To Punish or Exclude? Children's Responses to Unfair and Fair Advantages

Created in Competitive Contexts

January 30, 2022

Alexander P. D'Esterre, Arvid Samuelson, & Melanie Killen

University of Maryland

Cognitive Development

Correspondence:
Alexander Paul D'Esterre
Department of Human Development and Quantitative Methodology
University of Maryland
College Park, MD, 20742
Phone: 609.410.5756 Email: desterre@umd.edu

Abstract

Word Count: 150 / 150

To determine whether children will exclude or punish a peer who creates an unfair advantage in an intergroup team context, four- to ten-year-old participants ($N = 120$, $M_{age} = 6.87$) were assigned a team membership and evaluated unintentional and intentional unfair advantages created by a character. Children were more likely to endorse punishment and exclusion responses when reasoning about an opponent than a teammate. This difference between groups was not observed when in-group and out-group members reasoned about punishing a character who intentionally created an unfair advantage. Older children were less likely to endorse exclusion than younger participants. Further, older children and in-group members utilized punishment more frequently than exclusion. Taken together this demonstrates that the group identity and the age of the child influences the ways in which children endorse responses to transgressions. These findings increase our understanding regarding children's conceptions of fairness responses to transgressions in intergroup contexts.

KEY WORDS: morality, intergroup attitudes, social exclusion, punishment, cheating

To Punish or Exclude? Children's Responses to Unfair and Fair Advantages

Created in Competitive Contexts

Social exclusion from groups occurs for many reasons, such as rejecting group members

who deviate from the rules and norms of the group (Killen & Rutland, 2011; Nesdale et al.,

2005; Schmidt & Tomasello, 2012) as well as excluding those who betray expectations regarding

group loyalty (Abrams & Rutland, 2008). While often discussed as a form of punishment, social

exclusion is different from most forms of punishment as it affects not only the individual who

has violated group norms but also the group from which they have been excluded. Additionally,

social exclusion can be utilized without the intent to punish and can be used as a legitimate

means to create effective group functioning, such as when a group has a goal that requires certain

abilities of its members (e.g., excluding individuals who lack sports ability for a competitive

team, musical talent for a band, or knowledge of chess strategy for a chess team). However,

when social exclusion is used to maintain power or status at the expense of fair treatment of

others, such as when individuals are excluded solely due to their group membership (e.g.,

gender, race, or ethnicity), exclusion is often viewed as detrimental both to the psychological

health of individuals who are excluded and to the larger societal community (Mulvey, 2016).

A robust area of research has investigated when and how children understand different

aims and goals of groups in the context of exclusionary behavior (Abrams et al., 2008;

Burkholder et al., 2020; McGuire et al., 2018). This research has shown that from early

childhood (Cooley & Killen, 2015) to adulthood (Abrams et al., 2005), peer inclusion and

exclusion involve knowledge about group dynamics as well as the fair and equal treatment of

others. Children begin to differentiate interpersonal exclusion from intergroup exclusion, for

example viewing the exclusion of an aggressive child from a group as okay and more legitimate

than excluding a child because of their nationality, and they do so for reasons based on both an

understanding of group dynamics as well as fairness (Malti et al., 2012). Children also

understand that group loyalty is an important part of group identity, and they often refrain from

challenging group norms out of concerns of being excluded (Abrams et al., 2008; Mulvey, 2016;

Nesdale & Lawson, 2011).

Social exclusion is an efficient mechanism by which group norms may be enforced, but

children will also utilize other methods such as punishing individuals who violate the agreed

upon rules and customs (Schmidt & Tomasello, 2012; Tomasello, 2014). Reprimanding

another's behavior is a form of punishment which is directed at an individual with the goal of

changing their behavior to fit the group norms. In contrast, social exclusion from the group is

aimed at changing the dynamics of a group by ousting an individual who appears to be disloyal

to the group norms or antagonistic to the goals of the group (Abrams et al., 2005). Thus, while

punishment and social exclusion are often used interchangeably in the research literature, a goal

of this study was to determine whether children differentiate these forms of responses to a

transgression committed by an ingroup or outgroup member of a team.

With age, children's recognition of group dynamics leads to a wider variety of responses

when individuals deviate from group norms (Hardecker et al., 2016). Acquiring this

understanding of group dynamics could mean that social exclusion would be perceived as more

disruptive and negative for group harmony than punishing and individually admonishing one

member of the group. Importantly, decisions to exclude or punish a peer may involve a number

of factors: the intentions of the actor, the perceived consequences of the act, the group identity of

the peer in relation to the individual, and the social context of the transgression, among others.

Of interest for this study is whether and when children endorse punishing and/or excluding

individuals for a perceived wrong behavior as a function of the character's intentions and group

identity. Thus, the current study was developed in order to measure children's responses to situations that varied in terms of whether the action was intentionally or unintentionally committed as well as whether the action benefitted the child's in-group or the out-group.

Intergroup dynamics and competition

With age, children's knowledge about how groups work expands across various contexts. One context in which norm violations are particularly salient for children, as well as adults, is intergroup competition (McGuire et al., 2018). Intergroup competition captures aspects of group dynamics that often fuel prejudicial attitudes, such as in-group bias and out-group distrust. As indicated by Nesdale (2004), out-group dislike forms when groups are threatened by competition. In several studies, Nesdale and his colleagues (2004) have shown that competition can incite a dislike of the out-group when groups have an exclusive rather than an inclusive norm. Additionally, previous work has indicated that children's moral judgments can be influenced by these group identities, such that children respond more favorably to lies told to benefit a child's in-group than those than benefit an out-group (Sierksma et al., 2019). For these reasons, competition is likely to be a social context which heightens the importance of group identity and which in turn provides an ideal context to investigate the interaction between group identity and children's punishment and exclusion decisions.

Frequently, children justify intergroup exclusion based on interpersonal categories, often by conflating group membership with personality characteristics. For example, individuals may reject a group member based on their ethnicity due to an assumption that members of different ethnic groups do not share the same values (Hitti & Killen, 2015). In this situation, a link is made between group membership and personality characteristics that may be unfounded or the result of stereotypic expectations (Killen et al., 2013a). While there is research on the stereotypical

characteristics that are associated with group identity (e.g., smart/dumb, friendly/mean) less is known about the impact of group identity on children's understanding of contextually important, but non-stereotyped traits. For the purpose of this study, we were interested in investigating whether children differentially recognize the *intentions* of others due to their group identity, how children's age influences this relationship, and whether children's group identity influences their decision to respond to potential norm violations through the use of social exclusion or punishment.

Children's understanding of when to assign blame has been shown to be highly related to the extent to which they understand that intentions have to be differentiated from the consequences or outcomes of one's actions (Lagattuta & Weller, 2014). That is, someone can commit a bad act with good intentions, or a good act with bad intentions; in both cases, the intentions are the crucial information necessary for assigning blame. As an example, in a situation in which one child accidentally destroys another child's property, younger children are more likely than older children to assign blame to an accidental transgressor, to expect that the victim (property owner) will be angry at the accidental transgressor, and to expect that that the accidental transgressor intentionally destroyed the property (Killen et al., 2011). Further, research has shown that assignment of blame varies in contexts that involve group identity (D'Esterre et al., 2022). What has not been examined, however, is whether children would exclude or punish an individual who intentionally or unintentionally provides an advantage for their group in a competitive context.

In situations with high levels of group affiliation, such as found in a competitive team context, accusations are frequently made about intentions, particularly for members of the child's out-group (Yazdi et al., 2020). This raises an interesting and ecologically valid context to explore

whether children would determine that exclusion or punishment is necessary when evaluating an opposing team member's intentional or unintentional attempts to create an unfair advantage.

The Current Study

Thus, the present study was designed to examine the role of intentions and group identity regarding children's decisions about social exclusion and punishment in an intergroup context. To investigate the role of intentionality in a competitive intergroup context, participants were inducted into one of two teams (red team or the blue team) and then asked to evaluate different situations in which a character from either their in-group or the out-group: 1) unintentionally created an unfair advantage; 2) intentionally created an unfair advantage; and 3) created a fair advantage. Children were asked whether the team member should be punished and whether they should be excluded from the team.

For multiple reasons the present study was focused on children between the ages of four- to ten-years old. First, group identity has been shown to be extremely important to children during this age period (Abrams & Rutland, 2008; Nesdale, 2008) and recent research has shown that children's evaluations of fairness can be influenced by the group identity of relevant parties (McGuire & Rutland, 2020). Second, we know that concerns for fairness or quite salient during early and middle childhood (Smetana et al., 2014), indicating that the children care about the decision making that results in a fair or unfair outcome. Finally, children's understanding of intentionality (and mental state knowledge) changes during this age period (Lagatutta & Weller, 2014; Rizzo et al., 2018) suggesting that younger children would be less likely to differentiate intentional and unintentional advantages than older children from 4 – 10 years of age.

This last point in particular led us to sub-divide our sample into younger (4- to 6-years-old) and older (7- to 10-years-old) groups, in keeping with related research documenting

intentionality judgments and fairness considerations, and for comparisons of findings across studies (Killen et al., 2011; Glidden et al., 2021).

Methods

Participants

Participants were 120 children ages 4–10 (M_{Age} = 6.87 years, SD_{Age} = 1.81; 52% female), recruited from preschools and summer camps serving lower-middle to upper-middle income families in the Mid-Atlantic region of the United States. Participants were ethnically diverse (67% European American, 18% African American, 11% Asian American, and 4% Hispanic). Our participants were split by age into a group of 59 younger children (four- to six-years-old: 25 four-year-olds, 19 five-year-olds, and15 six-year-olds) and 61 older children (seven- to ten-years-old: 27 seven-year-olds, 15 eight-year-olds, 15 nine-year-olds, and 4 ten-year-olds). Previous research on these topics have revealed small to medium effects (Bernhard et al., 2020; McGuire et al., 2018), and thus an a priori power analyses using G*Power (Faul et al., 2009) indicated that at least 100 participants would be necessary to detect significant effects for our omnibus test.

Design

To examine the relationship between group identity, age, and advantage type, a 2 (group affiliation: in-group, out-group) x 2 (age: 4- to 6-years-old, 7- to 10-years-old) x 3 (advantage type: unintentional unfair, intentional unfair, or fair advantage) mixed-factorial design was utilized. Group affiliation and age were between-subject manipulations while advantage type was a within-subject manipulation.

Previous research on children's judgments of intentional and unintentional transgressions have indicated the existence of order effects, and therefore all participants were presented with the three advantage types presented in a fixed order. Specifically, it has been found that children who viewed an intentional transgression prior to an unintentional transgression were more likely to view the unintentional transgression as more acceptable than participants who heard the unintentional transgression first (D'Esterre et al., 2019). Presenting straightforward moral transgressions first provides children with a clearly unacceptable scenario, and causes children to quickly and simplistically reason that the unintentional transgressions are more acceptable. However, by providing the more nuanced unintentional transgression first, children were able to reason about this context without being primed with a simpler transgression beforehand. Therefore, in order to remove possible order effects from confounding the interpretation of observed differences all participants heard the unintentional unfair advantage first, followed by the intentional unfair advantage, and finally they were presented with the fair advantage context.

Procedure

This project was approved by the Institutional Review Board at [institution masked]. Individually administered interviews by trained research assistants were conducted in a quiet space in participants' schools, with each interview lasting approximately 15–20 minutes. Children were randomly assigned to the red (disadvantage) or blue (advantage) team. In order to increase the saliency of their team membership children selected a team logo in their team color to hold throughout the stories (see Killen et al., 2013b). Children also selected whether their team would get an ice cream or pizza party (hypothetically) at the end of the competition in order to further increase the saliency. A manipulation check asking children how happy or sad they would be if their assigned team won and how happy or sad they would be if the other team won

revealed that 87% of participants demonstrated an in-group preference even before hearing about

the results of the competition.

Research assistants showed participants a PowerPoint presentation and read a script of

the stories. Participants were explicitly represented in the story with a gender-matched silhouette

character entitled "you" (see Figure 1). The other characters were portrayed as approximately the

participant's age and were ethnically diverse. Stories began with an introduction about a

Figure 1.
Participants' view of team set-up for the competition
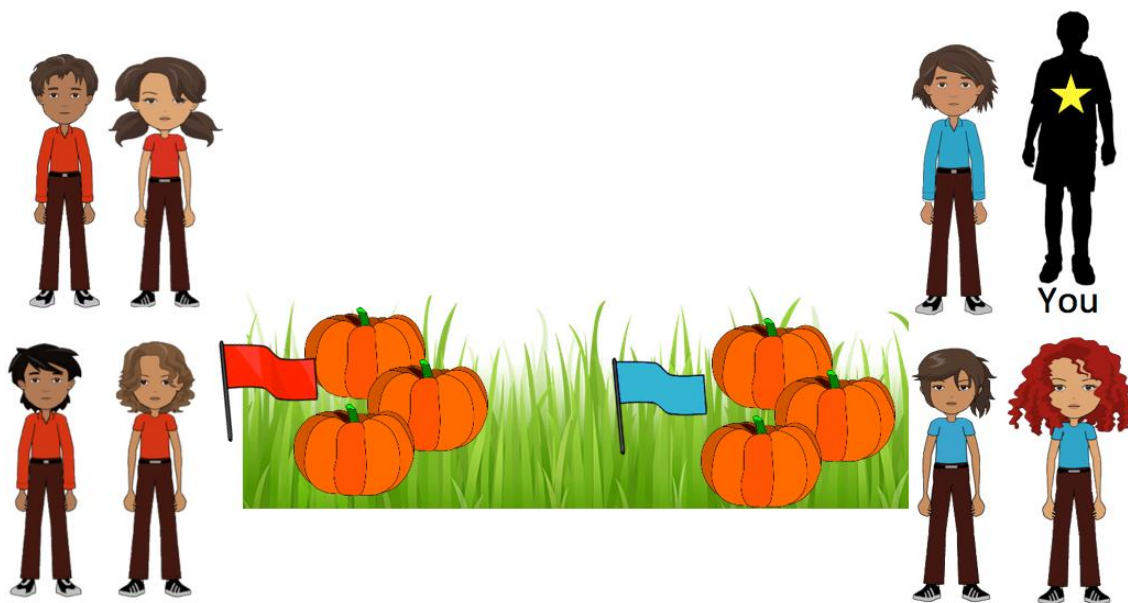
pumpkin growing contest between the red and the blue teams. Participants were told that the only

rule in the contest is that each team is only allowed to feed their team's pumpkins one cup of

plant food per day to help them grow. They were then shown the three stories: 1) unintentional

unfair advantage, 2) intentional unfair advantage, and 3) fair advantage. Each of the three stories

had accompanying animated illustrations that took them through the entire story and

accompanied the actions described by the experimenter. At the start of the stories, all participants

were reminded of the rule. All participants were presented with the stories in the same order

presented above (D'Esterre et al., 2019).

Stories

   Unintentional Unfair Advantage (UUA). Participants were told about Sam, a character

on the blue team. The story explained that it was Sam's turn to feed the pumpkins, but he could

not find the plant food, so he left to look for it. While Sam was away, his teammates found some

plant food. Both teams fed their pumpkins, and then everyone left. After they had gone, Sam

came back and fed his team's pumpkins. The blue team grew the biggest pumpkin and won the

contest. At this point, a memory check question was asked: "Is Sam on your team or is Sam on

the other team?". Children who answered correctly were asked a series of other questions.

Children who incorrectly responded had the manipulation repeated to them two or fewer times,

and all children successfully passed this memory check.

   Intentional Unfair Advantage (IUA). Children were introduced to Taylor, a blue team

member. Participants were told that it was Taylor's turn to feed the pumpkins, but she could not

find the plant food. They were informed that after she looked around, she found the plant food,

and both Taylor and the red team fed their pumpkins. Then, after everyone left, Taylor came

back with plant food for her team's pumpkins and then fed them again. Previous research

examining children's perceptions of intentions suggests that young children often mistakenly

attribute forgetfulness as a rationale for intentional transgressions (D'Esterre et al., 2019), and

therefore it was stated that Taylor remembered that she fed the pumpkins earlier. Children were

told that the blue team won the contest. At this point a memory check was asked: "Is Taylor on

your team or is Taylor on the other team?" After this, the procedure was the same as the

unintentional unfair advantage. All children successfully passed the memory check.

Fair Advantage (FA). In addition to comparing children's retributive attitudes toward unintentional and intentional unfair advantages, we wanted to provide children with a third comparison condition in which an advantage was created, but without a rule violation. This third condition was intended to allow us to determine how much of children's judgments was due to the unfairness of the previous scenarios as opposed to a simple dislike for the asymmetry present in an advantage.

To this end, the final story mentioned Casey, a member of the blue team, and participants were told that it was her turn to feed the pumpkins. Participants were informed that it was a very nice day outside and that the other contestants decide to go to the park instead of feeding their pumpkins, but Casey decided to stay and feed her team's pumpkins instead. Children were told that the blue team won. Once again, a memory check was administered: "Is Casey on your team or is Casey on the other team?" After this, the procedure was the same as the unintentional unfair advantage, and all children successfully passed the memory check.

Measures

After finishing each story, children were asked two comprehension measures regarding the team identity of the advantage creator ("Was [Sam/Taylor/Casey] a member of your team or the other team?") and the knowledge of the advantage creator ("Did [Sam/Taylor/Casey] think that the pumpkins had already been fed?"). Following the comprehension questions children were then administered social cognitive assessments consisting of questions about their own perspective regarding how to respond to the characters. These assessments included *exclusion* and *punishment.* These measures were selected as they provided insight into children's preferred reactions to the behavior of the characters following the act that created an advantage. Thus, the

goal was to determine what children think should happen *after* a character has created an advantage in contexts that varied by the knowledge and intentions of the advantage creator.

Punishment. Punishment of the character was assessed by the researcher asking, "Do you think [Sam/Taylor/Casey] should get in trouble for feeding the pumpkins?" Requests for elaboration were very infrequent, and in the few instances where elaborate was necessary the experimenter indicated that getting in trouble takes the form of being talked to by the teacher or put in a short time-out. This form of retribution affected the individual and was not group focused. Children's answers were coded as a binary "yes" or "no" response.

Exclusion. All children were given an assessment to inquire about exclusion from the team. The researcher asked, "Do you think that [Sam/Taylor/Casey] should be off the team?" Exclusion was utilized as it was believed to be a form of retribution which not only affected the individual, but also created a disruption in the group dynamics of the team. Children were provided with the options of "yes" or "no."

In addition to these measures, children were also asked to provide moral assessments of the characters' actions, but these questions were not analyzed as a part of the hypotheses covered in this manuscript. A full discussion of children's responses to these moral assessment measures can be found in the manuscript dealing with those hypotheses (CITATION OMITTED FOR BLIND REVIEW).

Hypotheses

The first three sets of predictions were focused on participants' utilization of both exclusion and punishment across the different advantage contexts to test whether participants were using information about intentionality and group identity in determining whether the peer should be excluded or punished. The fourth set of predictions were focused on participants'

differentiations between exclusion and punishment for each intentionality and group identity context. Age-related expectations were part of both sets of hypotheses. Due to previous research showing a shift in children's mental state understanding

Thus, we predicted that, overall, participants would be more likely to endorse both punishment and exclusion when an unfair advantage was created intentionally than unintentionally (H1a), and that participants would be more likely to endorse both exclusion and punishment in situations with an unfair advantage than a fair advantage (H1b).

We also predicted that participants who shared a group identity with the advantage creator would be less likely to endorse punishment and exclusion than would participants who belonged to the competing team, when averaging across the three advantage contexts (H2a). However, we also predicted that group differences in punishment and exclusion would be less pronounced for the intentional unfair advantage given that there was no dissociation between intentions and outcomes regarding the intentionally unfair effort to create an advantage (H2b). Further, we predicted that participants in the in-group condition would be more sensitive to the variations in intentions than would participants in the out-group, and thus we hypothesized that in-group participants would differ in their ratings of punishment and exclusion judgments across all three contexts, while participants in the out-group condition would view the fair advantage most positively but not distinguish between the unintentional and intentional unfair advantages (H2c).

With age, it was expected that participants would be less likely to endorse punishment and exclusion overall (H3a). It was further predicted that this difference would be largely driven by responses in advantage contexts with more complex intention information (e.g., unintentional unfair and fair advantages) and that differences would not be present for the straightforward

moral transgressions of the intentional unfair advantage (H3b). We also predicted that older

children would be better capable of encoding this complex intentional information and would

thus show different rates of punishment and exclusion across the three contexts while younger

children would only distinguish between the fair and unfair advantages (H3c).

Finally, we predicted that participants would endorse punishment and exclusion at

different rates, with punishment being more common than exclusion overall (H4a) as punishment

allows children to condemn the actions of an individual without negatively affecting the entire

group. Additionally, we predicted that in-group members would be less likely to endorse

exclusion than punishment as exclusion would influence their in-group's functioning while out-

group members were not expected to endorse these options at different rates (H4b). It was also

predicted that older participants would be less likely to endorse exclusion than punishment as

they would be more cognizant of the impact of exclusion on group dynamics and thus less likely

to endorse it, while we did not expect young children to distinguish between the two options to

the same extent (H4c).

Data analytic plan

In order to explore both the within and between-subject manipulations a Repeated

Measures General Linear Model (GLM) was used. While binary logistic regressions are often

utilized for binary DVs, this data analytic approach is not feasible with a repeated measures design.

Consultation with statisticians confirmed that a GLM was appropriate given the repeated measures nature

of our design and the fact that simulation studies have shown that repeated measures GLM are fairly

robust to violations of the normality assumption. This approach has also been further supported by

Oberfeld and Franke (2012). Participants' responses to the social cognition item (e.g., *Exclusion*)

for each of the three advantage contexts was entered as the within-subject variable, and group

affiliation (in-group and out-group) and children's age (4- to 6-years-old and 7- to 10-years-old)

were entered as between-subject factors. An independent-samples t-test comparing the age (in months) of participants in the in-group and out-group conditions revealed no significant difference, $t(118) = 1.074$, $p = .285$, thus a continuous measure of age was not entered as a covariate in subsequent analyses. In order to investigate the predicted interactions between advantage contexts and team membership and between advantage contexts and age, post-hoc analyses were conducted utilizing independent and paired samples t-tests.

Results

Children's Punishment Decisions

Children's Punishment Across Advantage Types. In line with the hypothesis that participants would show different rates of punishment endorsement for the different advantage types, a main effect of advantage type was found, $F(1,117) = 100.231$, $p < .001$, $\eta p^2 = .461$. In order to further examine this effect post-hoc pairwise comparisons indicated that participants did not differ in their rate of punishment endorsement between the unintentional and intentional unfair advantage contexts, when collapsing across group membership (UUA: 52.1%, $SE = 4.3\%$; IUA: 62.7%, $SE = 3.9\%$), $p = .102$, which was counter to our initial hypothesis (H1a). However, participants were less likely to endorse punishment for the advantage creator when the advantage was fair than when it was unfair (FA vs UUA: $p < .001$; FA vs IUA: $p < .001$) (H1b).

Children's Punishment Based on Group Identity. The second hypothesis, that participants would be less likely to endorse punishing the advantage creator in the in-group condition than the out-group was supported by the findings of a main effect of group, $F(1,117) = 21.698$, $p < .001$, $\eta p^2 = .156$. Post-hoc pairwise comparisons showed that participants in the in-group condition were significantly less likely to suggest punishment than participants in the out-group condition, (In-group: 31.5%; Out-group: 55.5%), $p < .001$ (H2a), as shown in Figure 2.

Interaction of Advantage Type and Group Identity for Punishment. A significant interaction between group identity and advantage type was not found, $F(1,117) = 1.319$, $p = .253$, $\eta p^2 = .011$, for the punishment measure. While this interaction was not significant, post-hoc analyses were conducted in order to investigate H2b and H2c for the punishment measure. First, in order to test the hypothesis that team membership would lead to a different exclusion endorsement pattern in the more socially complex contexts but not straightforward transgressions (H2b), independent samples t-tests were conducted. As shown in Figure 2, these analyses revealed that participants in the in-group were less likely to endorse punishment in the UUA, $t(118)=3.613$, $p < .001$, and FA, $t(118)=3.773$, $p < .001$, contexts. However, in-group and out-group participants endorsed punishment in the IUA context equally, $t(118)=1.804$, $p = .074$. This supports our prediction and suggests that, while group membership influenced punishment endorsement decisions for contexts in which intentions and outcomes differed (good/neutral intentions leading to an advantage for one team), in straightforward moral transgressions group identity did not change the way participants responded.

Next, paired sample t-tests were conducted to determine whether participants in the in-group and out-group conditions responded differently across the three advantage contexts (H2c). These tests showed that participants in the in-group condition punished significantly less in the unintentional context than the intentional context (UUA vs IUA), $t(57) = 2.385$, $p = .020$), and that they also punished less in the fair advantage condition than in the unintentional unfair

Figure 2.

Comparison of In-Group and Out-Group Punishment Decisions Across Advantage Types.
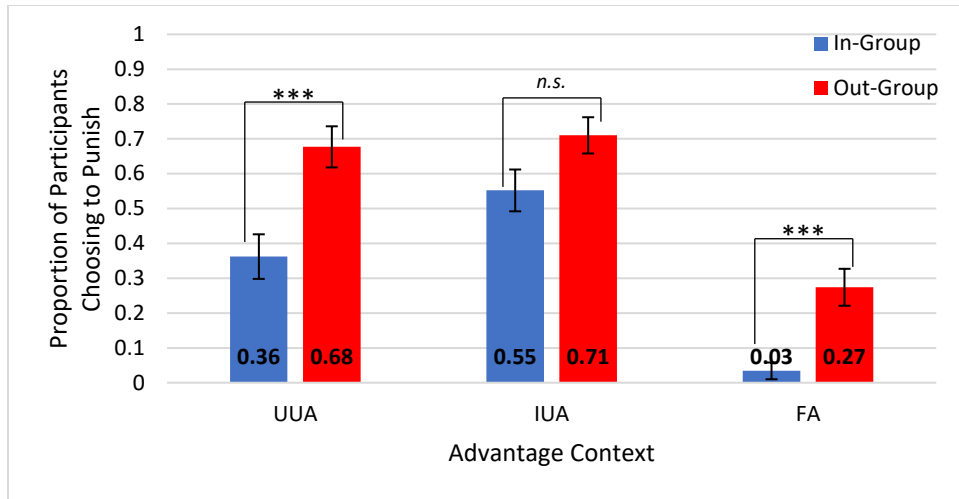
Note. N = 120. UUA = Unintentional Unfair Advantage, IUA = Intentional Unfair Advantage,

FA = Fair Advantage.

advantage condition (FA vs UUA), t(57) = 5.270, p < .001, and less in the fair advantage

condition than in the intentional unfair advantage condition (FA vs IUA), t(57) = 7.325, p < .001.

In contrast, participants in the out-group condition did *not* show different rates of punishment

endorsement for the unintentionally and the intentionally unfair advantage contexts (UUA vs

IUA), t(61) = .405, p = .687, but they did punish less in the fair advantage context than an in

unintentional unfair advantage context (FA vs UUA), t(61) = 5.180, p < .001, or the intentional

unfair advantage context (FA vs IUA), t(61)=5.331, p < .001. These results supported H2c and

suggest that participants in the in-group condition endorsed punishment differently as a function

of intentions while participants in the out-group condition did not.

Children's Punishment Based on Age. Primary age expectations (H3a) were not

supported by the analyses, F(1,117) = .147, p = .702, ηp^2 = .001. This result suggested that

participants had, across the three advantage contexts, no significant differences in their rate of

punishment endorsement based on their age (Younger: 42.5%; Older: 44.5%).

Interaction of Advantage Type and Age for Punishment. While a significant main effect was not found, a significant interaction between age and advantage type *was* found, $F(1,117) = 40.412$, $p < .001$, $\eta p^2 = .257$. To further elucidate the nature of this interaction, post-hoc analyses were conducted in order to investigate H3b and H3c for the punishment measure.

First, in order to test the hypothesis that age would lead to a different punishment endorsement pattern in the more socially complex contexts but not straightforward transgressions (H3b), independent samples t-tests were conducted. As shown in Figure 3, these analyses revealed that older participants were less likely to endorse punishment in the UUA, $t(118) = 2.230$, $p = .028$, and FA, $t(118) = 2.905$, $p = .004$, contexts. However, older children were

Figure 3

Comparison of Younger and Older Participants' Punishment Decisions Across Advantage Types.
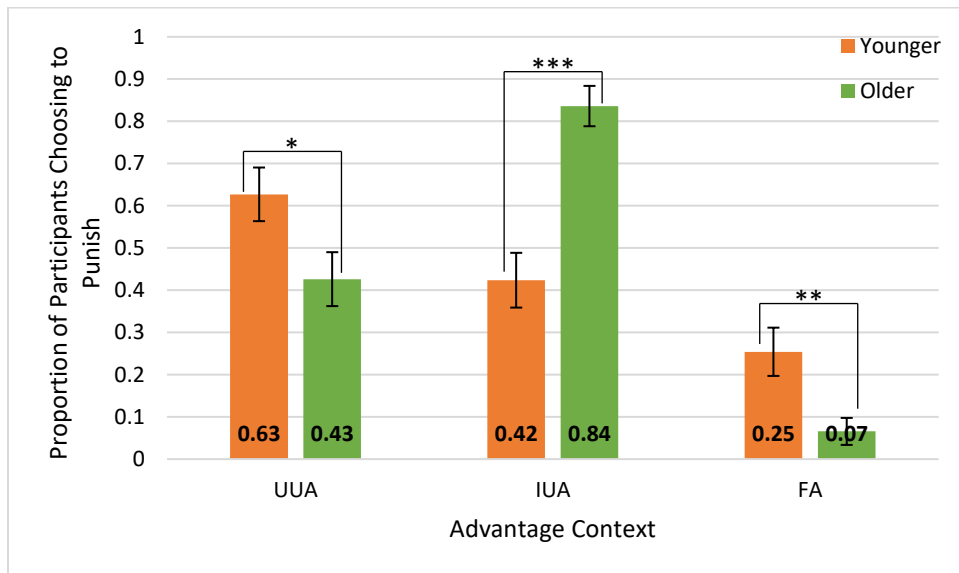
Note. $N = 120$. UUA = Unintentional Unfair Advantage, IUA = Intentional Unfair Advantage, FA = Fair Advantage.

actually *more likely* to endorse punishment in the IUA context than were the younger participants, $t(118) = -5.141$, $p < .001$. This suggests that, with age, children were less likely to

endorse punishment in contexts in which intentions and outcomes differed (good/neutral intentions leading to an advantage for one team over another), while being more likely to endorse punishment in response to straightforward moral transgressions. Next, paired sample t-tests were conducted to determine whether younger and older participants responded differently across the three advantage contexts (H3c). These tests showed that older participants punished significantly less in the unintentional context than the intentional context (UUA vs IUA), $t(60) = 6.057$, $p < .001$, and that they also punished less in the fair advantage condition than in the unintentional unfair advantage condition (FA vs UUA), $t(60) = 5.133$, $p < .001$, and less in the fair advantage condition than in the intentional unfair advantage condition (FA vs IUA), $t(60) = 13.035$, $p < .001$.

Unexpectedly, younger participants endorsed punishment for the unintentional unfair advantage more than they did for the intentional unfair advantage (UUA vs IUA), $t(58) = 2.839$, $p = .006$. Similarly to the older participants, younger children punished less in the fair advantage context than the unintentional unfair advantage context (FA vs UUA), $t(58) = 5.171$, $p < .001$, and less in the fair advantage than the intentional unfair advantage context (FA vs IUA), $t(58) = 2.317$, $p = .024$. These results largely supported H3c, however we did not expect younger children would differentiate between the unintentional and intentional unfair advantages, nor would we have predicted the difference in the observed direction.

Summary of Results for Punishment Assessment. These results support the hypothesis that children endorse punishment significantly less for fair advantages than unfair advantages and that children were less likely to endorse punishment for an in-group member than an out-group member. While participants responding to an advantage created by an in-group member endorsed punishment less for unintentional unfair advantages than intentionally unfair

advantages, members of the out-group did not show the same distinction and endorsed punishment equally for both contexts. Further, while group identity led to significant differences in the ways that participants responded to advantages created unintentionally and to fair advantages, participants showed equal levels of support for punishing a character who intentionally created an unfair advantage regardless of their own team membership. Additionally, younger and older children did not show any differences in punishment endorsement overall. Finally, it was revealed that older participants were much less likely to endorse punishment for accidental transgressions (unintentional unfair advantages) or fair advantages while they were more likely to endorse punishing an intentional transgression.

Children's Exclusion Decisions

Children's Exclusion Across Advantage Types. Supporting the hypothesis that participants would differ in their exclusion decisions for the different advantage types, a main effect of advantage type was found, $F(1,117) = 19.263$, $p < .001$, $\eta p^2 = .141$. Post-hoc pairwise comparisons revealed that, similar to their punishment decisions, participants did not differ in their endorsement of exclusion between the unintentional and intentional unfair advantage conditions when collapsing across group identity (UUA: 33.6%, $SE = 3.5\%$; IUA: 41.8%, $SE = 4.2\%$), $p = .353$. This was, once again, counter to our initial hypothesis (H1a). Also in line with the pattern of results for the punishment measure, participants were less likely to recommend that the target be excluded from the team in the fair advantage context than they were for either of the unfair advantages (FA vs UUA: $p < .001$; FA vs IUA: $p < .001$) (H1b).

Children's Exclusion Based on Group Identity. Our hypothesis that participants would be less likely to exclude in the in-group condition was supported by the findings of a main effect of group, $F(1,117) = 57.855$, $p < .001$, $\eta p^2 = .331$. This indicated that participants in the in-group

and out-group conditions differed in their exclusion responses when averaging across the three

advantage contexts. Post-hoc pairwise comparisons revealed that participants in the in-group

condition were significantly less likely to exclude than those in the out-group condition (In-

group: 12.0%; Out-group: 49.2%), $p < .001$ (H2a), as shown in Figure 4.

Interaction of Advantage Type and Group Identity for Exclusion. The interaction

effect between group identity and advantage type was also significant, $F(1,117) = 6.421$, $p =$

$.013$, $\eta p^2 = .052$. Therefore, this relationship was investigated further through post-hoc analyses.

First, in order to test the hypothesis that team membership would lead to a different

Figure 4

Comparison of In-Group and Out-Group Exclusion Decisions Across Advantage Types.
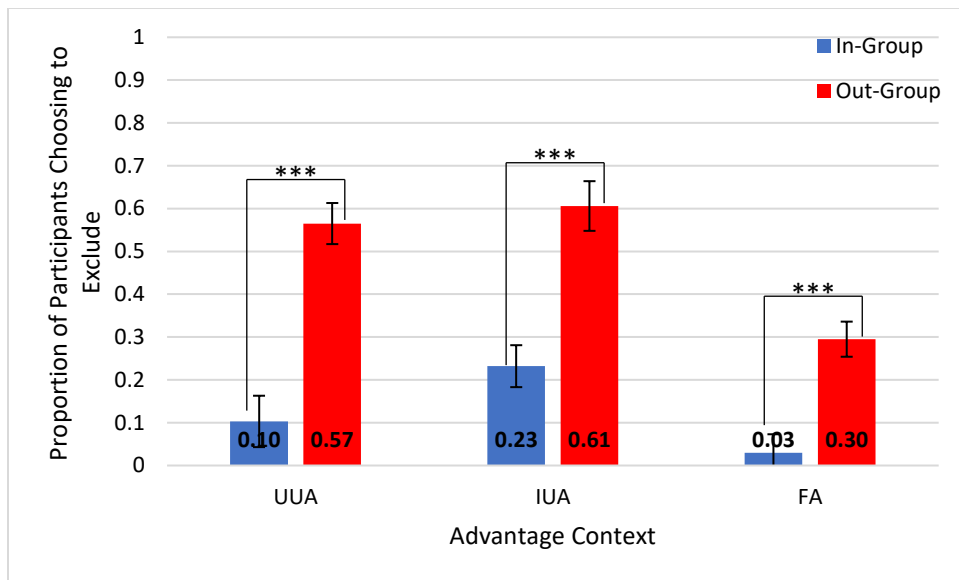
Note. $N = 120$. UUA = Unintentional Unfair Advantage, IUA = Intentional Unfair Advantage,

FA = Fair Advantage.

exclusion endorsement pattern in the more socially complex contexts but not straightforward

transgressions (H2b), independent samples t-tests were conducted. Again, as shown in Figure 4,

these analyses revealed that for each of the three advantage conditions, participants in the in-

group condition were less likely to exclude the target than were participants in the out-group condition (UUA: t(118) = 6.630, *p* < .001; IUA: t(118) = 4.182, *p* < .001; FA: t(118) = 4.570, *p* < .001). This differed slightly from the prediction that differences would only be found for the UUA and FA contexts but not in the IUA context and possible explanations will be discussed later.

Next, paired sample t-tests were conducted to determine whether participants in the in-group and out-group conditions responded differently across the three advantage contexts (H2c). These tests showed that participants in the in-group condition excluded significantly less in the unintentional unfair context than the intentional unfair context (UUA vs IUA), t(57) = 2.619, *p* = .011, and that they also excluded less in the fair advantage condition than in the unintentional unfair advantage context (FA vs UUA), t(57) = 2.055, *p* = .044, and less in the fair advantage than in the intentional unfair advantage context (FA vs IUA), t(57) = 4.058, *p* < .001. In contrast, participants in the out-group condition did *not* show different rates of exclusion decisions between the unintentionally and the intentionally unfair advantage contexts (UUA vs IUA), t(61) = .178, *p* = .859, but they did exclude less in the fair advantage context than in the unintentional unfair advantage (FA vs UUA), t(61) = 3.938, *p* < .001, and intentional unfair advantage context (FA vs IUA), t(61)=3.741, *p* < .001. These results supported H2c and suggest that participants in the in-group condition changed their exclusion decision based on the intentions of the advantage creator to a greater extent than did participants in the out-group condition.

Children's Exclusion Based on Age. Our hypothesis that participants would exclude differently as a function of their age (H3a) was supported by the finding of a main effect of age on children's exclusion endorsement, $F(1,117) = 7.707$, *p* = .006, $\eta p^2 = .062$. Post-hoc pairwise

comparisons revealed that, with age, participants were less likely to endorse exclusion (Younger: 37.4%; Older: 23.8%), $p = .006$.

Interaction of Advantage Type and Age for Exclusion. The interaction effect between age and advantage type was also significant, $F(1,117) = 13.622$, $p < .001$, $\eta p^2 = .104$. Therefore, this relationship was investigated further through post-hoc analyses. First, in order to test the hypothesis that age would lead to a different exclusion endorsement pattern in the more socially complex contexts but not straightforward transgressions (H3b), independent samples t-tests were conducted. As shown in Figure 5, these analyses revealed that older participants were less likely to exclude the target than were younger participants in both the unintentional unfair advantage (UUA), $t(118) = 4.005$, $p < .001$, and fair advantage contexts (FA), $t(118) = 3.117$, $p = .002$. However, in line with our prediction, no significant difference was found in exclusion endorsement for the straightforward transgression in the intentional unfair advantage context (IUA), $t(118) = .762$, $p = .448$.

Next, paired sample t-tests were conducted to determine whether participants responded differently across the three advantage contexts with age (H3c). These tests showed that older participants excluded significantly less in the unintentional context than the intentional context (UUA vs IUA), $t(60) = 3.573$, $p = .001$), and they also excluded less in the fair advantage context than in the unintentional unfair advantage context (FA vs UUA), $t(60) = 2.789$, $p = .007$, or the intentional unfair advantage context (FA vs IUA), $t(60) = 5.525$, $p < .001$. In contrast, younger participants did *not* show different rates of exclusion decisions between the unintentionally and

Figure 5

Comparison of Younger and Older Participants' Exclusion Decisions Across Advantage Types
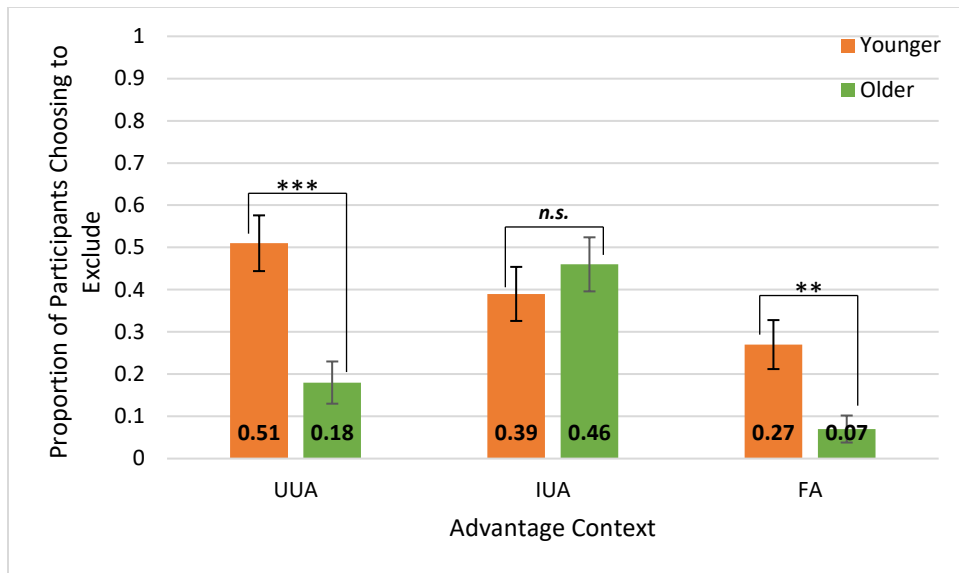
Note. N = 120. UUA = Unintentional Unfair Advantage, IUA = Intentional Unfair Advantage,

FA = Fair Advantage.

intentionally unfair advantage contexts (UUA vs IUA), $t(58) = 1.726$, $p = .090$, nor did they

show different rates of exclusion for the intentionally unfair and fair advantage contexts (IUA vs

FA), $t(58) = 1.990$, $p = .051$. Only when comparing the unintentional unfair advantage context to

the fair advantage did we find significant differences, with younger participants excluding

significantly less for the fair advantage (FA vs UUA), $t(58) = 3.399$, $p = .001$. The results of the

older sample fully supported our predictions in H3c, but once again the younger participants

displayed a pattern that was not in line with our predictions.

      Summary of Results for Exclusion Assessment. These results revealed that children

were less likely to support the removal of a character who created a fair advantage than a

character who created an unfair advantage. In addition, in-group members were less likely to

endorse exclusion than out-group members, and older children were less likely to endorse

exclusion than were younger children. In addition to these main effects, evidence was also found

to support the existence of significant interactions between advantage contexts and group identity

as well as an interaction between advantage contexts and age. Participants in the out-group endorsed exclusion at equal rates for the unintentional and the intentional unfair advantage context while those in the in-group endorsed exclusion less frequently for the unintentional unfair advantage than they did for the intentional unfair advantage.

This pattern suggests that children who did not share a group identity with the advantage creator were focused on whether or not a rule violation occurred but did not seem to consider the intentions of the advantage creator when making their exclusion decision. We also found that older children were less likely to endorse exclusion than younger children for the unintentional unfair and fair advantage conditions while children across the age range were equally likely to endorse exclusion for an intentionally created unfair advantage.

Comparisons between Punishment and Exclusion

Paired sample t-tests were conducted to determine if participants were more likely to endorse either punishment or exclusion in response to each of the three advantage conditions. Averaging across in-group and out-group participants, children were much more likely to endorse punishment than exclusion for the UUA, $t(119) = 3.875$, $p < .001$, and IUA contexts, $t(119) = 4.415$, $p < .001$; no significant differences were found for the FA context, $t(119) = .276$, $p = .783$. This partially supported the prediction (H4a) that participants would be less likely to endorse exclusion than they would to endorse punishment for the UUA and IUA contexts. It was unexpected that participants would not make the same distinction for the fair advantage context.

To further investigate the possibility that in-group and out-group members distinguished between exclusion and punishment to different extents we conducted additional paired sample t-tests to determine response patterns for in-group and out-group participants (H4b). These tests showed that participants in the in-group condition were significantly less likely to endorse

exclusion than punishment for the unintentional unfair, t(57) = 4.660, p = .044, and intentional

unfair advantage, t(57) = 4.700, p < .001, conditions. Again, participants did not differ in their

rates of punishment and exclusion endorsement for the fair advantage, t(57) = 1.000, p = .322.

This latter finding was likely due to a floor effect for both measures for the fair advantage

context (Exclusion: 3%; Punishment: 2%). In contrast, those in the out-group condition did not

show a significant difference in their endorsement of punishment or exclusion for any of the

three conditions (UUA: t(61) = 1.351, p = .182; IUA: t(61) = 1.725, p = .090; FA: t(61)= -.574, p

= .568), suggesting that participants who did not share a group identity with the transgressor

were less discerning in the level of retaliation that they would condone.

Finally, we also sought to investigate differences in punishment and exclusion

endorsement for younger and older participants (H4c). To this end, a set of paired sample t-tests

were conducted to compare responses on the two measures for each advantage context (Table 1).

Older participants were significantly less likely to endorse exclusion than punishment for the

unintentional unfair, t(60) = 3.223, p = .002, and intentional unfair advantage, t(60) = 5.645, p <

.001, contexts. Further, they did not differ in their rates of punishment and exclusion

endorsement for the fair advantage, t(60) = 0, p = 1.00, however this is likely due to a floor effect

(Exclusion: 7%; Punishment: 7%). Younger children only showed a significant difference in

their endorsement of punishment or exclusion for the unintentional unfair advantage condition,

t(58) = 2.176, p = .034, and did not differ for the intentional unfair, t(58) = .574, p = .568, or fair

advantage, t(58)= -.299, p = .766, contexts. This suggests that younger participants may

differentiate between the two forms of retribution to the same extent, as their older peers.

Table 1

Proportion of Participants Endorsing Exclusion Versus Punishment Across Advantage Contexts.

Advantage Context	Measure	Proportion of Participants Endorsing Retribution				
		Total	In-Group	Out-Group	Young	Old
Unintentional Unfair Advantage	Punishment	.53***	.36***	.68	.63*	.43**
	Exclusion	.34***	.09***	.58	.51*	.18**
Intentional Unfair Advantage	Punishment	.63***	.55***	.71	.42	.84***
	Exclusion	.43***	.24***	.60	.39	.46***
Fair Advantage	Punishment	.16	.03	.27	.25	.07
	Exclusion	.17	.02	.31	.27	.07

Note. N = 120. All significance indicators are for comparisons between values within a given cell

p < .05. **p* < .01. ***p* < .001.

Discussion

The novel findings of this study were that children's group membership, the characters' intentions, and their age had significant impacts on the ways in which they endorsed punishment and exclusion. The present study utilized ad-hoc groups in competitive contexts rather than pre-existing social identities, minimizing the potential for real world associations and pre-existing biases in children's judgments.

Children demonstrated different judgments regarding the two types of retribution, exclusion and punishment, when presented with different types of advantages: unintentional unfair, intentional unfair, and fair. Across measures and across contexts, participants in the in-group condition indicated that they were less likely to condemn the behavior of advantage creators than were participants in the out-group. While this pattern of results was expected and in line with previous research showing patterns of in-group preferences (McGuire et al., 2018; Rutland, 1999), the results presented here extended the previous literature by directly contrasting multiple advantage scenarios that varied by whether the act was intentional or unintentional, and fair or unfair. Further, the results of this study extend previous research on the impact of group

identity on children's willingness to condemn behaviors by suggesting that in-group members may be considering the impact of group functioning while making their decisions while out-group members are not. Moreover, with age, children were less likely to support exclusion across all three contexts, and this was shown for both the unintentionally unfair and fair advantage contexts. In addition, with age, children differentiated between exclusion and punishment, being less likely to support exclusion than punishment.

Research has shown that children develop an understanding of accidental transgressions and are more critical in their assessment of intentional transgressions than unintentional transgressions (D'Esterre et al., 2019). In the present study, children were given an affiliation with the relevant groups, moving them from the position of a third-person observer and putting them in a second-person perspective as has been done in previous research on group dynamics (McGuire et al., 2018). By inserting the participants directly in the presented contexts their pattern of responses shifted from what is typically seen in research on children's understanding of intentions. We found that participants who were in the out-group endorsed punishment and exclusion at equal rates for unintentional and intentional transgressions while participants in the in-group condition punished and excluded the unintentional transgressions less than the intentional transgressions. However, because the current study did not directly test children's mental state understanding these results should be treated cautiously until further research can be conducted.

Previous research has also shown that children are willing to punish unfair behavior even when they are unaffected third parties to the retributive behavior, that is, when they were not the recipient of the unfair treatment (McAuliffe et al., 2015). The current study extends this body of research by illuminating the relationship between children's relationship to the advantage creator

and their willingness to punish unfair behavior. These findings demonstrate that this relationship not only affects general willingness to engage in retribution but that it also influences what *type* of retribution children will endorse.

To date, no research has provided a systematic comparison of children's responses to unintentional, intentional, and fair advantages in competitive intergroup contexts. Having participants directly affiliated with the groups in these contexts added another level of novelty to the study and further enhances the real-world implications of these findings as children often find themselves firmly ingrained in the social contexts in which they are making these types of judgments. We now turn to a discussion of the findings for judgments about punishment, exclusion, and the direct comparison of both constructs for the role of group membership, advantage type, and age-related findings.

Punishment

The punishment measure utilized in the current study was selected as it allowed participants to reprimand a character for their actions without influencing the functioning of the groups. As such, it was expected that children in the in-group condition would be more likely to endorse punishment than exclusion. Within this experimental context, punishment was defined as having relatively minor and short-term implications, while exclusion was presented as more severe and with long-term consequences for the group. Thus, while we expected the in-group to endorse punishment less than the out-group, we also predicted that the differences may not be quite as pronounced as exclusion and that the in-group and out-group may look more similar in their punishment endorsement for certain contexts.

Broadly, the pattern of results obtained by this study revealed that participants in the in-group endorsed punishment less than those in the out-group, further reinforcing the findings of

previous research into in-group preferences (McGuire et al., 2018; Rutland, 1999). However, a closer examination reveals further nuances of the interactions between group identity and advantage contexts. Participants in the in-group condition endorsed punishment less often than those in the out-group condition overall, but within the intentional unfair advantage context no differences were found between these two groups. This lack of a significant difference was in line with our hypotheses, and it is supported by previous research showing that straightforward moral transgressions tend to be less prone to interpretation and children's moral judgments of these scenarios tend to be fairly uniform across social contexts (Smetana et al., 2014). Taken together this suggests that, while group identity does lead to differences in children's punishment decisions, this is primarily found in situations which are socio-morally complex and with room for interpretation. In the case of a clear transgression, group biases are not strong enough to influence children's decisions.

Investigations into the interaction between group identity and advantage context also revealed an interesting pattern: participants in the in-group condition endorsed punishment at different rates for each of the three advantage contexts, while the rates that out-group participants endorsed punishment did not differ between the unintentional and intentional unfair advantage contexts. This was, once again, in line with our predictions and supported by previous work suggesting that children are more sensitive to the mental states and intentions of in-group members than out-group members (Gönültaş et al., 2020). Previous work has also shown that children who successfully encode mental state information are likely to differentiate in their judgments of unintentional and intentional transgressions while those who fail to encode this mental state information respond similarly across these contexts (D'Esterre et al., 2019). Taken together it suggests that participants in the out-group condition may not put in the same effort to

encode the difference in intentions between these two contexts, and that this lack of intention encoding led them to endorse punishment equally for both transgressors. Alternatively, it is also possible that children in the outgroup are capable of recognizing intentions but weighed the importance of that information differently than did participants in the in-group condition. Future research utilizing an explicit measure of mental state understanding would be beneficial in disentangling these two alternate theories.

In addition to the impact of children's group identity on their punishment decisions, this study was also designed to assess age-related changes in children's punishment endorsement. We predicted that older children would be less likely to endorse punishment overall, but that this would be primarily driven by differences in contexts. However, children did not display significantly different rates of punishment endorsement as a function of their age across all three contexts. Upon a closer inspection of the three advantage contexts, it was discovered that older children punished significantly less in the two contexts that we had predicted. Further, there was a significant *increase* in older children's punishment endorsement for the intentional unfair advantage context. This pattern is in line with a large body of research showing that older children are better able to coordinate context-specific information in their moral evaluations than younger children (Smetana et al., 2014). It is interesting that older children were not only less likely to punish unjustly but were also more likely to punish justly in the case of a clear moral transgression.

Exclusion

As stated previously, the exclusion measure differed from the punishment measure in that we anticipated exclusion would be seen as having a significant, and negative, impact on the group functioning of the advantage creator's group. Given the previous research into children's

understanding of exclusion on group dynamics (e.g., Cooley & Killen, 2015), we predicted that

children's endorsement of exclusion would look different from their endorsement of punishment.

However, as no research to date has looked at the impact of group identity on children's

exclusion decisions, especially across multiple intentional contexts, the interaction of group and

intentions on children's exclusion decision was an open empirical question. Drawing on the

literature on children's understanding of group loyalty (Abrams et al., 2008; Mulvey, 2016;

Nesdale & Lawson, 2011), we hypothesized that participants in the in-group condition would be

less likely to endorse exclusion than would participants in the out-group condition.

Indeed, the findings on participants' endorsement of the exclusion option supported our

predictions: participants in the in-group condition endorsed exclusion less often than did

participants in the out-group condition. However, once again, looking more closely at the

interaction of group identity and advantage contexts revealed an interesting pattern of interplay

between these social factors. Participants in the in-group endorsed exclusion for the unintentional

advantage creator less often than the intentional advantage creator while participants in the out-

group condition did not differ in their rates of exclusion endorsement across these two advantage

contexts. This finding supports previous research demonstrating that children take intentions into

account when evaluating transgressions (D'Esterre et al., 2019; Margoni & Surian, 2017; Nobes

et al., 2017). However, unlike the punishment measure, participants in the in-group were

significantly less likely to endorse exclusion for the intentional unfair advantage creator. This

difference was counter to our expectations that children in the in-group and out-group condition

would not differentiate their evaluations of exclusion or punishment for the intentional unfair

condition. We take this different pattern of responses as evidence supporting our interpretation

that in-group participants are concerned with the negative impact of exclusion on group

functioning when making their decisions.

   While unexpected, this finding is in line with existing literature showing children's early

understanding of group loyalty (Abrams & Rutland, 2008; Nesdale & Lawson, 2011) and their

recognition of the impact that exclusion has on group functioning (Cooley & Killen, 2015) which

together may have been enough to deter the in-group participants from endorsing exclusion. This

pattern of results seems to suggest that, while group biases were not sufficient to influence

children's punishment endorsement, these same biases were enough to affect their exclusion

endorsement behavior. This possibility opens the door to several empirical questions, such as:

would in-group members endorse exclusion at the same rate as out-group members if the

transgression were more serious (e.g., physical harm) and would in-group and out-group

participants endorse exclusion at the same rate if the context were not competitive? These

questions remain a topic for future consideration and empirical study.

   We were also interested in investigating age-related differences in children's exclusion

endorsement. We had anticipated that older children would be more cognizant of the negative

effects of exclusion (Mulvey, 2016) and thus would be less likely to endorse the exclusion of

another individual. We found a significant main effect of age, such that older children were

significantly less likely to endorse exclusion than were younger children in general. This pattern

also held across each of the three advantage contexts which children viewed. While the main

effect of age was in line with our expectations, we did not predict that there would be differences

between old and young children in the intentional unfair advantage context. However, this

finding is in line with previous research showing that, as they grow older, children gain

experience of excluding and being excluded, developing an increased understanding of the

negative impact of exclusion (Abrams & Killen, 2014). Taken together these findings suggest that, as children grow older and gain increased knowledge about group dynamics, they are less likely to endorse exclusion across a wide range of contexts.

Comparisons between Punishment and Exclusion

While both punishment and exclusion are similar in that they are retributive responses to a perceived transgression, we expected that children would endorse the utilization of these strategies differently due to the different implications for group functioning that each response would entail. Specifically, it was expected that punishment would be more frequently endorsed than exclusion, and that participants in the in-group condition would show this pattern more strongly than out-group participants as exclusion would negatively impact the functioning of their own group. We also predicted that older children would be more likely to show different patterns in their punishment and exclusion endorsement than younger children would be due to older children's increased understanding of the negative impacts of exclusion making them hesitant to endorse such a strategy. Upon closer inspection, a strong and consistent pattern of differences in punishment endorsement and exclusion endorsement were found, although with three notable exceptions.

First, participants were more likely to endorse punishment than exclusion for the unintentional and intentional unfair advantage contexts, but levels of punishment and exclusion did not differ for the fair advantage context. Looking closer at the levels of punishment and exclusion endorsement for the fair advantage context shows that regardless of group membership or participant age, children were equally unlikely to endorse punishment or exclusion. This finding is not unexpected as the fair advantage scenario was included as a "baseline" situation where no transgression occurred and thus levels of both punishment and exclusion were

predicted to remain low for all conditions and ages. The pattern of results found indicates that this fair advantage context behaved as expected and that most participants seemed to be cognizant of the fact that even though one group was advantaged over another that this asymmetry was created legitimately and therefore was acceptable.

The second pattern observed when looking for differences between punishment and exclusion endorsement is that participants in the in-group condition were much more likely to endorse punishment than exclusion (in the unintentional unfair and intentional unfair advantage contexts), while participants in the out-group condition did not differ in their rates of punishment or exclusion across any of the three advantage contexts. Once again, this pattern was in line with our prediction as the participants in the in-group showed a hesitance to endorse exclusion in response to the transgressions, even after showing a readiness to endorse punishment. This supports one of the primary goals of this paper, and it provides evidence that children will consider the implications of exclusion on group functioning when it affects their own group, but that this consideration is not as relevant for children responding to the actions of an out-group member. To the best of our knowledge this pattern is a novel finding, but one with the potential to have far reaching implications on research looking at disparate responses to perceived transgressions. For example, this pattern suggests that children are more likely to endorse conflict resolution which is more severe if they view the transgressor as different from them, even if it leads to additional hardships for members of the out-group. If this pattern continues into adulthood it could be applied to issues such as differences in disciplinary action for ethnic majority and minority children, or to issues of racial biases and juror selection during criminal trials. Future research should explore this relationship between group membership and type of retribution. Extensive research has demonstrated in-group biases in adulthood (Dovidio et al.,

2015). Examining whether these biases persist in situations that vary in unintentional and intentional unfair advantage situations would provide more information regarding the contextual nature of bias.

Finally, it appears that older children were much more likely to punish than exclude in the unintentional unfair and intentional unfair advantage contexts while younger children were very similar in their rates of punishment and endorsement for each of these contexts. This pattern is in line with the broader literature as well as with our own expectations as older children are likely to have a firmer grasp on the implications of exclusion on group functioning as well as to have more experience with excluding and being excluded (Abrams & Killen, 2014). In turn, this appreciation for the implications of exclusion likely contributed to their reluctance to utilize exclusion in response to these transgressions.

The fact that the punishment and exclusion measures showed such differences between advantage contexts, between in-group and out-group members, and even in the pattern of use *within* participants in the in-group and out-group conditions further demonstrates the importance of including both measures and recognizing the meaningful distinction between these two measures of retributive attitudes. We interpret the differences in punishment and exclusion endorsement as reflective of the ways that punishment only affects an individual while exclusion influences their whole team, however it's also possible that children are responding to the perceived severity of punishment and exclusion. It is possible, and even likely, that these observed differences in retributive attitudes would look different for different types of moral responses, and that the patterns of these response types may change for different types of transgressions. Indeed, previous research has shown that children evaluate transgressions more negatively when the transgression is more serious or more visually salient (e.g., Weisberg &

Leslie, 2012). Therefore, future research into children's retributive attitudes, intergroup cognition, or intentions would likely benefit from the inclusion of diverse assessments and diverse transgressions in order to better understand these nuanced differences.

Conclusion

There were several novel findings in this study which have implications for research on children's understanding of group identity, intentions, fairness, and retributive attitudes. By investigating the intersection of intentions, intergroup attitudes, age, and retributive attitudes, these results were able to provide insight into the ways in which social factors can influence perceptions of reasonable responses to perceived transgressions. Importantly, this study showed that children endorsed punishment and exclusion differently based on the type of advantage created and based on the relationship between the group identity of the child and that of the target. Out-group members endorsed both punishment and exclusion at greater rates than in-group members for the unintentional unfair and fair advantage contexts, while the intentional unfair advantage contexts showed differences for exclusion but not for punishment. Older children, relative to their younger counterparts, also seemed to be better able to match their punishment decisions to the socio-moral context while simultaneously showing a reluctance to endorse exclusion regardless of the context provided.

In addition to presenting several novel findings, this line of research also suggests numerous future directions of inquiry that would be worth pursuing in the future. One such area of future research would be to determine whether children's punishment judgments differ significantly based on how strongly or weakly they identify with these novel groups. Further, it would be worthwhile to investigate possible social and emotional correlates with children's tendency to affiliate strongly with novel groups. Additionally, the group activity presented in the

current study was still capable of being performed by an individual or a smaller group, even though it would result in a greater strain on each individual member. It's possible that children's responses would differ in a scenario where the activity was impossible without a full group (e.g., a relay race) or that is made substantially more difficult with less members (e.g., soccer), and this type of distinction would be an interesting topic for future research.

Finally, the current study has limitations that could be further expanded upon in subsequent research. One such limitation of the current study was that we did not systematically test how participants conceptualized punishment and exclusion, and thus it's possible that children did not fully recognize the impact that exclusion has on group functioning that was intended with the exclusion measure we utilized. Future research on this topic may benefit from a clear example of what these types of responses would entail, as well as by gathering reasoning data from participants in order to gain greater insight into children's rationale for their choices. Additionally, while we measured children's understanding of the character's beliefs, we did not directly measure children's reports of intentions, and future research looking to investigate the connection between children's theory of mind abilities and their punishment and exclusion decisions would benefit from a more thorough theory of mind measure. Finally, for this study we decided to split our sample by age, in order to show differences between younger and older children and to highlight the hypothesized age where this shift typically occurs. However, it's also possible to investigate these questions with age as a continuous variable and this would be an important endeavor for future studies.

By demonstrating that different levels of punishment are endorsed for in-group and out-group members this study shows that, starting in childhood, biases are revealed when evaluating group advantages under different conditions of intentionality. Group membership and a sense of

group identity can be powerful tools which can encourage cooperation, mutual support, and

many other examples of prosocial behavior. However, in other contexts this same group identity

can lead to bias, discrimination, and many forms of prejudice. Therefore, it is important to

understand the social and contextual factors that cause either prosocial or antisocial behaviors to

be more likely and the implications of these behaviors across these contexts. By investigating

these issues within the scope of childhood we are able to gain a sense of how these attitudes

develop, with the hope that it will provide insight into the underlying mechanisms which

contribute to their developmental trajectory, which in turn would have the ability to provide a

greater sense of how and when to intervene before in-group favoritism and out-group bias

solidify into fully developed prejudicial and biased attitudes.

References

Abrams, D., Hogg, M., & Marques, J. M. (2005). A social psychological framework for understanding social inclusion and social exclusion. In D. Abrams, M. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion.* Psychology Press.

Abrams, D., & Killen, M. (2014). Social exclusion of children: Developmental origins of prejudice. *Journal of Social Issues, 70,* 1-11. https://doi.org/10.1111/josi.12043

Abrams, D., & Rutland, A. (2008). The development of subjective group dynamics. In S. R. Levy & M. Killen (Eds.) *Intergroup attitudes and relations in childhood through adulthood* (pp. 47-65). Oxford University Press.

Abrams, D., Rutland, A., Ferrel, J. M., & Pelletier, J. (2008). Children's judgments of disloyal and immoral peer behaviour: Subjective group dynamics in minimal intergroup contexts. *Child Development, 79,* 444-461. https://doi.org/10.1111/j.1467-8624.2007.01135.x

Bernhard, R. M., Martin, J. M., & Warneken, F. (2020). Why do children punish? Fair outcomes matter more than intent in children's second- and third-party punishment. *Journal of Experimental Child Psychology, 200,* 104909 DOI: 10.1016/j.jecp.2020.104909

Burkholder, A. R., Elenbaas, L., & Killen, M. (2020). Children's and adolescents' evaluations of intergroup exclusion in interracial and interwealth peer contexts. *Child Development, 91,* 512-527. https://doi.org/10.1111/cdev.13249

Cooley, S., & Killen, M. (2015). Children's evaluations of resource allocation in the context of group norms. *Developmental Psychology, 51,* 554-563. https://doi.org/10.1037/a0038796

D'Esterre, A. P., Rizzo, M.T., & Killen, M. (2019). Unintentional and intentional false statements: The role of morally-relevant theory of mind. *Journal of Experimental Child Psychology, 177,* 53-69. https://doi.org/10.1016/j.jecp.2018.07.013

D'Esterre, A., Woodward, B., & Killen, M. (2022). Children's group identity is related to

their assessment of fair and unfair advantages. *Journal of Experimental Child

Psychology, 214.* ID: 105292. Doi: 10.1016/j.jecp.2021.

Dovidio, J. F., Gaertner, S. L., & Saguy, T. (2015). Color-blindness and commonality: Included

but invisible? *American Behavioral Scientist, 59,* 1518-1538.

https://doi.org/10.1177/000276421558059

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using

G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research

Methods*, *41*, 1149–1160. https://doi.org/10.3758/brm.41.4.1149

Glidden, J., D'Esterre, A., & Killen, M. (2021). Morally-relevant theory of mind mediates

the relationship between group membership and moral judgments. *Cognitive

Development.* https://doi.org/10.1016/j.cogdev.2020.100976

Gönültaş, S., Selçuk, B., Slaughter, V., Hunter, J. A., & Ruffman, T. (2020). The capricious

nature of theory of mind: Does mental state understanding depend on the characteristics

of the target? *Child development*, *91*, 280-298. https://doi.org/10.1111/cdev.13223

Hardecker, S., Schmidt, M. F. H., Roden, M., & Tomasello, M. (2016). Young children's

behavioral and emotional responses to different social norm violations. *Journal of

Experimental Child Psychology, 150,* 364-379. https://doi.org/10.1016/j.jecp.2016.06.012

Hitti, A., & Killen, M. (2015). Expectations about ethnic peer group inclusivity: The role of

shared interests, group norms, and stereotypes. *Child Development, 86,* 1522-1537.

https://doi.org/10.1111/cdev.12393

Killen, M., Mulvey, K. L., & Hitti, A. (2013). Social exclusion in childhood: A developmental

    intergroup perspective. *Child Development, 84,* 772-790.

    https://doi.org/10.1111/cdev.12012

Killen, M., Mulvey, K. L., Richardson, C. B., Jampol, N., & Woodward, A. (2011). The

    "accidental transgressor": Morally-relevant theory of mind. *Cognition, 119,* 197-215.

    https://doi.org/10.1016/j.cognition.2011.01.006

Killen, M., & Rutland, A. (2011). *Children and social exclusion: Morality, prejudice, and group*

    identity. Wiley-Blackwell. https://doi.org/10.1002/9781444396317

Killen, M., Rutland, A., Abrams, D., Mulvey, K. L., & Hitti, A. (2013). Development of intra-

    and intergroup judgments in the context of moral and social-conventional norms. *Child*

    Development, 84, 1063-1080. https://doi.org/10.1111/cdev.12011

Lagattuta, K., & Weller, D. (2014). Interrelations between theory of mind and morality: A

    developmental perspective. In M. Killen & J. Smetana (Eds.), *Handbook of moral*

    development (pp. 385-408). Psychology Press.

Malti, T., Killen, M., & Gasser, L. (2012). Social judgments and emotion attributions about

    exclusion in Switzerland. *Child Development, 83,* 697-711.

    https://doi.org/10.1111/j.1467-8624.2011.01705.x

Margoni, F., & Surian, L. (2017). Children's intention-based moral judgments of helping agents.

    Cognitive Development, 41, 46-64. https://doi.org/10.1016/j.cogdev.2016.12.001

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

McGuire, L., Rizzo, M.T., Killen, M., & Rutland, A. (2018). The development of intergroup

    resource allocation: The role of cooperative and competitive in-group norms.

    Developmental Psychology, *54*, 1499-1506. https://doi.org/10.1037.dev0000535

McGuire, L., & Rutland, A. (2020). Children and adolescents coordinate group and moral

    concerns within different goal contexts when allocating resources. *British Journal of*

    Developmental Psychology, 38(4), 529–542. https://doi.org/10.1111/bjdp.12342

Mulvey, K. L. (2016). Evaluations of moral and conventional intergroup transgressions. *British*

    Journal of Developmental Psychology, 34, 489-501. https://doi.org/10.1111/bjdp.12145

Nesdale, D. (2004). Social identity processes and children's ethnic prejudice. In M. Bennett & F.

    Sani (Eds.), *The development of the social self* (pp. 219-245). Psychology Press.

    https://doi.org/10.4324/9780203391099_chapter_8

Nesdale, D., Durkin, K., Maass, A., & Griffiths, J. (2005). Group status, outgroup ethnicity and

    children's ethnic attitudes. *Journal of Applied Developmental Psychology, 25,* 237-251.

    https://doi.org/10.1016/j.appdev.2004.02.005

Nesdale, D., & Lawson, M. J. (2011). Social groups and children's intergroup attitudes: Can

    school norms moderate the effects of social group norms? *Child Development, 82,* 1594

    -1606. https://doi.org/10.1111/j.1467-8624.2011.01637.x

Nobes, G., Panagiotaki, G., & Enelhardt, P. E. (2017). The development of intention-based

    morality: The influence of intention salience and recency, negligence, and outcome on

    children's and adults' judgments. *Developmental Psychology, 53,* 1895-1911.

    https://doi.org/10.1037/dev0000380

Oberfeld, D., & Franke, T. (2012). Evaluating the robustness of repeated measures analyses: The case of small sample sizes and nonnormal data. *Behavior Research Methods, 45,* 792-812. https://doi.org/10.3758/s13428-012-0281-2

Rizzo, M. T., Cooley, S., Elenbaas, L., & Killen, M. (2018). Young children's inclusion decisions in moral and social–conventional group norm contexts. *Journal of Experimental Child Psychology, 165,* 19–36. https://doi.org/10.1016/j.jecp.2017.05.006

Rutland, A. (1999). The development of national prejudice, in-group favouritism and self-stereotypes in British children. *British Journal of Social Psychology, 38,* 55-70. https://doi.org/10.1348/014466699164031

Schmidt, M. F. H., & Tomasello, M. (2012). Young children enforce social norms. *Current Directions in Psychological Science, 21,* 232.236. https://doi.org/10.1177/0963721412448659

Sierksma, J., Spaltman, M., & Lansu, T. A. M. (2019). Children tell more prosocial lies in favor of in-group than out-group peers. *Developmental Psychology, 55,* 1428-1439. https://doi.org/10.1037/dev0000721

Smetana, J. G., Jambon, M., & Ball, C. (2014). The social domain approach to children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp 23-45). Psychology Press.

Tomasello, M. (2014). The ultra-social animal. *European Journal of Social Psychology, 44,* 187-194. https://doi.org/10.1002/ejsp.2015

Weisberg, D. S., & Leslie, A. M. (2012). The role of victims' emotions in preschoolers' moral judgments. *Review of Philosophy and Psychology, 3,* 439-455. https://doi.org/10.1007/s13164-012-0101-8

Yazdi, H., Barner, D., & Heyman, G. D. (2020). Children's intergroup attitudes: Insights from

Iran. *Child Development, 91,* 1733-1744. https://doi.org/10.1111/cdev.13363