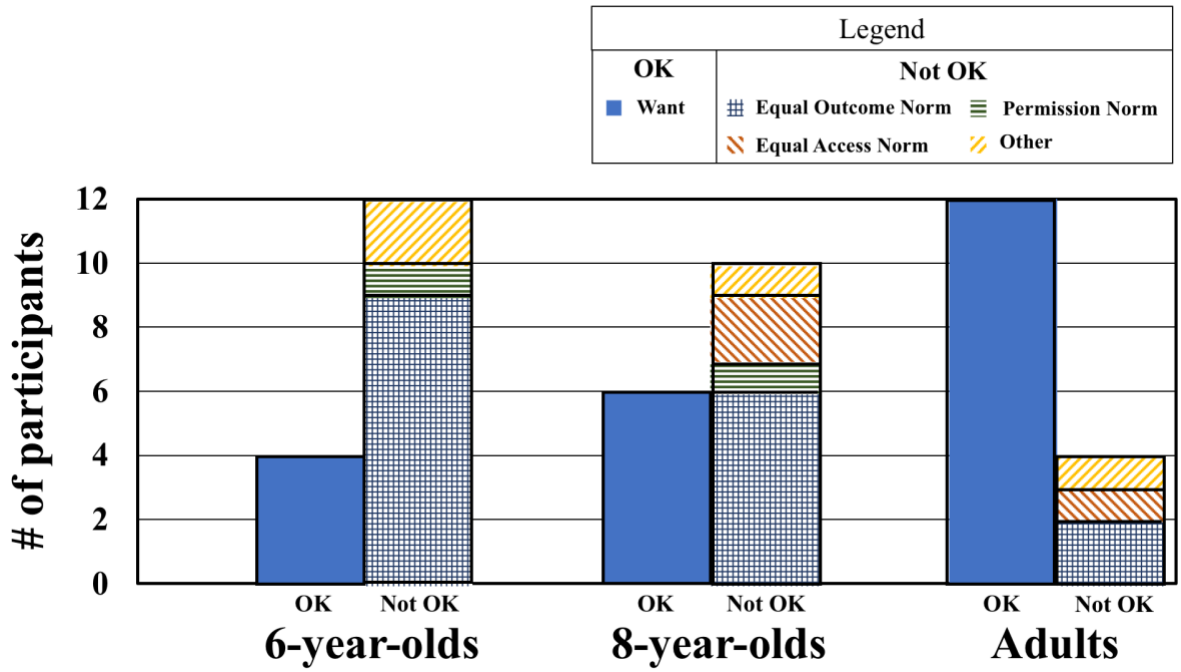

Figure 5. Participants' justifications in Study 1 as a function of their first choose-one evaluation trial.

"Okay" justifications all referenced the fact that the character who ended up with one had only wanted one. "Not okay" justifications fell into one of four categories: a) violations of an *equal outcome norm*: it was not okay because the two characters did not get the same amount[3]; b) violations of a *permission norm*: it was not okay because the character who took three should have asked the other character for permission; c) violations of an *equal access norm:* it was not okay because the first character might change his mind and would no longer have access to his fair share of half the resources; d) other comments. Table 3 shows examples of each category.

[3] In principle saying it is not fair could refer to multiple aspects of the scenario. However, since the default explanation refers to outcome, I have decided to subsume fairness with the outcome category.

The first author and a second coder independently coded the justifications. Agreement was nearly perfect (agreement = 97%, Cohen's kappa = .96, $p < .0005$). The one disagreement was resolved through discussion.

Table 3
Study 1 First Choose-One Trial "Not Okay" Evaluation Justification Categories and Examples

Category	Age	Examples
Equal Outcome Norm	6-year-old	"Because he only got one and he got three."
	6-year-old	"Because they didn't get the same amount."
Permission Norm	6-year-old	"[The boy with three] should ask [the boy with one if he wants another one."
	8-year-old	"[The boy with three] should have said, "Can I have the rest or are you sure that's all you want?"
Equal Access Norm	8-year-old	"If [the boy with one] changes his mind, there will be none left."
	Adult	"Now that [the boy with three] has taken all the chocolate, the other boy can't really go back and change his mind."
Other	Adult	"It was opportunistic, but they were only Hershey kisses so not that bad."

As Figure 5 shows, most 6-year-olds who judged the first choose-one trial to be "not okay" described it as an equal outcome norm violation, explaining that it was "not okay" because one character got fewer (or more) than the other. While many of the 8-year-olds who objected also described an equal outcome norm violation, a few raised concerns about the violation of the permission or equal access norms. I will return to the concerns 8-year-olds raised in Study 3.

Discussion

There are two primary findings from Study 1. First, participants at all ages objected to inequality on choose-three trials. When the first character took three items, leaving only one for the second character, participants overwhelmingly indicated that it was "not okay," and they judged the character who ended up with three items as happier. Second, there was a developmental shift between 6 years of age and adulthood in response to inequality on choose-one trials. When the first character took one item and left three for the second, most adults judged this to be okay and evaluated the two characters as equally happy. In contrast, 6-year-olds

objected to choose-one trials as often as they did to choose-three trials, rated the violation as serious as they did on choose-three trials, and judged the character who ended up with three items as happier as often as they did on choose-three trials. Eight-year-olds' acceptance of inequality that occurred by choice and their evaluations about how the two characters felt fell somewhere between the 6-year-olds and adults.

Why did 6-year-olds respond in the same way on the choose-three and choose-one trials on all three measures ("not okay" judgments, severity ratings, relative emotion judgments), objecting to both types of trials to the same extent? One possibility is that they did not understand (or believe) on choose-one trials that the first character's stated desire for one tasty treat would lead him to take just one when four were available. If this were the case, 6-year-olds may have judged the choose-one trials to be "not okay" because they expected the first character would take his fair share of the available resource (i.e., two) regardless of his or her stated desire. On this account, because the depicted outcome showed the first character taking just one, they objected. This explanation seems unlikely in light of the earlier described research showing that children younger than those in my study expect people to act according to their desires even if those desires do not match the children's own (e.g., Repacholi & Gopnik, 1997), and they expect people to be satisfied when their desires are fulfilled (e.g., Wellman & Woolley, 1990). However, the previous work has addressed children's understanding that people could want different *kinds* of things rather than different *amounts* of the same thing.

A second possibility is that 6-year-olds might have understood and believed that someone would take only one of four appetitive items, but also believed that that person ***should*** have taken

their fair share nonetheless. That is, perhaps they were responding in accordance with the equal outcome norm: It was not okay for a character to take just one item on choose-one trials because they should have taken two.

Study 2 was designed to investigate these two possibilities. Six-year-olds and adults (for comparison purposes) saw the pairs of characters and food items presented in Study 1 and heard the characters express the same verbal information. One group of participants was asked to predict how many items each character *will* take while another group of participants was asked how many items each character *should* take. The *will* question assessed whether children understood that two people could desire different quantities of the same resource while the *should* question addressed how many items children thought each character ought to take even after they had expressed their desire.

Study 2

Method

Participants. Sixteen 6-year-olds (M = 78.75 months, SD = 3.67 months; range = 72 months to 84 months; 8 girls) and sixteen undergraduates (M = 19 years, SD = 1 year; range = 18 to 20 years; 8 women) participated in a single 15-min session. Four additional children participated but three were excluded for failing one of the catch trials (described below) and one was excluded due to lack of fluency in English. Children were recruited from the same database of families as Study 1; undergraduate participants participated in exchange for course credit.

Procedure. Children sat at a table in front of a computer and next to an experimenter who explained to children that they would listen to some stories and answer some questions about them. Adults were provided with written directions about how to complete the study independently.

Design. Eight participants at each age took part in a will condition, and eight at each age took part in a should condition, balanced at each age for sex.[4] The average age of children in each condition was, respectively, 6;6 and 6;7.

The animations were exactly the same as those from Study 1 except that participants only heard the desire statements and did not actually see the characters take any items from the table. As in Study 1, participants were presented with 10 trials (8 test trials; 2 catch trials). The four choose-one trials occurred first: Participants heard one character express a desire for one of four available items and the other character say "Hmmm." The four choose-three trials occurred second: They heard the first character express a desire for three of four available items and the other character say, "Hmmm." On the final two catch trials, the first character expressed a desire for two of the four items, and the second character said, "Hmmm."

In the will condition, participants were simply asked, "How many will s/he [the first character] take? And how many will s/he [the second character] take?" In the should condition, they were asked, "How many should s/he [the first character] take? And how many should s/he [the second character] take?" Within each trial block, the character who spoke first was on the left (or right) on the first and fourth trials, and on the catch trials, the character who spoke first appeared once on the left and once on the right.

As noted earlier, four 6-year-olds were excluded from the final analyses; one was not yet a fluent English speaker and the other three failed one of the catch trials, by indicating that the first character would or should take more than the two s/he expressed a desire for. (The pattern of results reported here was the same regardless of whether they were included.)

[4] This study includes half the number of participants per condition included in the other studies because it primarily served as a control study to demonstrate that 6-year-olds understand two people can desire different amounts of a resource. Even with half the number of participants per condition, the results were clear.

Results and Discussion

Six-year-olds and adults in the will condition reported that a character who wants three items will take 2.91 (*SD* = .27) and 2.63 (*SD* = .52) items, respectively, and that a character who wants one item will take, 1.13 (*SD* = .27) and 1.03 (*SD* = .09) items, respectively. A 2 × 2 (Age × Trial Type) mixed ANOVA on these data yielded only a main effect of trial type, $F(1, 14) = 218.31$, $p < .0005$, $\eta_p^2 = .94$, and no main effect of age, $F(1, 14) = 2.70$, $p = .123$, $\eta_p^2 = .16$, or Trial Type × Age interaction, $F(1, 14) = .67$, $p = .425$, $\eta_p^2 = .05$. These results suggest that, consistent with research showing children expect that individuals will act on their desires when it comes to the kind of thing they want (Pesowski et al., 2016), 6-year-olds expect people to act on their desires when it comes to the amount of something they want. In both cases, this is true even when those expressed desires differ from what children themselves likely want.[5]

Results from the should condition suggest that 6-year-olds may have objected to the choose-one trials in Study 1 because they believed that individuals *should* take their fair share regardless of their stated desire—that is, that they objected because failing to take the same number of items represented a violation of the equal outcome norm. On choose-three trials, 6-year-olds and adults thought a character who wanted three items should only take 2.06 (*SD* = .40) and 2.19 (*SD* = .53) items, respectively. This is not surprising given that the character who wanted three was being greedy. However, whereas adults thought a character who wanted only one item should take that amount (*M* = 1.03 items, *SD* = .09), 6-year-olds thought a character who wanted only one item should take two (*M* = 1.94 items, *SD* = .18). A 2 × 2 (Age × Trial

[5] I focused the analyses exclusively on the average number of items that participants reported the first character will/should take within the two trial types. Due to interdependence, I did not analyze the number of items that participants reported the second character (who said "Hmmm") would/should take. Most participants reported that the second character would take all of the remaining items on the table. The number of items participants reported added up to four on all of the trials for 14 out of 16 six-year-olds and 12 out of 16 adults.

Type) mixed ANOVA on these data yielded a main effect of age, $F(1, 14) = 10.34$, $p = .006$, $\eta_p^2 = .43$, a main effect of trial type, $F(1, 14) = 27.30$, $p < .0005$, $\eta_p^2 = .66$, and a Trial Type × Age interaction, $F(1, 14) = 17.69$, $p = .001$, $\eta_p^2 = .56$. To follow up on the interaction, I conducted simple effects tests using Bonferroni correction. There was no difference between the number of items that 6-year-olds said the character who wanted three items and one item should take: They thought both characters *should* take the same number of items (i.e., two in both cases), $p = .483$, 95% CIs [-.25, .50]. In contrast, adults said that the character who wanted one item should take significantly fewer items than the character who wanted three (i.e., one when the character said s/he wanted one, and two when the character said s/he wanted three), $p < .0005$, 95% CI [.78, 1.53].

Six-year-olds' responses in the should condition are consistent with other studies showing that children this age can distinguish between what will happen versus what they think should happen (e.g., DeJesus, Rhodes, & Kinzler, 2014; Smith et al., 2013). For example, 6-year-olds know that they *should* share stickers equally with a peer but accurately predict that they *will* keep more stickers for themselves (Smith et al., 2013). While it makes sense that children would know people *should* be fair but often are not, I was surprised that children think people *should* take more than they want when their desire was less than their fair share. By this age, children know that someone who chooses a less desirable item has a preference for that item and know that individuals feel happy when their desires are fulfilled (Pesowski et al., 2016; Wellman & Woolley, 1990). Although the results from the will condition demonstrate that 6-year-olds understood the desire information conveyed by the character who wanted just one as well as adults, the results from the should condition indicate that their belief about what ought to happen differs: Whereas adults reported that the individual who expressed a desire for only one item

should only take one, 6-year-olds reported that the individual who wanted one item should take two. Six-year-olds' belief that two individuals should take the same amount even if they have different desires explains 6-year-olds' objection to the choose-one trials in Study 1 and the concerns they raised in the justification data about the violation of the equal outcome norm.

I turn now to investigating whether children's objections to the choose-one trials in Study 1 might have also been driven by concerns regarding other norm violations.

Study 3

I did not include 8-year-olds in Study 2 because, as a group, they did seem to distinguish between the choose-one and choose-three trials on all three measures ("not okay" judgments, severity ratings, and relative emotion judgments) in Study 1. That said, 8-year-olds in Study 1 objected to the choose-one trials as often as 6-year-olds. Why did 8-year-olds object to the choose-one trials? Did they believe, as 6-year-olds did, that the two individuals *should* abide by the equal outcome norm and both individuals should end up with the same number of items regardless of their stated desire? Were there other norms that they (and some 6-year-olds) felt were violated?

In Study 3, I address three possible norms that children may have felt were violated in Study 1. First, as most 6- and 8-year-olds who objected to choose-one trials in Study 1 explained in their justifications and as Study 2 demonstrated with 6-year-olds, the two characters ended up with different amounts, violating children's equal outcome norm that equally deserving individuals should get the same amount of a desirable resource (Shaw & Olson, 2012). Second, a few children were concerned that the character who took just one item would not be able to access his/her fair share (i.e., the second item) if s/he later decided s/he did want that fair share. That is, they may have objected to choose-one trials at least in part because the second character

took some of what they believed rightfully belonged to the first character, which I call the "equal access norm." Finally, although children did not articulate this in their justifications, it is possible that some objected to the choose-one trials in Study 1 because when the second character took more items than the first person, s/he violated the "situational norm" established by the first person who only took one. When taking from a shared pool of resources, adults are influenced by how many items individuals who preceded them took. For example, adult participants who were led to believe that other individuals have only taken a few items took fewer themselves than those who were led to believe that the individuals before them had taken a lot (Rutte, Wilke, & Messick, 1987).

Study 3 was designed to investigate which (if any) of these three norm violations contribute to children's objections to the choose-one trials. Participants were shown the same animations as those from Study 1: On choose-one trials, they saw the first character take one item and the second take three, and on choose-three trials, they saw the first character take three items and the second take one. In Study 3, however, the two characters took resources from a bowl of unlimited resources rather than a plate with only four items.

If children in Study 1 objected to choose-one trials because of concerns about the equal access norm, then they should not object to choose-one trials in Study 3 because the character who chose to take one item could (in principle) go back to the bowl to take more after the second character took three; after all, the bowl is has a limitless supply. Nor should they object to choose-three trials because the second character chooses to take just one item from the bowl even though she could have taken more and can also can also go back to the bowl later. If children objected to choose-one trials in Study 1 because of concerns about the "situational norm," then they might object to choose-one trials in Study 3 because the second character takes

more than the standard set by the first character and this may be perceived as greedy. Although, in principle, it is possible that they would also object to choose-three trials in Study 3 because the second character took less than the standard set by the first character, this seemed unlikely because choosing to disadvantaged oneself is not indicative of selfishness. If children objected to choose-one trials in Study 1 because of a violation of the equal outcome norm, then they should object to both choose-one and choose-three trials in Study 3 because in both cases, the number of resources the two characters end up with differs.

Method

Participants. Sixteen 6-year-olds (*M* = 77.28 months, *SD* = 3.60 months; range = 73 months to 83 months; 8 girls), sixteen 8-year-olds (*M* = 101.40 months, *SD* = 3.42 months; range = 96 months to 105 months; 8 girls), and sixteen undergraduates (*M* = 19 years, *SD* = 1 year; range = 17 to 20 years; 12 women**)** participated in a single 15-minute session. Two additional participants (one 6-year-old; one adult) participated but were excluded for failing the catch trials. Twenty-four children (10 six-year-olds; 14 eight-year-olds) were recruited from the same database of families as Studies 1 and 2. The remaining children (6 six-year-olds; 2 eight-year-olds) were recruited from a nearby science museum. Undergraduate participants participated in exchange for course credit.

Procedure. The procedure for Study 3 was similar to Study 1 with five exceptions. First, there was a bowl rather than a plate with four items (see Figure 6 for the final frame from a test trial). Second, half of the participants saw the four choose-one trials first whereas the other half saw the four choose-three trials first. Trial type blocks were counterbalanced in the event that choose-one and choose-three trials influenced each other. (However, as will be described below, no order effects were found). Third, I chose not to include the relative emotion judgment

measure from Study 1 because it followed the same pattern as the "not okay" judgments and did not offer any additional insight into children's evaluations of the two trial types. Fourth, the catch trials were changed so that in the first catch trial both characters took one item and in the second catch trial both took three items. Fifth, at the end of the session, I asked participants to justify their response to both the first choose-three and choose-one trials (i.e. the first and fifth trial).

Figure 6. Example of image depicting final allocation in Study 3.

As noted earlier, one 6-year-old and one adult were excluded from the final analyses because they said that one of the equal outcome catch trials was "not okay." (The pattern of results reported here was the same regardless of whether they were included.)

Results

Preliminary analyses revealed no main effect of or interactions with gender or block order on any of the measures; therefore, gender and block order will not be considered in the analyses below.

"Not okay" judgments. Figure 7 shows the average number of times that participants at each of the three ages said the final distribution of items was "not okay" as a function of trial type. As the figure shows, on choose-three trials, where the first character took three items from the bowl and the other character took one, an age difference emerged. Six- and 8-year-olds were

ambivalent about whether the inequality was "okay" or "not okay;" the number of times they evaluated it as "not okay" did not differ from chance, *t*s(15) < 1.32, *p*s > .206, *d*s < .33, 95% CIs [-.38, 1.63] and [-.69, 1.32], respectively. However, adults reported that the inequality was "not okay" fewer times than chance (i.e., they said it was "okay" more than chance), $t(15) = 2.64$, $p = .019$, $d = .66$, 95% CI [-1.92, -.20].

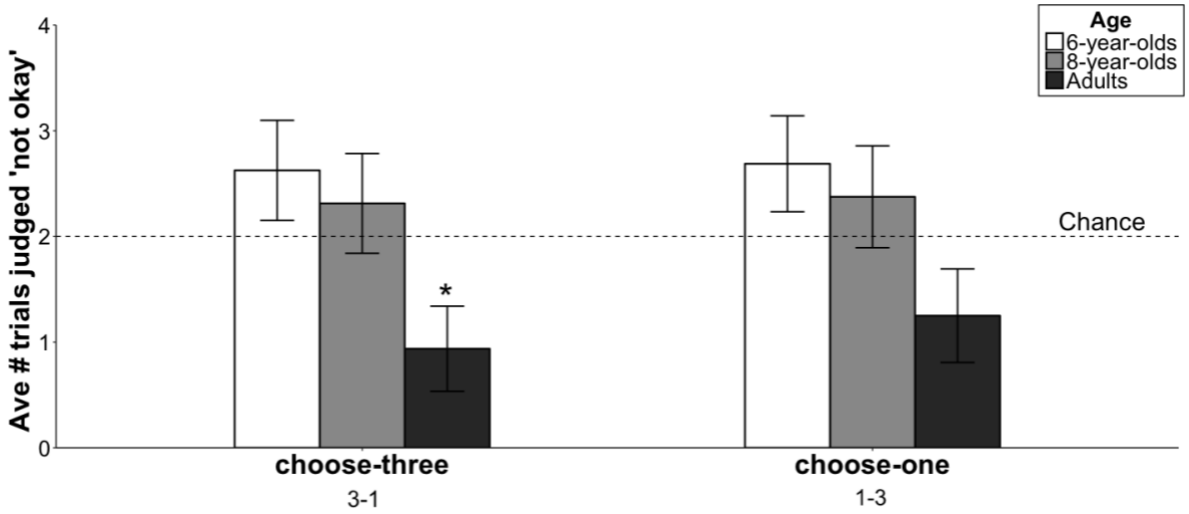

Figure 7. Average number of trials on which participants evaluated unequal outcomes as "not okay" in Study 3. Dashed line indicates chance level. Error bars show SEM. * $p < .05$

On choose-one trials, when the first character took one item from the bowl and the other character took three, 6-year-olds, 8-year-olds, and adults were ambivalent about whether the resulting inequality was "okay" or "not okay;" the number of times they evaluated it as "not okay" did not differ from chance, *t*s(15) < 1.70, *p*s > .111, *d*s < .42, 95% CIs [-.28, 1.65], [-.65, 1.40], and [-1.69, .19], respectively.

A 3 × 2 (Age × Trial Type) mixed ANOVA on the data in Figure 7 yielded only a main effect of age, $F(2, 45) = 4.14$, $p = .022$, $\eta_p^2 = .16$. There was no main effect of trial type, $F(1, 45)$

= .38, p = .541, η_p^2 = .01, and no Age × Trial Type interaction, $F(2, 45)$ = .124, p = .884, η_p^2 = .01. Bonferroni post-hoc comparisons showed that 6-year-olds were more likely than adults to say unequal outcomes were "not okay," p = .028, 95% CI [.13, 2.99]. Eight-year-olds did not differ from 6-year-olds, p = 1.000, 95% CI [-1.74, 1.12], or from adults, p = .105, 95% CI [-.18, 2.68].

Analyses at the individual level also showed that an age difference existed in terms of the number of trials judged to be "not okay:" Most children objected to inequality on both types of trials, whereas most adults did not. Table 4 shows the number of participants who judged the choose-three and choose-one trials to be "not okay" on 0-1, 2, or 3-4 trials as a function of age. On both choose-three and choose-one trials, the distribution of 6-year-olds differed either significantly or marginally from adults, Fisher Exact Tests, p = .025 and p =.066, respectively. The distribution of 8-year-olds on choose-three and choose-one trials did not differ from 6-year-olds, ps > .716, and was marginally different from adults, p = .066 and p = .081, respectively.

Table 4

Distribution of Participants' "Not Okay" Evaluations in Study 3

	Choose-Three Trials			Choose-One Trials		
	0-1 trials	2 trials	3-4 trials	0-1 trials	2 trials	3-4 trials
6-year-olds	5	1	10	5	1	10
8-year-olds	7	-	9	6	-	10
Adults	12	1	3	10	2	4

Severity ratings. To remain consistent with Study 1, I only examined the severity data from children who evaluated both types of trials as "not okay" on at least two or more trials. Figure 8 shows the average severity ratings as a function of trial type and age for subjects who judged both choose-three and choose-one trials to be "not okay" at least two times. As the figure shows, both trial types were rated as equally bad among both 6- and 8-year-olds. A 2 × 2 mixed

ANOVA (Age × Trial Type) on the data in Figure 8 confirmed that choose-one and choose-three trials were rated as equally bad and there was no difference in the ratings of 6- and 8-year-olds, *Fs* < 2.60, *ps* > .126, ηs_p^2 < .90.

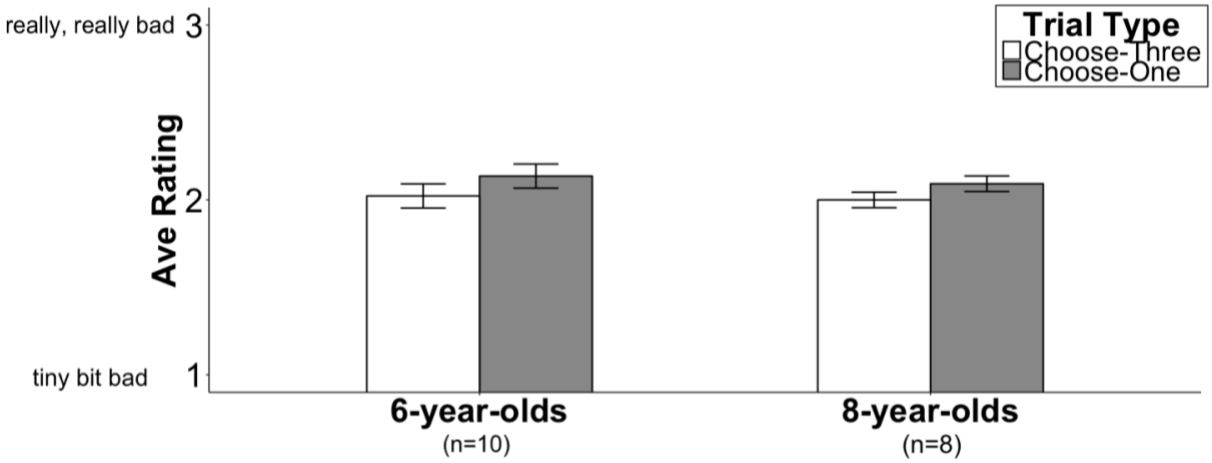

Figure 8. Average severity ratings on choose-three and choose-one trials among participants who rated choose-one and choose-three trials as "not okay" on at least two out of the four trials in Study 3. Error bars show SEM.

Justifications. After the catch trials, participants watched their first choose-three and first choose-one trial again (in the order that they saw them), were reminded of their evaluation, and asked to justify that judgment.[6] Figures 9A and 9B show the number of participants who said the first choose-one trial was "okay" or "not okay," as a function of age. As the figure shows, the pattern is the same as the data from the average number of choose-three and choose-one trials that were evaluated as "not okay" in Figure 7: Most 6- and 8-year-olds evaluated the first choose-one and choose-three trials as "not okay" while fewer than half the adults did so.

[6] Justification data is missing from one 8-year-old's "okay" choose-three evaluation due to noise at the museum, making the explanation inaudible. I included the evaluation in the "other" category.

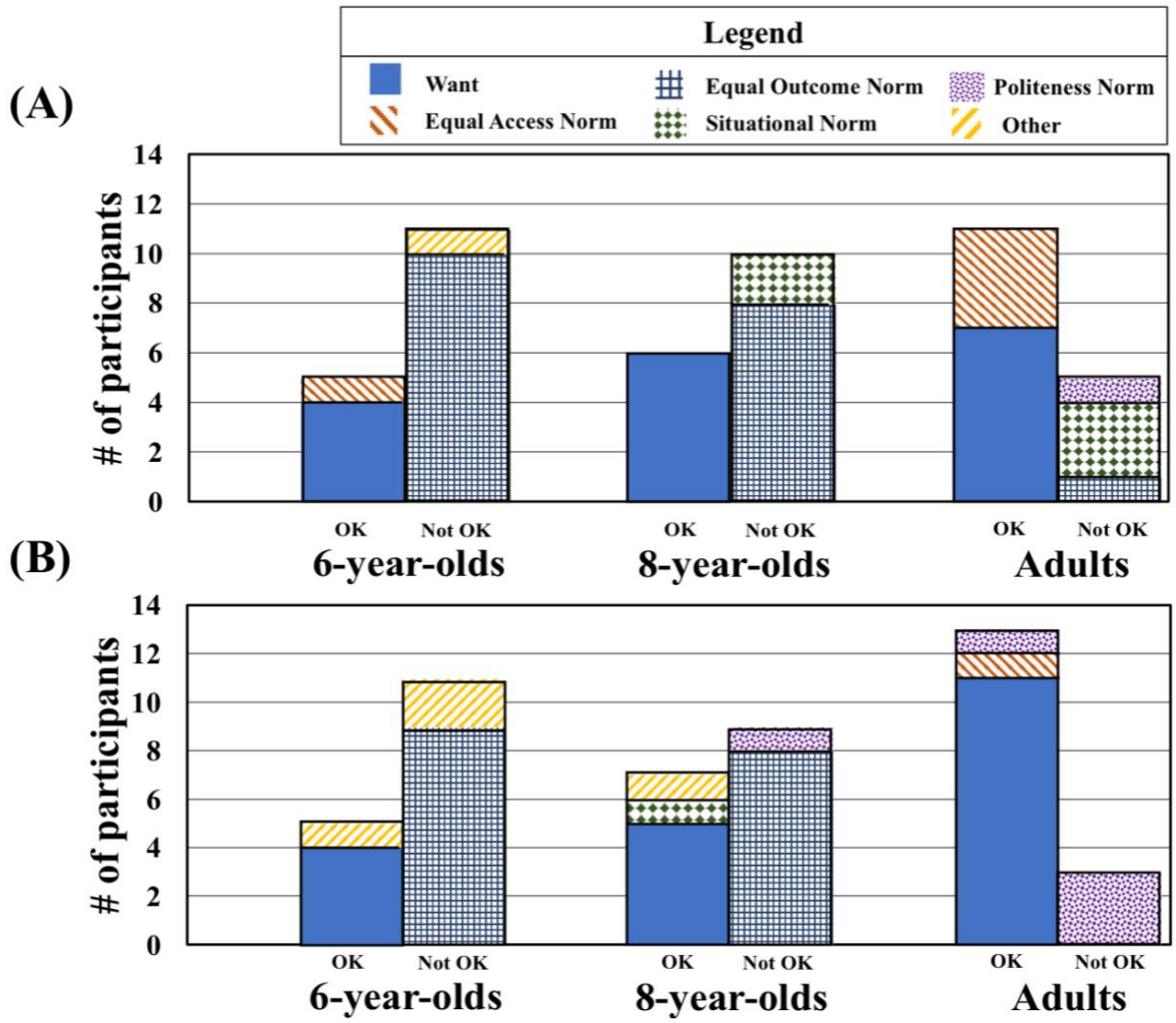

Figure 9. Participants' justifications in Study 3 as a function of their A) first choose-one evaluation trial and B) first choose-three evaluation trial.

Justifications fell into one of six categories: a) referencing that the characters got what they *wanted*; b) explaining that the two characters still have *equal access* to the resources and can always get more; c) referencing an *equal outcome norm* violation having to do with fairness or equality; d) referring to the *situational norm* in which the characters took one/three items

which was either less or more than the person who took first; e) stating a *politeness norm* that dictates each person ought to just take one item at a time; f) other comments. Table 5 shows examples of each category. A second coder and I independently coded each justification. Agreement was high (agreement = 93%, Cohen's kappa = .89, $p < .0005$). Disagreements were resolved through discussion.

Table 5
Study 2 First Choose-One and "Choose-Three" Evaluation Justification Categories and Examples

Category	Age	Trial Type	Evaluation	Examples
Want	6-year-old	Choose-One	Okay	Because he wants three and he wants one. Because they both want that many.
	8-year-old	Choose-One	Okay	They chose what they wanted.
Equal Access Norm	6-year-old	Choose-One	Okay	Because they can always get more.
Equal Outcome Norm	6-year-old	Choose-One	Not Okay	Because he got three and he got one.
Situational Norm	6-year-old	Choose-One	Not Okay	Because the first boy only took one so he thought the other one might only take take one but he took three.
Politeness Norm	Adult	Choose-Three	Not Okay	I think it would be more fair for the first person to take only one at first, let the other person get a chance to take one, and then the first person can choose again.
Other	6-year-old	Choose-Three	Okay	I think [the boy with three's] sweet tooth might get sweeter and [the boy with one's] sweet tooth might get a little sweeter.

As Figure 9 shows, most 6- and 8-year-old who evaluated a choose-one or choose-three scenario as "okay" explained that it was "okay" because both characters got what they wanted. Interestingly, as shown by the orange diagonal stripes, several adults said that the unequal outcomes were "okay" because, given the characters were choosing their own resources from a limitless supply in the bowl, the equal access norm was not violated and the character with less still had access to the resources and could always go back for more.

Figure 9 also shows there were three primary reasons participants objected to the unequal outcomes. The most common reason was that the equal outcome norm was violated when one character got more (or less) than the other, just as in Study 1 and as the results from the "should" condition of Study 2 indicated. Second, a few 8-year-olds and adults who evaluated the choose-one trials as "not okay" raised concerns that the "situational norm" was violated when the second character took more than the first character (see the green checkered function). Finally, one 8-

year-old and a few adults referred to a "politeness norm" when objecting to choose-one or choose-three trials, stating that individuals should only take one item at a time (see the purple function).

Discussion

Study 3 was designed to disentangle which one of the three norm violations (equal outcome norm, equal access norm, situational norm) best explains why children in Study 1 objected to the choose-one trials when the first character chose to take less than the second character.

One possibility was that on choose-one trials, children in Study 1 objected to a violation of an equal access norm that occurred when the second character took more than his or her fair share of the items, thus depriving the first character of access to his or her fair share. If this were the case, children would have accepted both choose-one and choose-three trials in Study 3 because the character who took three items never took more than his/her fair share (i.e., more than half the items in the bowl). Results from Study 3 showed that children did not accept both types of trials in Study 3.

A second possibility was that on choose-one trials, children in Study 1 objected to a violation of a situational norm: The second character took more items than the first and was knowingly creating inequality that advantaged himself/herself. If this were the case, children in Study 3 would have continued to object to the choose-one trials, but I thought it unlikely they would object to the choose-three trials because on the latter the character who takes less is disadvantaging himself/herself. In fact, children evaluated the same number of choose-one and choose-three trials as "not okay" and judged them to be just as bad.

A third explanation could be that on choose-one trials, children in Study 1 objected because the equal outcome norm was violated and they think both characters should take the same amount regardless of how many they said they wanted. In this case, they would be expected to object to both trial types in Study 3. Results from Study 3 (and Study 2) suggest that this is indeed, the most likely explanation. Even though it was clear resources were limitless and both characters were choosing the number of items they wanted, around 60% of 6- and 8-year-olds still objected on most trials.

As in Study 1, a developmental shift was evident: Overall, 6-year-olds objected more often to the unequal outcomes of both types of trials than adults. At the same time, there was a hint of a possibility that having the two characters take from a bowl with apparently limitless resources may have had some effect on 6-year-olds' judgments. In Study 1, 6-year-olds were more likely than chance to object to both trial types; in Study 3, they were merely at chance. Interestingly, however, having the two characters take from a bowl with apparently limitless resources did not influence the responses of 8-year-olds: They responded at chance levels in both Studies 1 and 3.

Responses provided by a few of the 8-year-olds during the exploratory follow-up questions suggest that a few of them interpreted the bowl of limitless resources in a different manner than expected. Of the eight 8-year-olds who objected to inequality by choice on at least two trials, three brought up the politeness norm during the exploratory follow-up questions. That is, they explained that everyone should get one item before it is okay for an individual to take more. For example, one 8-year-old said, "I would only take one and then everyone else take one and then if there was more I would get another one." It was not that she was necessarily averse to inequality per se (i.e., a violation of the equal outcome norm), but she wanted to make sure the

convention—common to Western classrooms and taking candy on Halloween—was followed. Although this raises interesting questions about how many resources individuals ought to take in a group setting, this was not my primary interest. My primary interest was in explaining why children object to inequality when one person chooses to take less than their fair share and the other chooses to takes more. I continue to investigate this question in Study 4.

Study 4

In Study 4, I set out to again investigate whether children's objection to the choose-one trials in Study 1 was driven by the violation of the equal outcome norm that the two characters possessed different amounts at the end of each trial or the violation of the equal access norm that the second character took more than half the resources available. I returned to using a plate—rather than a bowl— to eliminate any interpretation that the resources were being shared by a group. In contrast to Study 1, however, there were six rather than four items on the plate. On choose-one trials, the first character expressed a desire for and took one item and the second character took three, leaving two remaining on the plate. On choose-three trials, the first character expressed a desire for and took three items and the second character took one, again leaving two remaining on the plate. Importantly, across both trial types, the character who took three items never took more than his or her fair share and both trial types reflected inequality by choice.[7] I did not include a sample of adults because adults in Study 1 thought that it was okay for the second character to take more than his/her fair share when the first character only wanted one.

[7] Another design option would have been to keep the number of items available at four. On choose-one trials, the first character could have taken one, the second two, leaving one on the plate, for example. I chose to use six items in Study 4 so that the actual number of items the first and second character took would match those in Study 1.

Method

Participants. Sixteen 6-year-olds (*M* = 76.50 months, *SD* = 3.15 months; range = 72 months to 83 months; 8 girls) and sixteen 8-year-olds (*M* = 103.16 months, *SD* = 3.70 months; range = 96 months to 107 months; 8 girls) participated in a single 15-min session. Nineteen children (six 6-year-olds; thirteen 8-year-olds) were recruited from the same database of families as the previous studies. Thirteen children (ten 6-year-olds; three 8-year-olds) were recruited from a nearby science museum. One additional 8-year-old participated but was excluded for failing the catch trials. Children were primarily white and from middle-class backgrounds.

Procedure. Everything was the same as Study 3 with three exceptions. First, there were six items on a plate rather than a bowl full of unlimited resources. Thus, on both choose-one and choose-three trials, the character who ended up with three items took exactly his or her fair share (see Figure 10). Second, on the two catch trials, the two characters each took three items (i.e., half). Third, I always presented the four choose-one trials before the four choose-three trials because my primary interest was in the choose-one trials (as those were the ones in Study 1 where I saw a developmental shift) and I did not see a difference between the choose-one and choose-three trials in Study 3.

Figure 10. Example of image depicting final allocation in Study 4.

As noted earlier, one additional 8-year-old was excluded from the final analyses for failing a catch trial. (The pattern of results reported here was the same regardless of whether this participant was included.)

Results

Preliminary analyses revealed no main effect of or interactions with gender; therefore I collapsed across these factors below.

"Not okay" judgments. Figure 11 shows the average number of times that 6- and 8-year olds said the final distribution of items was "not okay." As in Study 1, 6-year-olds evaluated both the choose-three and choose-one trials as "not okay" on more trials than expected by chance, *t*s(15) > 3.58, *p*s < .003, *d*s > .89, 95% CIs [.46, 1.80] and [.99, 2.01], respectively. Eight-year-olds were ambivalent about whether both trial types were okay or not, *t*s(15) < 1.17, *p*s > .262, *d*s < .29, 95% CIs [-1.59, .47] and [-1.04, .92], respectively.

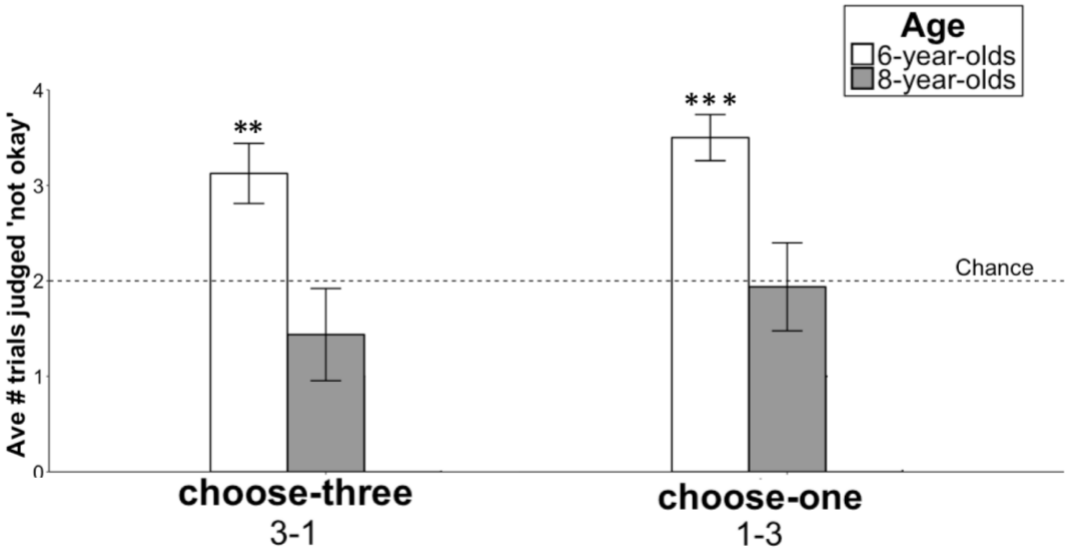

Figure 11. Average number of trials on which participants evaluated unequal outcomes as "not okay" in Study 3. Dashed line indicates chance level. Error bars show SEM. ** $p < .01$ *** $p < .0005$

A 2 × 2 mixed ANOVA (Age × Trial Type) on the data in Figure 11 yielded a main effect of age, $F(1, 30) = 10.16$, $p = .003$, $\eta_p^2 = .25$. Eight-year-olds evaluated fewer trials as "not okay" than 6-year-olds when the person who took more items did not take more than half. There was also a main effect of trial type, $F(1, 30) = 4.62$, $p = .040$, $\eta_p^2 = .13$, with more choose-one than choose-three trials evaluated as "not okay."[8] There was no Age × Trial Type interaction $F(1, 30) = .09$, $p = .761$, $\eta_p^2 < 01$.

At the individual level, the distribution of responses also varied by age. Table 6 shows the number of participants who judged the choose-three and choose-one trials to be "not okay" on 0-1, 2, or 3-4 trials as a function of age. Fisher exact tests showed that the distribution of 6- and 8-year-olds who evaluated the choose-one and choose-three trial types as not okay 0-1 versus 3-4 times differed across both choose-one and choose-three trials, *p*s < .035.

Table 6

Distribution of Participants' "Not Okay" Evaluations in Study 4

	Choose-one trials			Choose-three trials		
	0-1 trials	2 trials	3-4 trials	0-1 trials	2 trials	3-4 trials
6-year-olds	1	2	13	3	2	11
8-year-olds	7	1	8	10	-	6

Severity Ratings. Given that children in Study 3 who evaluated inequality by choice as "not okay" were not sensitive to the situational norm set by the first character, rating choose-three and choose-one trials as equally bad, I did not expect children's severity ratings to differ between the two trial types in Study 4. Although trial types were not counterbalanced, if children rated choose-one trials as worse than choose-three trials, it might provide some evidence that

[8] Because the choose-one trials came first, many children began by saying that these trials were "not okay," but by the choose-three trials some of them decided the inequality was "okay" since the person with more was only taking half. One six-year-old and three 8-year-olds switched their answers partway through the test trials from "not okay" to "okay."

they were sensitive to the situational norm set by the first character. I only examined the data from participants who evaluated both types of trials as "not okay" on at least two or more trials. Thirteen 6-year-olds and six 8-year-olds contributed data to this analysis.

Figure 12 shows the average severity ratings as a function of trial type and age for subjects who judged choose-three and choose-one trials to be "not okay" at least two times. As the figure shows, 6-year-olds rated the unequal outcomes as worse than 8-year-olds and both trial types as equally bad. A 2 × 2 mixed ANOVA (Age × Trial Type) confirmed that 6-year-olds evaluated the unequal outcomes as marginally worse than 8-year-olds, $F(1, 17) = 3.41$, $p = .082$, $\eta_p^2 = .17$. Additionally, there was no main effect of trial type, $F(1, 17) = 2.55$, $p = .129$, $\eta_p^2 = .13$, showing that choose-three and choose-one trials were rated as equally serious, and despite a suggestion of one in the data, no Age × Trial Type interaction, $F(1, 17) = 1.84$, $p = .193$, $\eta_p^2 = .10$.

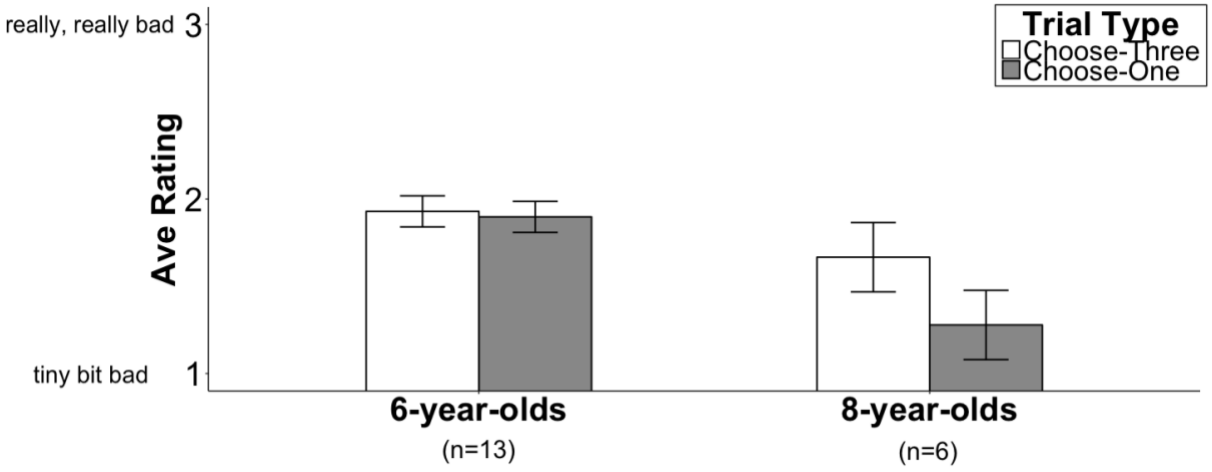

Figure 12. Average severity ratings on choose-three and choose-one trials among participants who rated choose-one and choose-three trials as "not okay" on at least two out of the four trials in Study 4. Error bars show SEM.

Justifications. After the catch trials, participants were reminded of whether they said the first choose-one test trial was "okay" or "not okay," and then were asked to explain that choice. As Figure 13 shows, the number of 6- and 8-year-olds who evaluated the first choose-one trial as "not okay" follows the same pattern as the data from the average number of choose-one trials that were evaluated as "not okay" in Figure 11: Most 6-year-olds and half the 8-year-olds evaluated the first choose-one trial as "not okay."

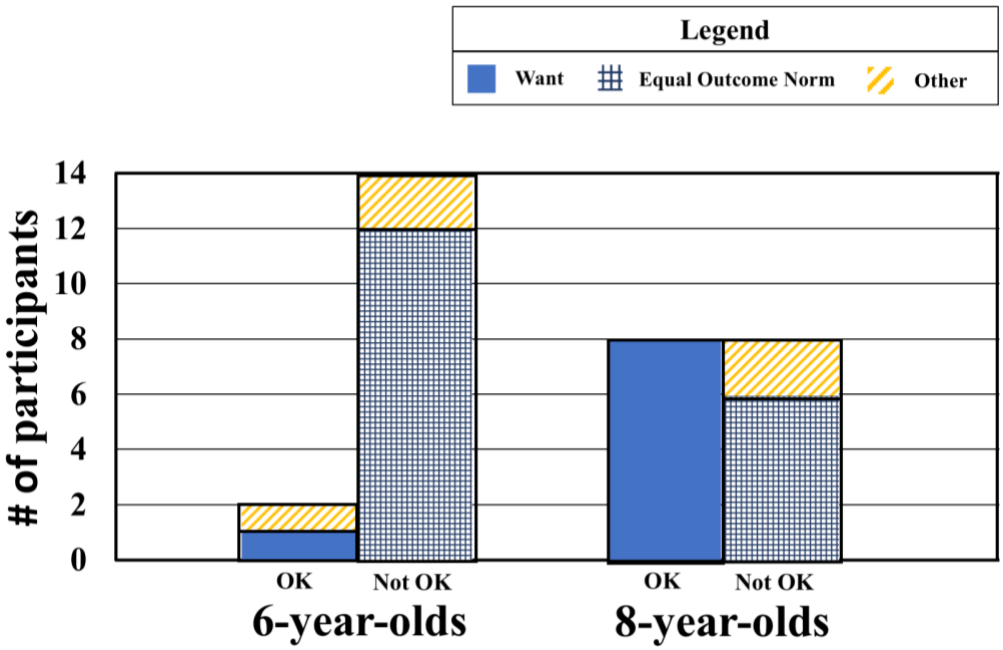

Figure 13. Participants' justifications in Study 4 as a function of their first choose-one evaluation trial.

The same coding scheme used in Study 1 was applied to the justifications in Study 4. A second coder and I independently coded the justifications. Agreement was perfect (agreement = 100%). As Figure 13 shows, "okay" justifications again almost always referenced the fact that the character who ended up with one had only wanted one. "Not okay" justifications most often referenced an equal outcome norm violation that the two characters did not have the same

number of items. Importantly, no participant in Study 4 justified their "not okay" response by referencing a violation of the equal access norm or permission norm, which had been raised by some 8-year-olds in Study 1 when the second character took more than half the resources.

Study 1 versus Study 4 Comparison. In contrast to Study 1 where 6- and 8-year-olds evaluated the same number of choose-one trials as "not okay," 6-year-olds in Study 4 evaluated the choose-one and choose-three trials (which here both reflect inequality by choice) as "not okay" more often than 8-year-olds. In order to investigate whether the manipulation in Study 4 had an impact on 6- and 8-year-olds' evaluations, I compared the number of "not okay" evaluations on choose-one and choose-three trials across Study 1 and Study 4.

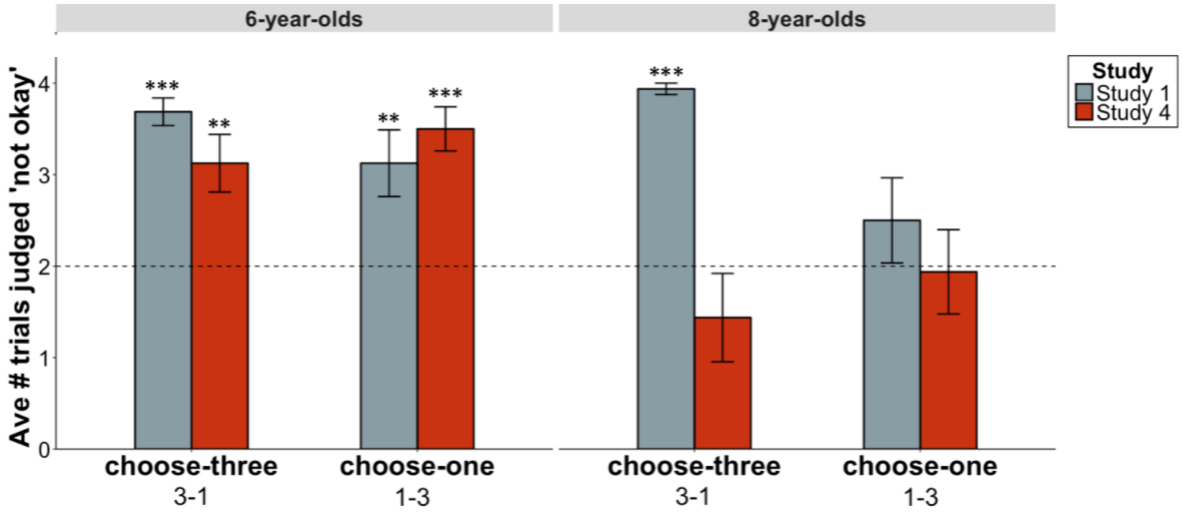

Figure 14. Average number of trials on which 6- and 8-year-olds evaluated unequal outcomes as "not okay" in Study 1 and Study 4. Dashed line indicates chance level. Error bars show SEM. ** $p < .01$ *** $p < .0005$

Figure 14 shows 6- and 8-year-olds' choose-three and choose-one evaluations in Study 1 and Study 4. As the figure shows, 6-year-olds evaluated both choose-three and choose-one trials as "not okay" at above chance levels across both Study 1 and Study 4. In contrast, 8-year-olds

only evaluated the choose-three trials of Study 1, when the first character took three, leaving only one item for the second character, as "not okay" at above chance level. All the other trial types, which reflect inequality by choice, were evaluated as "not okay" at chance levels.

A 2 × 2 × 2 mixed ANOVA (Age × Trial Type × Study) on the number of trials evaluated as "not okay" was conducted. There were two significant main effects. First, there was a main effect of study, $F(1, 60) = 7.20$, $p = .009$, $\eta_p^2 = .11$, showing that children evaluated more trials as "not okay" in Study 1 ($M = 3.31$, $SD = 1.32$) compared to Study 4 ($M = 2.5$, $SD = 1.74$). Second, there was a main of age, $F(1, 60) = 8.96$, $p = .004$, $\eta_p^2 = .13$, showing that 6-year-olds evaluated more trials as "not okay" overall than 8-year-olds (($M = 3.36$, $SD = 1.12$) vs. $M = 2.45$, $SD = 1.85$).

There were also two significant interactions. First, there was a Study × Age interaction, $F(1, 60) = 5.64$, $p = .021$, $\eta_p^2 = .09$. Simple effects tests with Bonferroni correction showed that whereas 6-year-olds evaluated the same number of trials as "not okay" in Study 1 and 4, $p = .827$, 95% CI [-.76, .95], 8-year-olds evaluated more trials as "not okay" in Study 1 compared to Study 4, $p = .001$, 95% CI [.68, 2.39]. Second, there was a Trial Type × Study interaction, $F(1, 60) = 16.53$, $p < .0005$, $\eta_p^2 = .22$. Simple effects tests with Bonferroni correction showed whereas more choose-three than choose-one trials were evaluated as "not okay" in Study 1, $p < .0005$, 95% CI [.51, 1.50], there were marginally fewer choose-three than choose-one trials evaluated as "not okay" in Study 4, $p = .082$, 95% CI [-.93, .058].

Discussion

Study 4 was conducted to investigate whether 6- and 8-year-olds objected to the choose-one trials in Study 1 because the equal outcome norm was violated or because the equal access norm was violated. The characters who took more resources in Study 4 did not violate the equal

access norm because they only took half the resources. If children continued to object to the choose-one trials in Study 4 as often as in Study 1, it indicates that their objection to the choose-one trials in Study 1 can be explained by the violation of the equal outcome norm.

Indeed, results from Study 4 showed that 6- and 8-year-olds' objection to the choose-one trials in Study 1 occurred because they violated the "equal outcome norm." Six-year-olds evaluated as many trials as "not okay" in Study 1 when the character with more resources took more than his or her fair share as they did in Study 4 when the character with more only took his or her fair share. Although 8-year-olds evaluated fewer trials than 6-year-olds as "not okay" in Study 4 when the character with more resources only took half, 8-year-olds objected at chance levels to both the choose-one trials in Study 1 when the second character took more than half the resources and the choose-one trials in Study 4 when the character only took half the resources. Therefore, results from Study 4 showed that seeing the character with more items only take half did not reduce the likelihood that 8-year-olds evaluated inequality by choice as "not okay."

However, the results from Study 4 provide two additional pieces of evidence that support an age difference between 6- and 8-year-olds' evaluations of inequality by choice. Recall that results from Study 1 were mixed regarding whether there was an age difference between 6- and 8-year-olds' judgments about inequality by choice. Although 8-year-olds distinguished between the choose-one and choose-three trials on all three measures ("not okay" judgments, severity ratings, relative emotion judgments) and 6-year-olds did not, 6- and 8-year-olds evaluated the same number of choose-one trials as "not okay."

The first piece of evidence from Study 4 that an age difference exists between 6- and 8-year-olds comes from children's "not okay" judgments. In contrast to Study 1 where there was no difference between the number of times 6- and 8-year-olds evaluated the choose-one trials as

"not okay," 6-year-olds evaluated more choose-one and choose-three trials in Study 4, both of which reflect inequality by choice, as "not okay" than 8-year-olds. Thus, when the character who takes more only takes half the resources, 8-year-olds are more comfortable with inequality by choice than 6-year-olds. The second piece of evidence comes from the analysis comparing Study 1 and Study 4. Whereas 8-year-olds evaluated fewer trials as "not okay" in Study 4 when both trial types reflected inequality by choice compared to Study 1 when choose-three trials did not reflect inequality by choice, 6-year-olds evaluated the same number of trials in Study 1 and Study 4 as "not okay." Thus, 6-year-olds evaluated inequality that occurs because the first person wants less as "not okay" as often as inequality that occurs because the first person is greedy.

Study 5

Accepting unequal outcomes that occur as a result of the choice of someone who takes less requires overriding the equal outcome norm. This may be especially difficult for children to do when the inequality is visible, as it has been in the previous studies. That is, in Studies 1, 3, and 4, one character was pictured with one item and the other character was pictured with three items. Children were asked to decide whether the final outcome was fair as they looked at the display of this inequality. There is some reason to expect that children might be better able to inhibit their aversion to inequality when the appetitive items are not in view.

In Carlson, Davis, and Leach (2005), for example, 3-year-olds were less successful at a reverse contingency task when they performed the task using actual treats than when they used representations of treats (see also Mischel and Ebbesen, 1970, showing that preschoolers wait longer when the treats are not visible). In this task, 3-year-olds needed to point to a smaller number of treats in order to receive a larger number of treats. One group of children saw the actual treats whereas another group of children were shown a picture of a mouse and elephant

meant to represent a small and large number of treats, respectively. While 3-year-olds who saw the actual treats pointed to the larger amount (the suboptimal response) around 45% of the time, 3-year-olds who indicated their choice using the pictures pointed to the mouse (the optimal response) 70% of the time. Thus, children who were symbolically distanced from the appetitive items were better able to override their desire to point to the larger number of treats, which was the amount they desired.

In one condition of Study 5 (the "hidden" condition), I covered the resources to cognitively distance 6-year-olds from the appetitive items. Just as 3-year-olds in Carlson and colleagues were better at inhibiting their prepotent response in the reverse contingency task when the treats were not visible, it is possible that 6-year-olds might be better able to override the equal outcome norm when the appetitive items are not in view.

In a second condition of Study 5 (the "permission+generosity" condition), I investigated whether 6-year-olds might also be more likely to accept inequality in the choose-one trials if the character who ultimately ended up with less both gave permission to the second character to take the remainder and also actively gave the remaining items to him or her. After the first character expressed a desire for just one (of the four) items and took one item, s/he gave the second character permission to have the rest (e.g., "You can have the rest.") and pushed them toward him or her. Thus, the second character both received permission from the first character to end up with more resources and was merely a passive recipient of the first character's generosity. As discussed in the Introduction, 6-year-olds have been found to object less to generous offers compared to selfish ones (McAuliffe et al., 2015). By incorporating two additional cues that highlight the first character's choice to take less, I believed it might help 6-year-olds recognize that the choose-one trials were permissible.

Finally, in a "standard" condition, I also included a third group of children who watched the same scenarios as those presented in Study 1. This condition served as a replication of Study 1 and also as a comparison group against which I could compare the choose-one trial data from the children in the hidden and permission+generosity conditions.

Method

Participants. Forty-eight 6-year-olds (*M* = 77.97 months, *SD* = 3.57 months; range = 72 months to 84 months; 24 girls) participated in a single 15-min session. Thirty children were recruited from the same database of families as the previous studies. Eighteen children were recruited from a nearby science museum. One child participated but was excluded for failing one of the catch trials. Children were primarily white and from middle-class backgrounds.

Procedure. Sixteen participants took part in each of the three conditions: standard, hidden, and permission+generosity, balanced at each age for sex. The average age of children in each condition was, respectively, 6;5, 6;5, and 6;6. The number of children from the museum who participated in each condition was, respectively 6, 7, and 5.

I returned to using a plate with four items (as in Study 1) rather than with six because the comparison in Study 4 revealed that 6-year-olds evaluated the same number of choose-one and choose-three trials as "not okay" regardless of whether there were four or six items on the plate. Thus, choose-one trials again reflected inequality by choice whereas choose-three trials did not. Choose-one trials were always presented before choose-three trials. Additionally, I only included the judgment question. I decided not to include the severity rating question because 6-year-olds evaluated both trial types as equally bad in the other three studies that included this measure.

In the standard condition, participants watched the same animations presented in Study 1. In the hidden condition, participants watched the characters take their items and drop them into brown paper bags so they were not visible as children decided whether the outcome was "okay" or "not okay." As in all the previous studies, the experimenter simply asked children if it was "okay" or "not okay" that one character got one and the other character got three. Figure 15 shows an example of a final frame from the first choose-one hidden test trial.

Figure 15. Example of an image depicting the final allocation from the hidden condition in Study 5.

In the permission+generosity condition, on the choose-one trials, the first character expressed a desire for and took only one item, and then said (to the second character), "You can have the rest," as s/he handed the remaining three items to the second character. The second character responded by saying, "Hmmm" (as in all the previous studies). On choose-three trials, the first character expressed a desire for and took three items, and then said, "You can have the rest," as s/he handed the single remaining item to the second character (who responded by saying "Hmmmm"). In the permission+generosity condition, the items were visible as in the standard condition and as in each of the previous studies.

An example of a choose-one test trial from both the hidden and permission+generosity conditions is available at https://osf.io/wpbmj/?view_only=27c91fa131964e958c7922a93bb8ca66.

As noted earlier, one additional 8-year-old was excluded from the final analyses for failing a catch trial. (The pattern of results reported here was the same regardless of whether this participant was included.)

Results

Preliminary analyses revealed no main effect of or interactions with gender; therefore, I collapsed across these factors below.

"Not okay" judgments. Figure 16 shows the average number of times that 6-year-olds said the final distribution of items was "not okay" as a function of trial type across the three conditions. Consistent with the data from previous trials, 6-year-olds in the standard, hidden, and permission+generosity conditions evaluated both the choose-three and choose-one trials as "not okay" on more trials than expected by chance, $ts(15) > 2.97$, $ps < .010$, $ds > .74$.

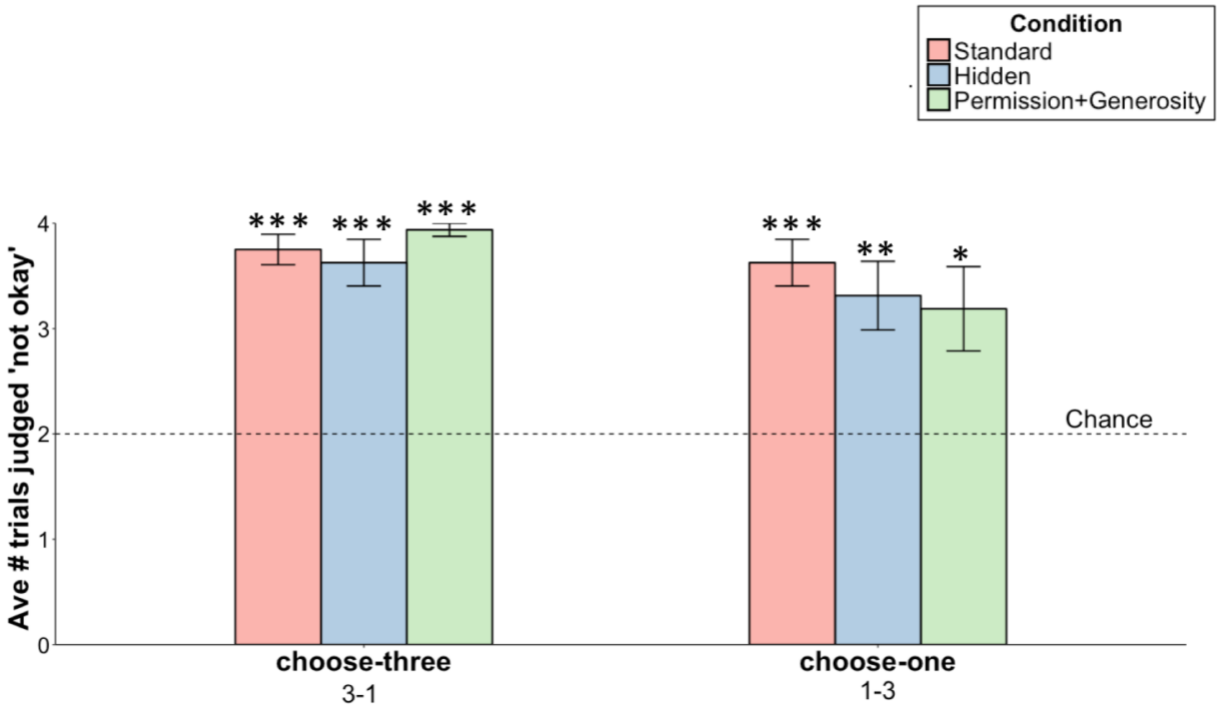

Figure 16. Average number of trials on which participants evaluated unequal outcomes as "not okay" in Study 3. Dashed line indicates chance level. Error bars show SEM. * $p < .05$ ** $p < .01$ *** $p < .0005$

A 3 × 2 (Condition × Trial Type) mixed ANOVA on the data in Figure 16 yielded a main effect of trial type, $F(1, 45) = 5.82$, $p = .020$, $\eta_p^2 = .01$, but no main effect of condition, $F(2, 45) = .270$, $p = .765$, $\eta_p^2 = .01$, and no Trial Type × Condition interaction, $F(2, 45) = 1.27$, $p = .290$, $\eta_p^2 = .05$. Overall, children evaluated more choose-three than choose-one trials as "not okay."

Analyses at the individual level yielded a similar pattern of results. Few children in any of the conditions distinguished between the choose-one and choose-three trials. Table 7 shows the number of participants in each condition who judged more choose-three than choose-one trials to be "not okay," the reverse, and the number judging the same number of each to be "not okay." In the standard and hidden conditions, there was no difference in the number of children

who judged more choose-three than choose-one trials as "not okay" than the reverse (Wilcoxon signed-rank: $Z = .82$, $p = .41$, $Z = 1.07$, $p = .29$, respectively); most judged the same number of each (i.e., four) to be "not okay." In contrast, marginally more children in the permission+generosity condition judged more choose-three than choose-one trials to be "not okay" than the reverse (Wilcoxon signed-rank: $Z = 1.84$, $p = .066$).

Table 7

Number of Participants in Each Condition in Study 5 Judging More Choose-Three or Choose-One Trials as "Not Okay," or the Number who Judged the Same Number of Each Trial Type as "Not Okay"

	More Choose-Three	More Choose-One	Same
Standard	2	1	13
Hidden	2	1	13
Generosity	4	0	12

Very few children in any of the conditions evaluated the choose-one trials as "okay." Table 8 shows the number of participants in each condition who said the choose-one trials were "not okay" 0-1 times, 2 times, or on 3-4 trials as a function of condition. Fisher exact tests showed that the distribution of children who evaluated the choose-one trials as "not okay" 0-1 versus 3-4 times did not differ in any of the three conditions, ps > .598. Thus, hiding the final distribution before the test question and adding additional cues highlighting the first character's choice to take less did not influence 6-year-olds' judgments about inequality that occurred by choice. Children in the hidden and permission+generosity evaluated just as many choose-one trials as "not okay" as they did in the standard condition.

Table 8

Distribution of Participants' Choose-One "Not Okay" Evaluations in Study 5

	0-1 trials	2 trials	3-4 trials
Standard	1	1	14
Hidden	2	2	12
Generosity	3	0	13

Justifications. After the catch trials, participants were reminded of whether they said the first choose-one test trial was "okay" or "not okay," and then were asked to explain that choice. As Figure 17 shows, almost all of the 6-year-olds in each condition evaluated the first choose-one trial as "not okay."

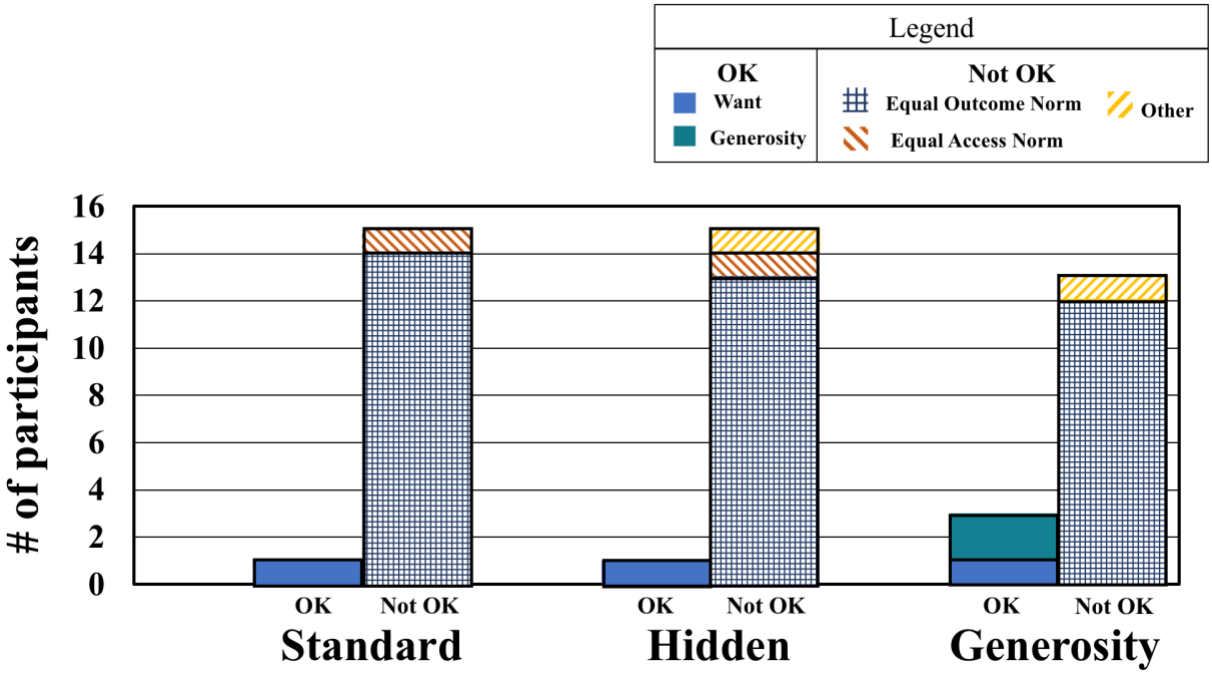

Figure 17. Participants' justifications in Study 5 as a function of their first choose-one evaluation trial.

The same coding scheme used Study 1 was applied to the justifications in Study 5 with one exception. There was one new category in the "okay" justifications, which was stating that the first character was *generous* and gave the items away (e.g., "He gave them to him."). A second coder and I independently coded the justifications. Agreement was nearly perfect (agreement = 98%, Cohen's kappa = .94, $p < .0005$). The one disagreement was resolved through discussion.

Figure 17 shows that, as in the previous studies, children who evaluated inequality by choice as "okay" justified their response by stating that the character only wanted one item. In the permission+generosity condition, two children who evaluated the unequal outcome as okay referenced the fact that the first character gave the items away. "Not okay" justifications most often referenced a violation of the equal outcome norm.

Discussion

Study 5 had two primary aims. The first aim was to investigate whether 6-year-olds might be better able to override the equal outcome norm when the appetitive items were hidden from view as they made their judgments. Results from the hidden condition showed this had no effect: Six-year-olds judged the choose-one trials to be "not okay" at above chance and judged the same number of choose-one trials as "not okay" as children who made their judgments when the items were visible. Additionally, 6-year-olds who evaluated the outcomes while the items were covered did not distinguish between the choose-three and choose-one trials, with most judging both to be "not okay" on all four trials. These results suggest that seeing the unequal distributions in Studies 1, 3, and 4 was not a major factor in why 6-year-olds judged them to be "not okay."

Second, I investigated whether 6-year-olds would be more likely to judge an unequal outcome as acceptable if the person who ended up with just one item gave permission for the other character to have more and actually gave the items to him/her. This manipulation also had no effect. Six-year-olds still evaluated the choose-one trials as "not okay" at above chance and there was no difference between the number of times children in the permission+generosity conditions evaluated the choose-one trials as "not okay" and the number of times children in the

standard condition did so. In line with previous research that children object to generous offers less than selfish ones (McAuliffe et al., 2015), there was a hint that children evaluated the unequal outcomes as "not okay" on more choose-three trials when the first character was greedy than on choose-one trials when the first character took less than his or her fair share. That said, 81% of 6-year-olds who saw the first character act generously still judged the choose-one trials as "not okay" on 3 or 4 out of 4 trials.

In sum, children's responses to the choose-one trials in Study 5 did not differ across the three conditions. These results provide further evidence for a robust aversion towards inequality by choice in early childhood and demonstrate the difficulty that 6-year-olds have overriding the equal outcome norm.

General Discussion

Young children subscribe to an equal outcome norm that individuals should receive and possess the same number of resources. Over the course of development, they learn that some types of inequality are acceptable in their culture. While a significant amount of research has examined the developmental timeline of children's responses to other types of acceptable inequality, such as merit and need, this dissertation, to my knowledge, represents the first set of studies to investigate whether children judge inequality that occurs by choice as acceptable.

Summary of Results

Over five studies, I found evidence for a developmental shift in response to inequality by choice. Six-year-olds objected to inequality by choice, 8-year-olds were ambivalent, and adults thought this type of inequality was permissible.

Study 1 found that 6-year-olds do not distinguish between scenarios in which inequality occurs because one person chooses to take less than half the resources (leaving more for the other person) with scenarios in which inequality occurs because one person took more resources (leaving less for the second person). They objected as often to these two types of inequality, rated the violation as equally serious, and judged the character who ended up with three items as happier as often. In contrast, 8-year-olds and adults distinguished between these two types of inequality: Eight-year-olds and adults judged more choose-three than choose-one trials to be "not okay," and judged the character with three items to be happier than the character with one item on more choose-three than choose-one trials. Additionally, when 8-year-olds evaluated the choose-one trials to be "not okay," there was a hint that they evaluated the choose-three trials as worse than the choose-one trials.

In Study 2, I investigated whether 6-year-olds' failure to distinguish between the two types of inequality was due to an inability to understand that two people can desire different amounts of a particular resource, or whether they think characters should take their fair share even if they do not want it. I found that 6-year-olds understand that two people can desire different amounts of a resource but think that individuals ought to take the same number of items regardless of their desire. This differed from adults who thought that individuals should take less if that is all they want.

Study 3 investigated which of three norms—equal access norm, situational norm, or equal outcome norm —explained children's objections to the inequality by choice trials in Study 1. Children's objections to the choose-one trials in Study 1 were not due to a situational norm: They objected as often when the second character took more from a limitless bowl of resources than the first character as when the second character took fewer. Although Study 3 also aimed to

tease apart whether children's objections to the choose-one trials in Study 1 were driven by the violation of the equal outcome norm or the equal access norm, seeing a bowl full of unlimited resources may have led some 8-year-olds to construe the situation as one requiring each individual to take just one item at a time.

Study 4 was conducted to investigate again whether children objected to the choose-one trials in Study 1 because they violated the equal outcome norm or the equal access norm; results suggested it was because they violated the equal outcome norm. Even when the character who took more resources only took half the items (allowing the other character access to his/her fair share), 6-year-olds objected as often in Study 4 as they had in Study 1, when the character who took more items took more than half. Eight-year-olds were ambivalent about whether the choose-one trials in both Study 1 and Study 4 were permissible. Thus, even when the equal access norm was not violated in Study 4, 8-year-olds' objections did not shift significantly. However, it is important to note that children's objections to the choose-one trials in Study 1 may have been influenced by the existence of multiple norm violations. Additionally, although 8-year-olds' objections seem to primarily be driven by the violation of the equal outcome norm, the fact that they were more comfortable than 6-year-olds with inequality by choice in Study 4 when the equal access norm was not violated but were as likely to object in Study 1 when the equal access norm was violated suggests that some 8-year-olds may have in fact been responding to the equal access norm violation in Study 1.

Study 5 investigated whether 6-year-olds would continue to object to inequality by choice if the amount of cognitive effort needed to override the equal outcome norm was reduced. Results showed that 6-year-olds were unable to override the equal outcome norm when the

appetitive items were covered as well as when the first character gave permission to take the rest and then actively gave the items to the second character.

Together, these studies demonstrate that different norms are prioritized across development in response to inequality by choice. Six-year-olds clearly prioritized the equal outcome norm, consistently evaluating inequality by choice as "not okay." Eight-year-olds were ambivalent: There was a lot of variability within the age group, representative of a transition stage. Like 6-year-olds, some 8-year-olds prioritized the equal outcome norm and consistently evaluated inequality by choice as "not okay." However, at least some 8-year-olds prioritized the equal access norm over the equal outcome norm because the results from Study 4 (when the equal access norm was not violated) showed that 8-year-olds were less likely than 6-year-olds to evaluate inequality by choice as "not okay." Finally, the results from Study 1 showed that some 8-year-olds responded to inequality by choice in the same manner as adults who were comfortable with it even when both the equal outcome and equal access norms were violated.

Young Children are Strict Norm Enforcers

Six-year-olds strictly enforced the equal outcome norm even in the presence of information that warranted overriding the norm. These findings are consistent with other research showing that 6-year-olds fail to integrate contextual information into their resource allocation evaluations, instead employing rule-based reasoning. In one study, for example, 6-year-olds inferred that a distributor who allocated an extra resource to one individual over another was better friends with the recipient even when the distributor acted impartially. In Liberman and Shaw (2017), 4- to 9-year-olds heard a vignette about a distributor who was splitting five erasers between two boys. After giving two erasers to each of the boys, the distributor allocated the final eraser to one boy in either one of two ways. In the partial condition,

children were told that the distributor chose which boy should receive the extra eraser. In the impartial condition, children were told the distributor rolled a die to decide which boy should receive the extra eraser. Afterwards, participants were asked whether the distributor was better friends with the recipient, the non-recipient, or both characters. Whereas 7- to 9-year-olds in the partial condition were more likely than 7- to 9-year-olds in the impartial condition to judge that the distributor was better friends with the recipient (86% versus 16%), 4- to 6-year-olds in the two conditions were as likely to judge that the distributor was better friends with the recipient (61% versus 59%). Thus, children 6 years of age and younger disregarded how the inequality was created and responded in accordance with the norm that giving a resource to one person over another reflects favoritism.

Similarly, 6-year-olds in the studies in this dissertation objected as often to inequality regardless of how it came about. They evaluated inequality that occurred when one person chose to take less as "not okay" as often as inequality that occurred when the person had no choice but to take less. They applied their equal outcome norm to inequality by choice and—despite understanding desire and free will in other contexts—did not integrate that information into their judgments.

Just as the 6-year-olds in Liberman and Shaw (2017) assumed that the distributor favored the recipient regardless of the procedure he used to allocate the item, an exploratory follow-up question at the end of Study 4 indicated that 6-year-olds assigned fault to the character who took more resources regardless of context. Recall that in Study 4 there were six items on the plate. At the end of the study, I asked 6-year-olds which character they thought was at fault (e.g., Who do

you think did the wrong thing?"). They nearly always assigned blame to the character who took three resources even when s/he went first, took only half the resources, and could not have known that the second character would *choose* to take just one.

With age, children begin to integrate contextual cues into their resource allocation judgments and inferences. The developmental literature that examines children's responses to inequality under acceptable circumstances (e.g., when an impartial procedure is used) has documented the developmental shift that occurs in children's responses but has not yet discovered the mechanism that causes the shift. In regards to the shift that occurs in response to inequality by choice, there are at least three reasons that older children might be better able at integrating the contextual cues into their judgments. First, older children have greater executive function than younger children and are likely to be better able to inhibit the equal outcome norm when weighting the contextual information (Zelazo et al., 2013). Although the hidden condition in Study 5 was intended to reduce the level of inhibition 6-year-olds needed to override the equal outcome norm, it is possible that this condition was not robust enough to attain the desired effect. Whereas Carlson et al. (2005) used "cool" representations (a mouse and an elephant) to represent the treats, the hidden condition merely covered the items and may not have succeeded at dampening the "hot" affective response to the appetitive treats (e.g., Metcalfe & Mischel, 1999). Second, older children have arguably encountered inequality by choice more often in their environment and have had more opportunity to see how authority figures and peers on both the advantaged and disadvantaged side have responded.

Third, between age six and eight, children become better at recognizing that thoughts and beliefs are constructed and differ across individuals. That is, children begin to understand that two individuals can leave the same event with different interpretations about what transpired,

especially in the case in which an event or stimulus is ambiguous. In Lalonde and Chandler (2002), for example, 6- and 7-year-olds were shown a drawing and were then asked how two individuals presented only with a restricted and ambiguous view would interpret the image. In order to demonstrate an understanding that knowledge is constructed (i.e. an "interpretive theory of mind"), children had to predict that two naïve puppets would not guess the actual picture (which children had knowledge of) and would interpret the ambiguous stimuli differently from each other. Thus, children had to inhibit their own knowledge while recognizing that two people can have different interpretations of the same event. While 7-year-olds demonstrated better interpretive theory of mind than 6-year-olds (88% versus 33%) research in this area has shown that 7- and 8-year-olds still show a lot of variability in performance. For example, Carpendale and Chandler (1996) found that 8-year-olds were accurate on a similar task only 67% of the time.

The cognitive abilities required for interpretive understanding share some similarity to those needed to recognize inequality by choice as permissible. On the surface, it is ambiguous whether inequality by choice is permissible. The character with less is choosing to take less of his or her own free will, but the inequality that is created is still a violation of the firmly entrenched equal outcome norm. In order to evaluate inequality by choice as permissible, children need to inhibit the equal outcome norm while also recognizing that two people can come to the same resource distribution event with different interpretations of the items they wish to take. The fact that interpretive understanding seems to require some of the same cognitive abilities as those required to recognize inequality by choice as permissible and the pattern of performance across age is similar for both types of tasks suggests that there might be a similar mechanism driving both abilities.

I turn now to considering the value of children's evaluations.

The Value of Evaluations over Behavior

Research on how agency and choice influences 6- and 8-year-olds' willingness to *act* prosocially and disadvantage themselves in a resource allocation context is consistent with the age difference that I found here in children's evaluations. In (Carpendale & Chandler, 1996), for example, 4- to 8-year-olds were told that an experimenter had three erasers to distribute between the participant and a hypothetical peer. After distributing one eraser to each of the children, the experimenter asked the participant whether s/he should give the final eraser to a peer or throw it away. Consistent with the findings here that 8-year-olds were ambivalent when evaluating inequality by choice, around half of 7- to 8-year-olds were willing to give the extra eraser to a peer. Interestingly, also in line with the findings in this dissertation, 75% of the 4- to 6-year-olds chose to throw the extra eraser away. However, when asked if they would be happier if a researcher gave the extra eraser to a peer or threw it away, around three-quarters of children in both age groups reported they would be happier if the extra item were thrown away. Whereas 7- to 8-year-olds felt more comfortable getting less than someone else when they made the decision to be generous than when it was foisted upon them, 4- to 6-year-olds were never comfortable with the inequality.

However, 8-year-olds' willingness to act prosocially does not address whether they think this type of inequality is fair and children's evaluations have value separate from their behavior. Moral theorists argue that third-party evaluations are the true test of moral understanding because they are not influenced by self-interested motives, including selfishness or self-presentation concerns (e.g., Vaish, Missana, & Tomasello, 2011). It is possible that the 8-year-olds in Shaw et al.'s (2016) study were trying to appear generous to the experimenter.

While the studies in this dissertation demonstrate that 6-year-olds consistently evaluate inequality by choice as unacceptable, it is likely that 6-year-olds would accept inequality by choice if they were to benefit. Children age 6 and younger are happy and accept advantageous offers where they get more resources than another person but are sad and object when they get disadvantageous offers where they get less than another person (Blake & McAuliffe, 2011; LoBue, Nishida, Chiong, DeLoache, & Haidt, 2011). If they heard someone else only wanted one of four treats, it is likely that they would happily take the three remaining pieces. By asking children to evaluate inequality by choice in a third-party context, however, I am measuring their moral judgments without the interference of self-interested motives, including younger children's selfishness and older children's self-presentation concerns.

Moreover, individuals' evaluations matter because their perceptions about how fair their environment is influences their well-being. For example, the extent to which American adults evaluate others as trustworthy or fair influences how happy they rate themselves to be (Oishi, Kesebir, & Diener, 2011). Thus, children's evaluations of inequality by choice—a type of inequality they see in their daily lives and is permitted by authority figures—might influence children's sense of trust in their environment and overall happiness. Indeed, it is not uncommon for a child to proclaim, "That's not fair!" in situations where we as adults do not perceive an infraction. Although adults may ignore children's responses when they seem unjustified, the findings from the studies here reveal that children might really perceive some behaviors as moral violations and be upset by situations that adults do not interpret in the same manner.

Limitations and Future Research

Below I review four limitations of the current work and potential directions for future research. The four limitations include: 1) the items used in the studies here were treats and it is

unknown whether these findings would generalize to other items; 2) Study 2 did not include a group of 8-year-olds so it is not clear whether 8-year-olds think someone who wants less should take less or should take their fair share; 3) inequality by choice was only assessed in a Western, Educated, Industrialized, Rich, and Democratic (WEIRD) culture; 4) evaluations may have been driven by the demand characteristics within the paradigm.

Item value. First, one limitation of this work is that the items the two characters were deciding how to split were valuable treats. Perhaps young children think individuals ought to take the same number of items regardless of choice because the items will have value later. One question is whether the findings here would generalize to other types of resources. It is possible that children might have evaluated inequality by choice as permissible if the two characters had been deciding how to split other types of items. For example, the characters could have decided how to split items that children consider aversive, such as vegetables. Although it is possible that children might have interpreted inequality by choice as permissible with aversive items, exploratory questions at the end of some of the studies here showed that 6-year-olds either treated these items as if they were valuable (e.g., broccoli is healthy and everyone should have the same amount) or as if they were a punishment that both individuals should have to endure equally.

Furthermore, other research has shown that children *only* accept unequal distributions following different levels of effort if the items are luxury (fun) items like treats and do not accept unequal distributions when the items are necessary items needed to remain healthy. In Rizzo et al. (2016), 6- to 8-year-olds heard a story about a character who worked hard to find a lot of "blickets" and a character who was lazy and did not work to find any. Children in the luxury resource group were told that blickets made the characters happy and were fun but were not

essential because the characters could play with other things that made them happy. Children in the necessary resource group were told that blickets keep the characters healthy and without them the characters would get sick. Afterwards, children were asked to distribute six blickets between the hardworking and lazy character. Whereas children in the luxury resource group distributed more blickets to the hardworker, children in the necessary resource group distributed an equal number of blickets to both characters. It is surprising that children consider unequal distributions of luxury items to be permissible in the case of effort but not in the case of choice. While future research might examine whether 6-year-olds' interpretation of inequality by choice is different when item value is manipulated, if children do not accept inequality by choice with luxury items, it is not clear that they would endorse it with other types of items.

***Should* evaluations.** A second limitation is that in Study 2 I only investigated the number of items that 6-year-olds thought the character who wanted one item *should* take and did not include a group of 8-year-olds. Consequently, I do not know whether the majority of 8-year-olds would have responded as 6-year-olds, stating that both characters should take two items regardless of desire, or if they would have responded as adults and said the character who wanted one item should only take one. In addition to exploring this question with 8-year-olds, future research could employ a within-subjects design to investigate whether children's evaluations about what ought to happen line up with their evaluations of inequality by choice.

Cross-cultural research. The results from the studies here demonstrate that adults in a WEIRD culture think inequality by choice is acceptable. However, it is important to note that fairness evaluations can differ across culture. For example, children who live on a kibbutz have been found to distribute items more equally between themselves and a peer than urban children even when the child or the peer accomplished different amounts in a task (Nisan, 1984).

Therefore, individuals in other cultures may either a) not consider inequality by choice as an acceptable reason to deviate from the norm of equality or; b) have a different developmental trajectory of when this type of inequality is accepted. Future research should examine whether individuals evaluate inequality by choice as permissible in other cultural contexts.

Demand characteristics. In the study here, 6-year-olds objected to inequality by choice between two individuals who appeared to be peers of the same age and gender and were presumably equally deserving. These factors may have served as demand characteristics that influenced children's judgments—that is, given that the characters were identical in nearly every way, young children may have thought the experimenter expected them to enforce the equal outcome norm. Manipulating social information about the two characters, such as the relative dominance of the two individuals, might have influenced 6-year-olds' willingness to accept inequality by choice. For example, recent research has found that 17-month-olds look longer at a scenario in which a submissive and dominant individual each receive the same number of resources than at a scenario in which the dominant individual receives more resources (Enright, Gweon, & Sommerville, 2017). This suggests that children less than two years of age expect dominant individuals to receive more resources than submissive ones. Therefore, 6-year-olds might have been more willing to accept inequality by choice if the individual who had expressed a desire for and took less was submissive or was considered of lower rank than the individual who chose to take more.

Conclusion

In sum, the studies reported here show a developmental shift between early childhood and adulthood in response to inequality by choice. Six-year-olds prioritized the equal outcome norm and consistently evaluated inequality by choice as "not okay," eight-year-olds were

ambivalent and were more likely than 6-year-olds to accept inequality by choice when the equal access norm was not violated, and adults thought inequality by choice was "okay" even when both the equal outcome and equal access norms were violated. The results from this research add to a growing body of literature charting the developmental timeline of when children evaluate different types of inequality as acceptable.

References

Arneson, R. (1993). Equality. In R. Goodin & R. Petit (Eds.), *A companion to contemporary political philosophy* (pp. 489–507). Oxford: Blackwell.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford University Press.

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*, 492–498. https://doi.org/10.1037/a0026598

Blake, P. R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, *120*, 215–224. https://doi.org/10.1016/j.cognition.2011.04.006

Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more: Executive function and symbolic representation in preschool children. *Psychological Science*, *16*, 609–616.

Carpendale, J. I., & Chandler, M. J. (1996). On the Distinction between False Belief Understanding and Subscribing to an Interpretive Theory of Mind, *67*, 1686–1706. https://doi.org/10.1111/1467-8624.ep9704041043

Cohen, G. A. (1989). On the currency of egalitarian justice. *Ethics*, *99*, 906–944.

Damon, W. (1977). *The social world of the child*. San Francisco: Jossey-Bass.

DeJesus, J. M., Rhodes, M., & Kinzler, K. D. (2014). Evaluations versus expectations: Children's divergent beliefs about resource distribution. *Cognitive Science*, *38*, 178–193. https://doi.org/10.1111/cogs.12093

Dworkin, R. (1981). What is equality? Part 2: Equality of resources. *Philosophy and Public Affairs*, *10*, 185–246.

Enright, E. A., Gweon, H., & Sommerville, J. A. (2017). 'To the victor go the spoils': Infants expect resources to align with dominance structures. *Cognition, 164,* 8–21. https://doi.org/10.1016/j.cognition.2017.03.008

Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature, 13,* 1–25.

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99,* 689–723. https://doi.org/10.1037//0033-295X.99.4.689

Kenward, B., & Dahl, M. (2011). Preschoolers distribute scarce resources according to the moral valence of recipients' previous actions. *Developmental Psychology, 47,* 1054–1064. https://doi.org/10.1037/a0023869

Kushnir, T., Gopnik, A., Chernyak, N., Seiver, E., & Wellman, H. M. (2015). Developing intuitions about free will between ages four and six. *Cognition, 138,* 79–101. https://doi.org/10.1016/j.cognition.2015.01.003

Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology, 20,* 163–198. https://doi.org/10.1016/S0732-118X(02)00007-7

Li, V., Spitzer, B., & Olson, K. R. (2014). Preschoolers reduce inequality while favoring individuals with more. *Child Development, 85,* 1123–1133. https://doi.org/10.1111/cdev.12198

Liberman, Z., & Shaw, A. (2017). Children use partial resource sharing as a cue to friendship. *Journal of Experimental Child Psychology, 159,* 96–109. https://doi.org/10.1016/j.jecp.2017.02.002

LoBue, V., Nishida, T., Chiong, C., DeLoache, J. S., & Haidt, J. (2011). When getting something good is bad: Even three-year-olds react to inequality. *Social Development, 20,* 154–170. https://doi.org/10.1111/j.1467-9507.2009.00560.x

McAuliffe, K., Blake, P. R., Steinbeis, N., & Warneken, F. (2017). The developmental foundations of human fairness. *Nature Human Behaviour, 1,* 1–9. https://doi.org/10.1038/s41562-016-0042

McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition, 134,* 1–10. https://doi.org/10.1016/j.cognition.2014.08.013

Meristo, M., Strid, K., & Surian, L. (2016). Preverbal infants' ability to encode the outcome of distributive actions. *Infancy, 21,* 353–372. https://doi.org/10.1111/infa.12124

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review, 106,* 3–19. https://doi.org/10.1037//0033-295X.106.1.3

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology, 16,* 329–337. https://doi.org/10.1037/h0029815

Moore, C. (2009). Fairness in children's resource allocation depends on the recipient. *Psychological Science, 20,* 944–948.

Nisan, M. (1984). Distributive Justice and Social Norms. *Child Development, 55,* 1020–1029. https://doi.org/10.2307/1130153

Nisan, M. (1989). Children's perceptions of effort and productivity as granting a right for reward. *British Journal of Developmental Psychology, 7,* 307–319.

Oishi, S., Kesebir, S., & Diener, E. (2011). Income inequality and happinness. *Psychological Science, 22,* 1095–1100. https://doi.org/10.1177/0956797611417262

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108*, 222–231. https://doi.org/10.1016/j.cognition.2007.12.003

Paulus, M. (2014). The early origins of human charity: developmental changes in preschoolers' sharing with poor and wealthy individuals. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2014.00344

Paulus, M. (2016). It's payback time: Preschoolers selectively request resources from someone they had benefitted. *Developmental Psychology, 52*, 1299–1306. https://doi.org/10.1037/dev0000150

Pesowski, M. L., Denison, S., & Friedman, O. (2016). Young children infer preferences from a single action, but not if it is constrained. *Cognition, 155*, 168–175. https://doi.org/10.1016/j.cognition.2016.07.004

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14-and 18-month-olds. *Developmental Psychology, 33*, 12–21.

Rizzo, M. T., Elenbaas, L., Cooley, S., & Killen, M. (2016). Children's recognition of fairness and others' welfare in a resource allocation task: Age related changes. *Developmental Psychology, 52*, 1307–1317. https://doi.org/10.1037/dev0000134

Rizzo, M. T., & Killen, M. (2016). Children's understanding of equity in the context of inequality. *British Journal of Developmental Psychology, 34*, 569–581. https://doi.org/10.1111/bjdp.12150

Rutte, C. G., Wilke, H. A. M., & Messick, D. (1987). Scarcity or abundance caused by people or the environment as determinants of behavior in the resource dilemma. *Journal of Experimental Social Psychology, 23*, 208–216.

Schäfer, M., Haun, D. B., & Tomasello, M. (2015). Fair is not fair everywhere. *Psychological Science, 26,* 1252–1260.

Schmidt, M. F., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PloS One, 6*. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0023223

Shaw, A. (2013). Beyond "to share or not to share" The impartiality account of fairness. *Current Directions in Psychological Science*, *22*, 413–417.

Shaw, A., Choshen-Hillel, S., & Caruso, E. M. (2016). The development of inequity aversion: Understanding when (and why) people give others the bigger piece of the pie. *Psychological Science*, *27*, 1352–1359. https://doi.org/10.1177/0956797616660548

Shaw, A., DeScioli, P., & Olson, K. R. (2012). Fairness versus favoritism in children. *Evolution and Human Behavior*, *33*, 736–745. https://doi.org/10.1016/j.evolhumbehav.2012.06.001

Shaw, A., & Knobe, J. (2013). Not all mutualism is fair, and not all fairness is mutualistic. *Behavior and Brain Sciences*, *36*, 100–101.

Shaw, A., & Olson, K. (2014). Fairness as partiality aversion: The development of procedural justice. *Journal of Experimental Child Psychology*, *119*, 40–53. https://doi.org/10.1016/j.jecp.2013.10.007

Shaw, A., & Olson, K. R. (2012). Children discard a resource to avoid inequity. *Journal of Experimental Psychology: General*, *141*, 382–395. https://doi.org/10.1037/a0025907

Sigelman, C. K., & Waitzman, K. A. (1991). The Development of Distributive Justice Orientations: Contextual Influences on Children's Resource Allocations. *Child Development*, *62*, 1367–1378. https://doi.org/10.2307/1130812

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*, 196–204. https://doi.org/10.1177/0956797611422072

Smetana, J. G. (1981). Preschool Children's Conceptions of Moral and Social Rules. *Child Development*, *52*, 1333–1336. https://doi.org/10.2307/1129527

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PLoS ONE*, *8*. https://doi.org/10.1371/journal.pone.0059510

Sommerville, J. A., Schmidt, M. F. H., Yun, J., & Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life: Fairness expectations and prosocial behavior. *Infancy*, *18*, 40–66. https://doi.org/10.1111/j.1532-7078.2012.00129.x

Temkin, L. S. (2011). Justice, equality, fairness, desert, rights, free will, responsibility, and luck. In C. Knight & Z. Stemplowska (Eds.), *Distributive justice and responsibility* (pp. 51–76). New York, NY: Oxford University Press.

Tisak, M. S., & Turiel, E. (1988). Variation in seriousness of transgressions and children's moral and conventional concepts. *Developmental Psychology*, *24*, 352–357.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge, England: Cambridge University Press.

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions: Children intervene in moral transgressions. *British Journal of Developmental Psychology*, *29*, 124–130. https://doi.org/10.1348/026151010X532888

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, *35*, 245–275.

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development, 78*, 16–33.