Children's Judgments of and Reasoning about People with Disabilities who Produce Norm Violations

Nicolette G. Granata*, Megan Wiebe, Jonathan D. Lane

Vanderbilt University, Peabody College, Dept. of Psychology & Human Development

*Corresponding Author: nicolette.g.granata@vanderbilt.edu

This is a pre-press draft of a manuscript accepted for publication at the *Journal of Experimental Child Psychology.*

Abstract

People with disabilities may behave in non-normative ways because they cannot act otherwise. This study explored whether U.S. children (3.00-8.99 years; $N = 105$) differ in their evaluations of people who commit norm violations when those persons have perceptual or physical disabilities. Across twelve scenarios, children were asked to explain different characters' non-normative behaviors and to evaluate each character's naughtiness. Characters were either typically-developing, had a physical disability, or had a hearing disability. Disabilities were described to participants but were not visually depicted. Across moral and conventional norm violations, children 4.5-years and older judged characters with disabilities as less naughty than characters without disabilities, whereas younger children (3- to 4-years) judged all characters as equally naughty. Children's explanations for characters' non-normative behaviors (acknowledging characters' physical/auditory limitations, and inferring negative attributes) significantly predicted their naughtiness judgments; this was true for participants across the sampled age range. Thus, preschool children demonstrated flexibility in their moral judgments across a variety of everyday behavioral violations, tempering their negative evaluations of persons who committed non-normative behaviors when those persons had unseen disabilities that could reasonably account for their actions. Parents and teachers may be able to build upon these early moral intuitions to foster greater acceptance of persons with disabilities.

Keywords: social cognition, disability concepts, moral evaluations

Children's Judgments of and Reasoning about People with Disabilities who Produce Norm Violations

Approximately 1 in 6 children in the United States have been diagnosed with a developmental disability, and most attend school with typically-developing (TD) classmates. These children are victimized much more often than their TD peers--with victimization rates approaching 50%--and this has serious implications for their physical and psychological well-being (Koller et al., 2018; Rose & Gage, 2016). Even within inclusive settings, children with disabilities often develop fewer and lower quality relationships with peers than TD children (Diamond et al., 1993; Koller & San Juan, 2015). One explanation for why children with disabilities may experience lower levels of social inclusion is that their integration depends, in part, on the beliefs and expectations of their peers (Diamond et al., 2008; Diamond & Kensinger, 2002; Koller et al., 2018). Though limited, research on children's judgments of the behaviors of persons with disabilities suggests that there may be a critical set of behavioral cues that guide children in preferentially seeking TD classmates as playmates (Diamond et al., 1993; Koller et al., 2018).

In the current study, we examine how children in the United States evaluate other children with disabilities who produce non-normative behaviors. Often, individuals with disabilities must engage in non-normative behaviors to pursue everyday goals, and these behaviors may (unintentionally) inconvenience or upset others. For example, someone with a physical (walking) disability might move slowly on a bustling city sidewalk, and someone with a perceptual (hearing) disability might have to stare at a speaker's lips (instead of their eyes) to comprehend their message. As well, individuals with disabilities may be *unable* to engage in behaviors that are otherwise normative and socially expected; a person with a physical disability

may be unable to help a person who has dropped their groceries, and a person with a perceptual (hearing) disability may be unable to hear someone say hello in passing and thus may not say hello in return.

In all of these cases, the person producing the behavior *may* be evaluated negatively. In the current study, we ask how children evaluate people with physical (walking) and perceptual (hearing) disabilities, as well as people with no disabilities, after they produce non-normative behaviors, and we explore how these evaluations differ between children ages 3 to 8 years. As well, we explore children's *reasoning* behind their evaluations. The results inform literatures on children's moral reasoning, and their concepts of physical and perceptual ability. In what follows, we discuss past research on children's concepts of disability and their reasoning about people who produce non-normative behaviors. We then outline the current study.

Children's Understanding of and Evaluations of Physical and Perceptual Disabilities

By 3-4 years of age, children often have an *initial* understanding of certain perceptual abilities and constraints. For example, young children appreciate that people may vary in their ability to see or hear, and that the functioning of these senses may limit their ability to obtain information through those senses (Lane et al., 2010; Lane et al., 2016; Moll, et al., 2014; Pillow, 1989; Williamson et al., 2013). They also understand that people's visual and auditory perception can be limited by distance between the perceiver and the stimulus and by the magnitude of the stimulus--e.g., someone may not hear something if they are not near the sound or if the sound is soft (Lane et al., 2016; Moll et al., 2014).

To what extent do children recognize and understand the implications of different types of *disability*, and does this depend on how the disabilities are portrayed? Much of the work on this topic has evaluated U.S. preschoolers, often as a single group ranging from 3-to-5 or 3-to-6

years (e.g., Diamond & Hestenes, 1996; Diamond, Hong & Tu, 2008; Diamond & Hong, 2010; Huckstadt & Shutts, 2014). This research has yielded important insights into children's concepts of disability, but often does not explore *developmental patterns* in these concepts. In an influential study, Diamond and Hestenes (1996) showed children ages 3-6 years photographs of other children using disability equipment (i.e., a wheelchair, hearing aid, eye patch). Children were asked open-ended questions about the characters and their abilities, including the question "Tell me about this girl/boy." As a group (age-related differences were not analyzed), most children (73%) spontaneously mentioned a disability in response to this question for the character with the *physical disability* (in a wheelchair). Other work by Diamond and colleagues has demonstrated that children in this age range also appreciate that particular activities are more challenging or unachievable for those with physical disabilities (Diamond et al., 2008; Diamond & Hong, 2010). For example, they appreciate that a child who uses a wheelchair will have more difficulty dancing than completing a puzzle.

However, fewer than half (41%) of the children interviewed by Diamond and Hestenes (1996) commented on the disability of the character with a visual impairment (illustrated by an eye patch), and very few (11%) did so for the character with a hearing impairment (illustrated by a hearing aid). Because perceptual disabilities are often *not* accompanied by salient physical markers (e.g., large medical equipment), children (and adults) may be less aware that others possess these disabilities and may be less understanding of the impact of those disabilities on people's everyday lives and behavior. Thus, preschoolers better recognize disabilities with more obvious physical markers and better understand the limitations imposed by those disabilities.

Large, salient physical markers such as wheelchairs or crutches may help children (and adults) recognize and remember that an individual is disabled. Unfortunately, these same

markers may also negatively skew children's impressions of persons with disabilities (e.g., Diamond et al., 2008; Huckstadt & Shutts, 2014). In a Visible Preference task, 3- to 5-year-olds were presented two photographs side-by-side: a child in a wheelchair and a child with no wheelchair were matched for attractiveness, age, hair color, race, and gender (Huckstadt & Shutts, 2014). Participants preferred to befriend the child who appeared with *no* wheelchair. This suggests that the presence of disability equipment (the wheelchair) alone influenced children's social preferences. Thus, for studies exploring children's evaluations of the *actions* performed by persons with disabilities—as we do in the current experiment—internal validity may be increased by controlling for the presence of disability equipment, so that children's attitudes about the *equipment* do not affect their responses. There are also benefits to external validity when controlling for the presence of disability equipment. Namely, a disability is often an aspect of someone that persists even when their equipment is not visible—someone who is blind is still blind even when they remove their dark glasses; someone who cannot walk is still unable to walk even when they are not sitting in a wheelchair—the disability is an aspect of the person that endures across contexts and time even when there is no visual marker. To examine children's concepts or judgments of persons with *enduring* disabilities, it is useful to not depict medical equipment alongside those persons (who, in reality, may not always utilize equipment).

Thus, in the current study, all persons that children are asked to evaluate are depicted without disability equipment; instead, their disabilities are *described in detail verbally* (i.e., children are given multiple examples of how characters' physical/perceptual functioning could be affected), and they have no apparent physical differences other than features that vary across typically and atypically developing persons (e.g., hair color, clothing color). Prior research has demonstrated that preschoolers can reason about persons' *exceptional* perceptual abilities when

provided detailed descriptions of those abilities. For example, Lane et al. (2010) found that preschoolers understood that characters with exceptional vision would have greater knowledge of certain information. One of these characters, a cat (depicted as an ordinary cat, with yellow eyes) was described as having, "special eyes that let him see in the dark." Children as young as 4-years reported that the cat would know the contents of darkened containers whereas people with ordinary vision would not know the contents. Thus, preschoolers correctly drew inferences about what others could see based primarily on a verbal description of their perceptual abilities. Using verbal descriptions in the current study has the added benefit of revealing how well children can learn about and draw inferences from verbal descriptions of people's *unseen disabilities*.

Evaluating the Non-normative Behavior of Persons with and without Disabilities

Atypical behaviors are commonly produced by people with certain physical disabilities because of salient differences in how they perform everyday activities (such as eating food or moving about). As well, persons with intellectual and perceptual disabilities may produce novel behaviors—e.g., hand flapping, ticks. There are also many situations in which a person with a disability simply *cannot* produce particular normative behaviors, even when others expect them to, because of their disability. How do children evaluate people who produce non-normative behaviors?

Moral norms involve potentially universal concerns, such as avoiding harm, upholding justice, and respecting individual rights (Tisak & Turiel, 1988). Conventional norms are conceptualized as specific to individual social systems; they build upon arbitrary, mutual expectations for behavior (Ball et al., 2016). For example, in one society, but not another, it may be inappropriate to laugh during a funeral. Children in the U.S. and Germany, as young as 3-4

years, judge both conventional violations and moral violations as wrong (Dahl & Kim, 2014; Göckeritz et al., 2014). But children often conceptualize moral norms as more rigid across contexts (e.g., hurting someone is *always* wrong), and often judge moral violations as more punishable (Hardecker et al., 2016). We are interested in evaluating how *robustly* children take into account people's disabilities when judging their naughtiness, and so we begin by including violations that would traditionally be considered moral as well as violations that would traditionally be considered conventional among participants in the United States.

Dating back to Piaget's foundational work on children's moral development (Piaget, 1965), researchers have repeatedly found that, with development, children increasingly account for norm-violators' intent when evaluating their actions. Whereas 4-5-year-olds' evaluations are primarily influenced by the presence and magnitude of violations (violations are bad, and more extreme violations--e.g., physically injuring others--are especially bad), older children (7-8 years) often consider additional factors, such as violators' motives and intentions (e.g., Cushman et al., 2013). Recent work demonstrates that the transition to considering actors' intentions may occur as early as 5-years (Killen et al., 2011) for children in the United States. An issue that is especially relevant to the current study is whether and when children consider factors that are *outside actors' control* (i.e., obstacles or limitations) when interpreting and judging their behavior.

In one of the only studies to directly explore how children judge people with disabilities who commit behavioral violations, children 3-5 years were told that characters (one visually impaired (VI), the rest typically-developing (TD)) were going to play a game where the goal was to *look through* a viewfinder to see what was inside a box (Huckstadt & Shutts, 2014). The VI character wore opaque glasses and was introduced with, "This kid's eyes don't work, so he can't

see anything." The experimenter explained that this character could *not* play the usual way because of their disability. Each character then played the game: one TD character played the game according to the norm (by *looking through* the viewfinder), and another TD character violated the norm as did the VI character (by *reaching into* the viewfinder). Using a 3-point scale ("Really OK", "Sort of OK", or "Really not OK"), children rated the norm-violating actions of the VI character and the TD character as *equally naughty*; both were judged naughtier than the other TD character's norm-abiding action. These findings may suggest that preschoolers evaluated behaviors based exclusively on whether a violation was committed, regardless of actors' abilities. The authors conclude that these children, "were not forgiving of norm violations committed by individuals with disabilities" (p. 109). Importantly, these data were not analyzed to identify potential age-related differences between 3- and 5-years; thus, it is unclear if these results are typical of children across the entire preschool period or if these results might mask early developmental patterns.

Our hypotheses, regarding children's evaluations of persons who produce non-normative behaviors, are informed by work on children's concepts of abilities and disabilities (reviewed earlier; Diamond & Hestenes, 1996; Diamond et al., 2008; Lane et al., 2010; Williamson et al., 2013) and work on children's evaluations of actors' non-normative behaviors based on actors' intentions and extenuating circumstances (Cushman et al., 2013; Killen et al., 2011; Piaget, 1965). We predict that children 3-4 years old will evaluate persons with and without disabilities as *equally* naughty after their non-normative behavior, and that by at least 7-8 years of age, children will evaluate persons with disabilities as *less* naughty than persons without disabilities after their non-normative behavior. As well, given prior work on the development of disability concepts, we tentatively expect to see differences in judgments of characters with physical versus

perceptual disabilities--because young children recognize and understand physical disabilities better than perceptual disabilities (Diamond & Hestenes, 1996; Diamond et al., 2008), young children may be more lenient towards characters with physical disabilities, resulting in somewhat lower naughtiness ratings for these characters.

Previous work does not clearly inform hypotheses about developmental shifts in children's evaluations of persons with disabilities between 3 and 6 years; the most relevant work has either not included 5- and 6-year-olds or has combined data from children ranging from 3 to 6 years (e.g., Diamond & Hestenes, 1996; Diamond et al., 2008; Diamond & Hong, 2010; Huckstadt & Shutts, 2014). However, important developments take place in children's understanding of auditory perception (e.g., Lane et al., 2010; 2016) and in their general moral reasoning (Killen et al., 2011) during this period—developments that might have a bearing on how children evaluate persons with disabilities. For example, recent work demonstrates that 5-year-olds judge unintentional non-normative behavior as less naughty than intentional non-normative behavior (e.g., Cushman et al., 2013; Killen et al., 2011). Thus, we also included children ranging from 5- to 6-years in age, so that we could identify potential shifts in children's concepts of the implications of disabilities on non-normative behavior *throughout* early childhood.

Finally, we have no strong hypotheses about whether children's accounting for actors' abilities in their naughtiness judgments might vary based on the specific type of behavioral violation. Thus, for exploratory purposes and to evaluate the robustness of our findings, we include a variety of violations, including some violations that might be construed as "moral" and others that might be construed as "conventional" based on past work with children from the U.S. and Europe.

Reasoning about Non-Normative Behaviors

In addition to evaluating whether children more often excuse (i.e., less harshly evaluate) persons with disabilities who produce non-normative behaviors, we also examine children's explanations for *why* these persons produce aberrant behaviors. Crucially, we then evaluate how such explanations relate to children's judgments of persons' naughtiness. We are particularly interested in whether children account for non-normative behavior in terms of people's disabilities (e.g., that a person did not help someone in need because they could not physically reach the person), and whether they infer that non-normative behaviors reflect people's negative traits or motives (e.g., that a person did not help someone because they are mean). Given early developments in children's understanding of the links between people's perception, knowledge, and behavior (Lane et al., 2010; Lane et al., 2016; Moll et al., 2014; Pillow, 1989; Williamson et al., 2013), we anticipate a general age-related *increase* in children's inferences that non-normative behaviors are a function of characters' disabilities (for the characters that indeed have disabilities). Because of developmental shifts in children's moral judgments (Cushman et al., 2013; Killen et al., 2011; Piaget, 1965) and shifts in children's tendency to infer hostile intentions (e.g., Dodge, 1980; Dodge et al., 1986; Orobio De Castro et al., 2002; described in more detail below), we anticipate an age-related *decrease* in children's attributions of non-normative behavior to characters' negative traits or intentions (for all characters, including those with and without disabilities).

More critically, we anticipate that children who more often cite characters' disabilities when explaining their non-normative behavior (i.e., who appreciate the role of characters' disabilities in their non-normative behavior) will judge those characters as less naughty. Conversely, we anticipate that children who more often infer that non-normative behavior

reflects characters' negative traits or motives will judge those characters as *more* naughty. This hypothesis was based on research on the side-effect-effect (Leslie et al., 2006) and children's evaluations of ambiguous social circumstances with negative outcomes (e.g., Dodge, 1980; Dodge et al., 1986; Orobio De Castro et al., 2002)—individuals (including children) who more often interpret behavior and outcomes of behavior as negative are more prone to infer that actors have negative qualities or negative intentions. Characters who are attributed these negative qualities are deemed naughtier. We have no strong predictions about *when* specifically in development these relations between children's reasoning and their naughtiness judgments will emerge, but our wide age range (3-8 years) allows us to address this question.

The Current Study

The current study was designed to address the following research questions: (1) How do children's judgments of others' naughtiness compare when violations are committed by persons with disabilities versus persons with no disabilities? (2) To the extent that children's naughtiness judgments are moderated by persons' disabilities, is this tendency robust across moral *and* conventional violations? (3) How do children *explain* the non-normative behavior of persons with and without disabilities? and (4) How do children's explanations of that behavior (particularly, their acknowledging characters' disabilities and their negative attributions) predict children's naughtiness judgments?

Prior studies exploring whether children account for actors' abilities and constraints in their reactions to actors' non-normative behavior have used experimenter-created games and study-specific norms and violations (Huckstadt & Shutts, 2014; Josephs et al., 2016). For example, Josephs et al. (2016) taught a puppet how to sort marbles using a novel apparatus; the puppet's hands were either free or tied (restricting use of its arms) and the puppet played with the

apparatus in a way that violated the researcher's directions. Preschoolers (averaging 47-months in age) accounted for the physical constraint, protesting or correcting violations *less* when the puppet was constrained. This might suggest that 4-year-olds evaluated the non-constrained puppet as more naughty than the constrained puppet, but that remains unclear, as children were not asked to evaluate the puppet (one may attempt to correct another's behavior—as a teacher might—without judging that person as bad). Importantly, because the same researcher who *taught* children about the norms was present when the norms were violated, children's behaviors might have reflected their *desire to appease and conform* to the researcher's expectations. In the current study, we ask children to evaluate people who violate norms that children are likely already familiar with, thus we do not need to teach children about the norms.

Based on personal classroom observations and detailed discussions with teachers, we identified scenarios that are likely to occur in everyday life for students with disabilities in the United States. We believe that this method increases the external validity of the study and reduces demand characteristics. Specifically, participants reasoned about characters who either did not help a fallen classmate, talked too loud in the classroom, rejected a classmate's request to play, or did not complete a classroom activity. These particular violations were chosen because each one could conceivably be committed by someone as a function of a physical disability or as a function of a hearing disability. For example, when a classmate falls down and yells for help, a plausible reason that a person with a physical disability might not help is because they cannot *walk* over to help, and a plausible reason that a person with a hearing disability might not help is because they cannot *hear* the classmate yell for help.

Method

Participants

Children ages 3.00-8.99 years (N = 105, 53 girls) were recruited from a medium-sized city in the Southeastern United States, as part of a larger study exploring children's concepts of persons with disabilities. An additional four participants were interviewed, but their data were excluded because they failed to correctly answer four or more of the eight memory-check questions (described later). Of the 105 participants, 49 were recruited by calling parents of children in the target age-range (3-8 years), who were living in the greater metropolitan area (using contact information from State birth records). Children whose parents expressed interest and consented to participate completed the study in a quiet room in our lab on the University's campus. The remaining 56 participants were recruited by distributing informed consent documents to a local private school. Children whose parents offered consent were tested in a quiet room at their school, away from distractions.

Our goal was to recruit a sample large enough to fulfill the requirements of an a priori power analysis (using G*Power 3.1; Faul, Erdfelder, Lang, & Buchner, 2007) for ANOVAs that included age group (3 levels, between-subjects), character disability type (3 levels, within-subjects), and violation type (2 levels, within-subjects). To examine age-related differences, we planned to analyze three age groups: 3.00-4.99 years (M_{age} = 4.08*;* n =30; 15 girls; 15 boys), 5.00-6.99 years (M_{age} = 5.98*;* n = 43; 18 girls; 25 boys), and 7.00-8.99 year (n = 32; M_{age} =7.97*;* 20 girls; 12 boys). The power analysis determined that we required a minimum of 93 participants to detect medium effect sizes ($fs \geq .25$; Cohen, 1992) with statistical power $\geq .80$ and $\alpha = .05$. We exceeded this recruitment goal because our consent rate was greater than expected, and we interviewed all participants whose parents provided consent.

To help characterize our sample, we asked parents to complete a voluntary questionnaire, for which they reported family demographics. Most of the participants (86.7%; n = 91) were identified by their parents as "White/Caucasian", followed by 6.7% (n = 7) as "Asian/Asian American", 5.7% (n = 6) as "Black/African American", 2.9% (n = 3) as "Hispanic or Latino", and 1% (n = 1) as "Other". These categories were not mutually exclusive; parents could select more than one. Of these parents, 32.4% (n = 34) reported having a Bachelor's degree, 32.4% (n = 34) reported having a Master's degree, 28.6% (n = 30) reported earning a Doctorate, and 4.8 % (n = 5) reported having completed some college. Two parents did not report their education level. On these questionnaires, parents also reported on their children's exposure to persons or media characters with disabilities—these descriptive data are presented in Supplementary Materials, Table S1 (evaluating associations among these variables was beyond the scope of our research questions, so these data are not considered further). Participant recruitment, parent consent, child assent, and all study procedures were approved by [*blinded for review*] University's Institutional Review Board (*blinded for review*).

Materials

Materials included laminated vector graphics (approximately 1.5 x 2.5 inches) depicting seated characters (18 girls, 18 boys) with differing physical features (e.g., different hair color, hair style, eye color, clothing color, skin tone). All characters were seated, to avoid visually distinguishing the characters with physical disabilities (who must be seated) from the other characters (who could conceivably stand). Additionally, four laminated graphics depicted a fallen girl, and four graphics depicted a fallen boy (approximately 2.5 x 1.5 inches). Two scenes (one of a classroom and one of an outdoor playground) were printed on 8.5 x 11-inch, laminated paper. Appendix A depicts the two scenes and examples of how graphics were placed on the

scenes for the Behavior Violation Scenarios (described later). Study sessions were recorded via a small audio recorder (if parents consented to audio recording), so that children's responses could later be transcribed.

Procedure

Introduction to Disabilities. Before beginning the study, each child spent several minutes building rapport with the experimenter (E). Once the child was comfortable, E directed them to sit in a chair on E's left, so that E and the child were seated on the same side of a table. The study began with E randomly selecting two character graphics (matched to the participant's gender), one-at-a-time. One character was described as having a physical (walking) disability: "This boy's/girl's legs don't work. They can't get out of their chair and move around if they want to. They can't run around the playground. They can't walk to the front of the classroom to ask the teacher questions if they need help." E asked comprehension-check questions about their disability (e.g., "So what part of this boy's/girl's body doesn't work?"; "Can this boy/girl walk?"), and either affirmed children's answers (e.g., "Yeah, their legs don't work."; "Yeah, because their legs don't work."), or corrected their answers (e.g., "Actually, their *legs* don't work."; "Actually, they *can't* walk, because their legs don't work."). The other character was described as having a perceptual (hearing) disability: "This boy's/girl's ears don't work. They can't hear if a firetruck is coming down the street. They can't hear the school bell at the end of the day. They can't hear their friends yelling on the playground." Children were asked similar comprehension-check questions about this character, and their answers were either affirmed or corrected accordingly. These questions ensured that all children had some exposure to information about physical and perceptual disabilities before completing the focal part of the study, on children's inferences about and evaluations of these persons. Overall, 85.7% of

participants correctly answered both introductory questions about physical disability and 90.5% correctly answered both introductory questions about perceptual disability. The order in which the two characters (and disabilities) were presented was counterbalanced across participants within each age group.

Behavior Violation Scenarios. Children were presented four scenarios--two moral violations and two conventional violations--in orders counterbalanced across participants within each age group. The four scenarios and corresponding graphics are presented in Appendix A. For each scenario, E placed a scene graphic--a classroom scene or a playground scene--in front of the participant. For each *set* of scenarios (a set includes one moral violation and one conventional violation), E randomly selected graphics of seated characters (three or four, depending upon the scenario), along with other graphics that were relevant to that scenario. All graphics were placed on the scene, with the three main characters placed from left to right on the left side of the scene. For each scenario, E described the context. For example, in one scenario (a moral violation), E narrated: "These three kids [*point to scene*] are in the same classroom…Look! Another boy/girl [*point to fallen character*] trips and falls down. He/she screams for someone to help." Then, for each of the three seated characters, in turn, participants were told the character's name, their abilities or disabilities, and their non-normative behavior. For example, the first character (who always had no disability), was introduced as follows: "This is [*name*]. [*Name*] can walk and he/she can hear. [*Name*] doesn't help the boy/girl who fell [*point to fallen character*]." To direct the participant's focus to that specific character, E moved their graphic beside the scene, away from the other characters, and closer to the participant. E asked children to *explain* the character's behavior, "Why did [*name*] not help the boy/girl who fell?", and then (using language similar to that used in prior work; Baird & Astington, 2004; Cushman et al., 2013; Hebble, 1971;

Hewitt, 1975; Piaget, 1965/1932) to *evaluate* the character's naughtiness; "Is [*name*] naughty for not helping?"; if the child responded "Yes", they were also asked: "So you think [*name*] is naughty. Is he/she a little naughty, or very naughty?" After these questions, E moved the character's graphic back onto the scene.

The same procedure was repeated for the characters with physical and perceptual disabilities, who were presented either second or third (in counterbalanced order across participants within each age group). For example, for the character with a physical disability, E provided a description of the disability, similar to that given during the *Introduction to Disabilities* portion of the study; e.g., "Remember when we talked about boys/girls whose legs don't work? [*Name*] is one of those boys/girls. [*Name*] can hear, but he/she can*not* walk." To ensure that children were paying attention and remembered this fact about the character, they were asked, "So what part of [*name*]'s body doesn't work?" Across all four scenarios, children were asked eight of these memory-check questions. Correct answers included "legs" or "ears," respectively. Other responses ("feet", "knees", or "I don't know") were recorded as incorrect. Participants' memory performance was near ceiling, on average, having correctly answered 7.44 of the 8 questions (as noted earlier, four participants were excluded from analysis because they correctly answered fewer than half of these questions). Children's answer to each question was either affirmed or corrected, accordingly. Then, E asked children to *explain* the character's behavior (e.g., "Why did [character name] not help the boy/girl who fell?") and to *evaluate* their naughtiness. This same procedure was repeated for the three remaining scenarios (see Appendix A). The order in which the four scenarios were presented was counterbalanced across

participants within each age group.

The current protocol was part of a larger study on children's concepts of disabilities; the protocol described above was delivered either at the very beginning (n = 51) or at the end (n = 54) of the larger study session (with order counterbalanced across participants within each age group). The other tasks in the study session, which are unrelated to the current research questions, assessed children's attributions of knowledge and beliefs to *different* characters (with different appearances and different names) who had disabilities. The "Introduction to Disabilities" portion of the protocol, described above, always came at the beginning of the larger protocol. The entire study session lasted approximately 20 minutes (with the tasks reported herein lasting approximately 10-12 minutes), after which participants chose a small toy as a gift.

Scoring

Naughtiness ratings. Responses to the questions about whether each character is naughty were assigned scores of '0' for "no", '1' for "yes" followed by "a little naughty", and '2' for "yes" followed by "very naughty". Six *naughtiness* composite scores were computed by averaging across responses for each of the six pairs of violation-ability scenarios: No Disability-Conventional; No Disability-Moral; Perceptual Disability-Conventional; Perceptual Disability-Moral; Physical Disability-Conventional; Physical Disability-Moral. Scores for each composite could range from 0-2.

Explanations. Responses to the open-ended questions about why each character committed each violation were coded for use of eight explanation categories: *Physical or Auditory Limitation, Unaware, Conflicting Desire, Negative Motive, Negative Person/Action, Apathy, Diffusion of Responsibility,* and *Other* (see Table 1). For each reasoning category, responses were assigned scores of '0' if children did not use that category of reasoning, or '1' if

children did. Each response could be coded into multiple categories. Responses were transcribed in a separate file and categorized by two coders (who were blind to participants' ages and to characters' abilities). Coders achieved 99.33% inter-rater reliability across approximately 20% of the data (2384 of 2400 codes matched). Inter-rater reliability by category was 99.6% (299/300) for *Physical or Auditory Limitation,* 99.33% (298/300) for *Unaware,* 98.66% (296/300) for *Conflicting Desire (Preference),* 99% (297/300) for *Negative Motive,* 99.33% (298/300) for *Negative Person/Action,* 99.33% (298/300) for *Apathy,* 100% (300/300) for *Diffusion of Responsibility,* and 99.33% (298/300) for *Other.* Having achieved high inter-rater reliability, one coder categorized the remaining 80% of the data. Ultimately, because of conceptual similarities, we combined the categories *Negative Motive* and *Negative Person/Action* into one category (*Negative Attributions*), resulting in seven categories. A total of 21 composite scores were computed, each representing how often children used a specific form of reasoning for a specific character (the character with no disability, a physical disability, or a perceptual disability), across the four scenarios. Scores for each composite could range from 0-100%.

Results

In preliminary analyses, we checked whether findings differed depending upon when children were presented these non-normative behavior scenarios—either at the beginning or the end of the study session. Those statistical models were similar to those reported herein, but also included an 'order' variable, revealing no significant main effects or interaction effects involving order, for any of the following analyses. Thus, presentation order is not considered further.

Naughtiness Ratings

Our first set of research questions concern whether children's evaluations of persons who committed non-normative behaviors differ depending upon persons' abilities, and how those

evaluations compare across children ranging from 3- to 8-years in age. As well, we were interested in testing the robustness of these patterns, by exploring whether similar patterns emerge across different types of behavioral violations—moral and conventional. The two moral violations were "Not Helping Fallen Classmate" and "Rejecting a Classmate's Request to Play." The two conventional violations were "Talking too Loud in Class" and "Not Completing Classroom Activity" (see Appendix A).

A 3 (Participant's Age: 3-4, 5-6, 7-8 years) X 3 (Character's Disability: None (Typical), Physical disability, Perceptual disability) X 2 (Violation Type: Moral vs. Conventional) ANOVA revealed significant effects of participants' Age ($F(2, 102) = 15.96$, $p < .001$, $\eta^2_p = .24$) and Violation Type ($F(1, 102) = 13.31$, $p < .001$, $\eta^2_p = .12$) on children's Naughtiness ratings. As expected, children in the Youngest age group (3-4 years) judged characters as naughtier ($M = 1.24$, $SE = .08$) than children in the Middle and Oldest age groups (5-6 years: $M = .78$, $SE = .07$, $p < .001$; 7-8 years: $M = .60$, $SE = .08$, $p < .001$). There was no significant difference in naughtiness scores between the Middle and Oldest age groups, $p = .10$. Children judged the characters who committed conventional violations as naughtier ($M = .94$, $SE = .05$) than those who committed moral violations ($M = .80$, $SE = .05$), $p < .001$.

More critically, this analysis revealed a significant effect of Character's Disability ($F(2, 101) = 73.19$, $p < .001$, $\eta^2_p = .59$), which was subsumed under an interaction of Age X Character's Disability ($F(4, 204) = 14.93$, $p < .001$, $\eta^2_p = .23$). As depicted in Figure 1, the Youngest children (3-4 years) judged characters with no disability as equally naughty as characters with a physical disability ($p = .64$) or perceptual disability ($p = .22$). The youngest children rated characters with physical disabilities as somewhat naughtier than characters with perceptual disabilities, $p = .02$. In contrast, children in the Middle (5-6 years) and Oldest (7-8

years) age groups judged characters as *much less* naughty when they possessed a physical or perceptual disability, relative to characters who had no disability, all p's < .001 (there were no significant differences in their evaluations of characters with physical vs. perceptual disabilities, Middle: p = .32; Oldest: p = .054). More fine-grained, supplementary analyses confirmed that children began conceptualizing the disabled characters' behaviors as less naughty (relative to typical characters' behavior) somewhere between the ages of 4 and 5 years. The sample of 4-year-olds (M_{age} = 4.53 years, n = 14) judged the three characters as equally naughty, $F(2, 13) = 3.35$, p = .067, η^2_p = .34. However, an equivalently-sized group of the youngest 5-year-olds (M_{age} = 5.37 years, n = 14) differed significantly in their evaluations ($F(2, 13) = 11.68$, p = .001, η^2_p = .64); judging the character without a disability (M = 1.35, SE = 1.54) as significantly more naughty than the character with a perceptual disability (M = .77, SE = .18, p < .001) and the character with a physical disability (M = .80, SE = .18, p < .001). Postestimation tests of supplementary multilevel regression analyses (see Supplementary Materials) confirmed this apparent shift between 4 and 5 years: By 4.5 years children judged characters with a perceptual disability ($\chi^2(1) = 20.69$, p < .001) *and* characters with a physical disability ($\chi^2(1) = 11.24$, p < .001) as less naughty than characters with no disability.

Thus, age-related trends differed significantly depending on whether characters had a disability (physical or perceptual) or had no disability. Children's evaluations of characters with no disabilities were similar across the age-range. However, there were significant age-related *decreases* in the naughtiness that children ascribed to characters with physical or perceptual disabilities; children 4.5-years and older took into account people's disabilities, judging persons with disabilities less harshly relative to persons with no disabilities who produced the same behaviors. These patterns emerged (almost identically) for children's evaluations of persons with

perceptual (auditory) disabilities and persons with physical (walking) disabilities (see Figure 1), and whether characters had committed moral violations or conventional violations (see Figures 2 and 3 for a comparison). Indeed, in our primary analyses, there were no significant interactions involving violation type: Ability X Violation: $F(2, 101) = 1.89$, $p = .16$; $\eta^2_p = .02$; Violation X Age: $F(2, 102) = 2.68$, $p = .07$; $\eta^2_p = .05$; Ability X Violation X Age: $F(4, 204) = .95$, $p = .44$; $\eta^2_p = .02$. The patterns revealed with this ANOVA are similar to patterns found with multi-level linear regression analyses in which age is treated as a continuous variable. Similar patterns are also found using multi-level Poisson regression, which does not assume a normal distribution of residuals (see Supplementary Materials).

Explanations for Non-normative Behaviors

Next, we examine children's *explanations* for characters' non-normative behaviors, and whether those explanations--particularly, their reasoning that characters' behaviors were restricted by their disabilities or that behaviors reflected characters' negative traits and motives--predict children's naughtiness ratings. Note that E never told participants why the characters performed non-normative behaviors, and never explicitly noted connections between characters' disabilities/abilities and their behavior for any of the norm-violation scenarios. Participants were asked to explain each of the three character's behaviors ("Why did [*name*] [*violation*]?"), for each of the four scenarios. One participant did not respond to *any* open-ended question, and so their data are not included in these analyses ($n = 104$). Table 2 depicts the frequency with which children used 7 types of explanations to account for the non-normative behavior of characters with a *perceptual disability*, a *physical disability*, or *no disability*. For illustrative purposes, we present these data with participants binned into three age groups. In general, the most frequent explanation for characters' behavior (aside from "Other") involved characters' *Physical or*

Auditory Limits (e.g., "he/she can't walk/hear"; 47.29% of total responses), *Attributions of Negative Traits or Motives* to the characters (e.g., "she is doing it on purpose"; 14.61% of total responses), and inferences that characters were *Apathetic* (e.g., "she doesn't want to"; 11.27% of total responses).

Our research questions and hypotheses primarily concern children accounting for characters' behavior in terms of their *Physical or Auditory Limitations*, or in terms of characters' *Negative Attributes*. Thus, we directly examine how children's use of these explanations vary across participants' age and across characters' abilities. An additional four participants did not provide responses to one or more open-ended questions about one or more of the three characters, and so their data are not included in these analyses (n = 100). Because there are 4 fewer participants in these analyses, descriptive statistics reported in this section differ slightly from those reported in Table 2.

We first evaluate whether the frequency of mentioning characters' *Physical or Auditory Limitations* varies depending upon participants' ages and characters' abilities, using a 3 (Participant Age: 3-4, 5-6, 7-8 years) X 3 (Character Disability: None (Typical), Physical disability, Perceptual disability) mixed-effects ANOVA, with Character Disability as a within-subjects factor. This analysis revealed a significant effect of Character Disability, $F(2, 96) = 155.17$, $p < .001$, $\eta^2_p = .76$. As depicted in Table 2, across all three age groups, children accounted for characters' non-normative behavior in terms of their *Physical or Auditory Limitations* far more often for characters with physical (M = 64%, SD = 3%) or perceptual disabilities (M = 67%, SD = 3%) than characters with no disability (M = 9%, SD = 2%), ps < .001 (children mentioned limitations equally often for characters with physical and perceptual disabilities, p = .24). However, this pattern became more exaggerated with age, as indicated by a

significant effect of Age ($F(2, 97) = 5.83$, $p = .004$, $\eta^2_p = .11$), subsumed under a significant interaction of Age X Character Disability, $F(4, 194) = 7.48$, $p < .001$, $\eta^2_p = .13$. Post-hoc analyses indicated that even the youngest children (3-4 years) mentioned limitations more often for characters with physical ($M = 47.8\%$, $SD = 6.2\%$) or perceptual ($M = 45.2\%$, $SD = 6.5\%$) disabilities, vs. characters with no disabilities ($M = 13.5\%$, $SD = 3.8\%$), ps $< .001$. However, this distinction was greater among 5-6 year-olds (physical: $M = 68.6\%$, $SD = 4.8\%$; perceptual: $M = 68.2\%$, $SD = 5\%$; vs. no disability: $M = 11.2\%$, $SD = 3\%$, ps $< .001$) and 7-8 year-olds (physical: $M = 75\%$, $SD = 5.7\%$; perceptual: $M = 87.4\%$, $SD = 5.9\%$; vs. no disability: $M = 3.2\%$, $SD = 3.5\%$, ps $<.001$). Indeed, by 8-years, children *typically* mentioned such limitations to account for the non-normative behavior of characters with disabilities.

A similar analysis was performed for participants' inferences that non-normative behavior reflected characters' *Negative Attributes*. This analysis revealed a main effect of Character Disability ($F(2, 96) = 7.66$, $p = .001$, $\eta^2_p = .14$), subsumed under an interaction of Character Disability X Participants' Age, $F(4, 194) = 5.9$, $p < .001$, $\eta^2_p = .11$. The youngest children (3-4 years) did not differ in their attributions of negative qualities to characters, whether or not characters had disabilities: no disabilities ($M = 7.7\%$, $SD = 5.5\%$), physical disabilities ($M = 13.5\%$, $SD = 4.1\%$), perceptual disabilities ($M = 14.4\%$, $SD = 3.9\%$), all ps $> .21$. However, children ages 5-6 years made negative attributions significantly more for characters with no disabilities ($M = 20.7\%$, $SD = 4.3\%$) than characters with disabilities (physical: $M = 8.7\%$, $SD = 3.2\%$, $p = .02$; perceptual: $M = 10.5\%$, $SD = 3.0\%$, $p = .003$). As well, children 7-8 years made negative attributions significantly more for characters with no disabilities ($M = 28.8\%$, $SD = 5.1\%$) than characters with disabilities (physical: $M = 4.8\%$, $SD = 3.7\%$, $p < .001$; perceptual: $M = 0\%$, $SD = 3.6\%$, $p < .001$). For the youngest two age groups (3-4 years and 5-6 years), there

were no differences in negative attributions made to characters with physical vs. perceptual disabilities (*p*s > .40), and there was only a marginal difference among 7- to 8-year-olds ($p = .05$).

In sum, children's interpretation of characters' non-normative behaviors differed based on characters' abilities. Children described physical or auditory limitations *more often*, and mentioned negative attributes *less often*, when accounting for the behavior of characters with disabilities compared to characters without disabilities. These effects varied with age, with older children exhibiting these patterns more robustly than younger children. The patterns revealed with these ANOVAs were similar to patterns found using multi-level regression analyses, in which age was treated as a continuous variable (see Supplementary Materials).

Relations between Explanations for Non-Normative Behavior and Naughtiness Judgments

Next, we examine whether children's explanations for why characters with disabilities produced non-normative behavior—specifically, their noting characters' limitations or attributing negative qualities to characters—predict children's *evaluations* of those characters, and whether these associations vary with children's age. All but two children ($n = 103$) provided explanations for the behavior of characters with perceptual disabilities and physical disabilities.

We first focus on whether children's noting of characters' physical or auditory limitations is related to their evaluations of those characters. We conducted multi-level regression analyses here, as both our primary predictor variable and our primary outcome variable are non-categorical variables. In two separate regressions (one for characters with a perceptual disability, one for characters with a physical disability) we predicted children's evaluations of characters based on children's age (a continuous variable), frequency of mentioning those characters' physical or auditory limitations, and the interaction of these variables. We ran separate models to

evaluate how children's reasoning about the behavior of different characters (e.g., those with physical disabilities) is related to children's evaluations of those *same* characters. As well, previous research has identified differences in the way children understand and reason about physical disabilities and perceptual disabilities, especially among children at the younger end of our age range (Diamond & Hestenes, 1996; Diamond et al., 2008), and this might lead to differing associations between children's explanations and evaluations of non-normative behavior for different disabilities. Neither regression analysis revealed a significant interaction between children's age and their explanation usage, so this interaction effect was removed from both models. The reduced models revealed remarkably similar trends for characters with perceptual disabilities and characters with physical disabilities.

As revealed in earlier analyses, there were age-related decreases in how naughty children judged characters with disabilities (Perceptual Disability: β = -.15, *SE* = .04, *t* = -4.25, $p < .001$, 95% CI [-.23, -.08]; Physical Disability: β = -.19, *SE* = .03, *t* = -5.98, $p < .001$, 95% CI [-.25, -.12]). More critical to the current analysis, these models revealed significant, negative associations between children's mentioning characters' limitations and their judgments of characters' naughtiness (Perceptual Disability: β = -.64, *SE* = .15, *t* = -4.09, $p < .001$, 95% CI [-.95, -.33]; Physical Disability: β = -.79, *SE* = .15, *t* = -5.40, $p < .001$, 95% CI [-1.08, -.50]). As depicted in Figure 4, more often citing characters' physical or auditory limitations was associated with rating those characters as significantly *less* naughty. This was true for children across the age range (as indicated by the lack of significant interaction effects involving age).

Similar analyses were conducted to examine whether children's inferences about characters' *negative attributes* are associated with their evaluations of those characters. In two separate regressions (one for characters with a perceptual disability, one for characters with a

physical disability) we predicted children's evaluations of characters based on children's age (a continuous variable), frequency of mentioning characters' negative attributes, and the interaction of these variables. Neither analysis revealed a significant interaction between children's age and their explanation usage, so this interaction effect was removed from both models. The resulting, reduced models revealed almost identical trends for characters with perceptual disabilities and characters with physical disabilities. These analyses again identified age-related decreases in how naughty children judged characters with disabilities (Perceptual Disability: β = -.19, *SE* = .03, *t* = -5.38, $p < .001$, 95% CI [-.26, -.12]; Physical Disability: β = -.22, *SE* = .03, *t* = -6.83, $p < .001$, 95% CI [-.28, -.15]). More importantly, these models revealed *positive* associations between inferring negative attributes and naughtiness evaluations (Perceptual Disability: β = .79, *SE* = .24, *t* = 3.28, *p* = .001, 95% CI [.31, 1.26]; Physical Disability: β = .70, *SE* = .23, *t* = 3.01, $p < .01$, 95% CI [.24, 1.15]). As depicted in Figure 5, more often inferring characters' negative traits or motives was associated with evaluating those persons as significantly *naughtier*. This was equally true for children across the entire age range (as indicated by the lack of significant interactions with age).

Discussion

The current study was designed to examine 3- to 8-year-olds' evaluations of persons with disabilities who produce non-normative behaviors, and to examine children's explanations for these behaviors. We found intriguing age-related shifts in children's evaluations—beginning at around 4.5-years, children evaluated persons with disabilities as less naughty than typically-developing persons who produced identical behavior. We also found age-related shifts in children's explanations for persons' non-normative behavior and identified two commonly-used explanations—characters' limitations, and characters' negative attributes—that predicted how

harshly children evaluated those persons. In the following sections, we review these findings in greater detail, and integrate them with prior work on children's concepts of physical and perceptual abilities and work on the development of children's moral cognition. We also identify important directions for future research on these topics.

Children's *Evaluations* of Non-normative Behaviors Depends on Actors' Abilities

A primary goal of this study was to investigate whether, across early development (3-8 years), children judge persons with physical or perceptual disabilities as more or less naughty than TD persons when they commit identical moral or conventional violations. Children younger than 4.5 years generally did *not* modify their naughtiness ratings based on characters' abilities; they rated characters as equally naughty, even if they possessed disabilities that might account for their non-normative behavior. These findings for the youngest participants are generally consistent with those of Huckstadt and Shutts (2014), who found that preschoolers (ages 3-5 years) tended to not modify their judgments based on whether a norm violator had a perceptual disability (blindness). However, because Huckstadt and Shutts (2014) analyzed children ranging from 3-5 years as one group, they could not identify age-related differences that may exist across these early years. The current study identified that children as young as 4.5-years *did* account for persons' disabilities, judging characters with perceptual *and* physical disabilities as less naughty (compared to characters without disabilities) when they violated norms.

Children as young as 4.5 years (54 months) in the current study less harshly evaluated persons who behaved non-normatively if they possessed a disability, whereas Josephs et al. (2016) found that even younger children (averaging 47 months) were less likely to protest and correct the non-normative actions of physically-constrained persons. This apparent age discrepancy might be a function of measurement differences—we measured children's

evaluations of persons who violated norms, whereas Josephs et al. (2016) examined children's denial and correction of non-normative behavior; one may deny or correct another's action without deeming the actor naughty. The apparent age discrepancy might also reflect differences in the constraints experienced by actors. In the current study, characters' disabilities were not visible or alterable; they were hidden (characters used no visible medical equipment) and were described as being stable (rather than temporary), whereas Josephs et al. (2016) visibly limited a puppet's movement with a removeable rope.

The extent to which children (and adults) take others' constraints into account when evaluating their behavior is contingent upon their ability to *mentally represent* and *remember* such constraints. The tendency for 3-year-olds and young 4-year-olds to evaluate persons similarly whether or not they possessed disabilities may be a function of these cognitive demands. Or perhaps, as is true of most studies that ask children to evaluate others, there may be implicit demand characteristics, especially for the youngest children (e.g., children may assume that saying that the character is "naughty" is the "right" answer). As well, 3-year-olds have more difficulty than older children in understanding that unseen entities and qualities persist over time (e.g., that sugar still exists in water, even after it is dissolved; Rosen & Rozin, 1993), and also have a tendency to treat things that are unseen as less "real" (Woolley & Brown, 2015). Given the lack of visual cues to characters' disabilities, it is that much more impressive that children as young as 4.5-years in the current study *regularly* evaluated norm violators more leniently when they possessed non-visible physical or perceptual disabilities. We wish to note, though, that our goal was not to *pinpoint* a precise developmental moment when children begin to account for persons' non-visible disabilities in their moral evaluations, but to identify roughly when this tendency emerges within the preschool years and how it proceeds to develop during middle

childhood. We would not be surprised if children raised in other contexts might exhibit this tendency somewhat earlier or somewhat later than children in the current sample.

Thus, although visible markers of disabilities can increase the likelihood that others will notice and appreciate disabilities, visible markers were not *necessary* for U.S. preschoolers to understand (and then draw inferences from) disabilities. Rather, *descriptions* of the disabilities that included specific examples of how those disabilities would affect persons (e.g., for the hearing disability: "They can't hear if a firetruck is coming down the street. They can't hear the school bell at the end of the day. They can't hear their friends yelling on the playground") were sufficient for many children to account for those disabilities in their judgments of persons for their non-normative behavior. These findings expand upon work demonstrating that preschoolers understand some implications of *exceptional* abilities after they are provided detailed descriptions of those abilities (e.g., Lane et al., 2010).

Our findings may have implications for how young children are taught about disabilities, at least in one cultural context. The fact that preschoolers were able to use descriptions of physical and perceptual disabilities (in the absence of any visual markers) to calibrate their evaluations of persons for their non-normative behavior suggests that children are capable of learning about these types of disabilities at a young age, when provided appropriate descriptions. The descriptions used in the current study were well suited to test our specific research questions and to limit confounding variables, but these descriptions may be ecologically limited and are not necessarily the precise descriptions that should be used in real-world applications to introduce children to ideas about disability. Disability descriptions provided by parents, teachers, and in children's books are multi-faceted: they might describe non-normative behaviors, social implications, medical equipment, accommodations, and more. No research (to our knowledge)

has examined how each of these components or the order in which components are presented impacts children's comprehension of disabilities; that research is necessary before recommending specific educational interventions. A limitation of describing disabilities as stable (as in the current study) is the possibility of reinforcing essentialist notions of *immutability*: the idea that membership in a certain social category (in this case, having a particular disability) is not changeable (Bastian & Haslam, 2006). Essentialist notions about groups of people may be linked to higher levels of bias towards those groups (Bogart, Rosa, & Slepian, 2019), although we have no way of knowing if the descriptions used in the current study influenced essentialist notions or bias. Future research may be directed at identifying precisely what language is most effective in communicating to children ideas about different sorts of disabilities.

Children's *Reasoning* about Non-normative Behaviors Depends on Actors' Abilities

A second goal of this study was to investigate whether, across early development, children's *explanations* for others' non-normative behavior differ depending upon whether individuals have physical disabilities, perceptual disabilities, or no disabilities. For each scenario, participants were asked to explain *why* each character performed their non-normative behavior. Note that the experimenter never cited characters' disabilities as the reason for their non-normative behavior, so we were able to evaluate the extent to which children made these connections *themselves*. We purposely did not provide information regarding the characters' intentions because (1) in real-life, we typically do not know the intentions of actors when we witness non-normative behavior, and (2) we wanted to capture the extent to which children would attribute *different kinds* of intentions to characters with different abilities. We found that children as young as 4-years of age more often cited physical or auditory limitations to account for the non-normative behavior of persons with disabilities (vs. persons without disabilities), and

there was a marked age-related increase in children's tendency to cite such limitations. Thus, young preschoolers understood that disabilities may contribute to non-normative behavior, and this understanding increased substantially across early and middle childhood.

Another way that people make sense of others' undesirable behavior is by attributing to them negative traits or intentions (e.g., Dodge, 1980; Dodge et al., 1986; Leslie et al., 2006; Orobio De Castro et al., 2002). As one might expect, children used this reasoning most often for characters *without* disabilities. But this varied by children's age, with only children 5.5 years of age and older using this reasoning significantly more for characters without disabilities than with disabilities. Indeed, the oldest children (7-8 years) rarely used this reasoning for characters with physical disabilities, and *never* used this reasoning for characters with perceptual disabilities. Thus, older children were more prone to infer that typically-developing persons' non-normative behavior reflected negative traits or negative intentions, but these children refrained from making such inferences when persons had disabilities that accounted for their behavior.

Children's Evaluations of Non-normative Behaviors was associated with their Reasoning

A final goal of this study was to investigate whether children's reasoning about the non-normative behaviors of individuals with disabilities relates to their evaluations of these persons. Children who frequently mentioned *physical or auditory limitations* to account for the non-normative behavior of persons with disabilities rated these characters as significantly *less naughty* (vs. participants who rarely mentioned these limitations). In contrast, children who attributed *negative qualities or motives* to persons with disabilities rated these characters as significantly *naughtier* than those who did not mention these attributes in their responses. These effects did not interact with participants' age, and thus were equally strong among children across the age range of 3-8 years.

The frequency with which children used these two types of explanations differed significantly across development, and children's use of these explanations predicted how naughty children judged the characters as. Thus, one might be tempted to reason that age differences in children's explanations fully account for age differences in children's naughtiness evaluations. However, associations between these explanations and children's naughtiness evaluations were consistent across the age range, and the same age-related decreases in naughtiness evaluations emerged regardless of how often children used these explanations. This suggests that additional developments account for the age-related differences in evaluations of disabled persons who produce non-normative behavior. Whereas our explanation data demonstrate age-related differences in children's (1) recognizing the existence of characters' perceptual and physical disabilities, and (2) understanding the implications of those disabilities in the production of non-normative behavior, our data do not capture the extent to which children believe that having a disability is a *valid excuse* for non-normative behavior. Future work should examine how this factor might account for age-related shifts in children's evaluations of persons who perform non-normative actions. As well, there may be important shifts during the late preschool years (4-5 years of age) in how children apply their theory of mind (ToM) when evaluating others who produce non-normative behavior (Killen et al., 2011). The development and application of a ToM may be particularly critical in children's attempts to conceptualize the minds of persons with disabilities, whom they typically have less contact with and whose disabilities may be novel.

Additional Limitations and Future Directions

The current study was designed to investigate children's judgments of persons with disabilities (described verbally, rather than marked visually) when they produce non-normative

behaviors. Maximizing the internal validity of our study entailed sacrificing some external validity. For example, we specifically chose to not use visual markers of disability to reduce confounds associated with children's negative sentiments toward medical equipment and other physical markers of disability (Diamond et al., 2008; Huckstadt & Shutts, 2014). Yet, in the real world, these markers do exist at times (for some disabilities) and may sway persons' evaluations of non-normative behavior. We predict that if medical equipment is depicted, children will be less lenient towards characters with disabilities who produce non-normative behavior because of biases often associated with that equipment. This may be particularly true for children with little exposure to persons with disabilities (as was the case for our participants; see Supplementary Materials). Future work can test these possibilities. As well, in real-world contexts, disabilities may be accompanied by labels. For example, children may be told that a classmate has "Autism," or "Cerebral Palsy". Labels can be used to essentialize qualities of individuals -- to suggest that those qualities are inherent and stable (Gelman, 2004). Understanding how such labels influence children's ideas about and behavior toward individuals with disabilities is a critical next step.

Each violation in this study was designed to meet two criteria: (1) *both* mobility and hearing disabilities could be potential reasons for characters' non-normative behavior, and (2) violations were likely to occur in real-world settings for children with disabilities. Future studies may employ scenarios for which only one type of disability could account for non-normative behavior, to gauge the extent to which children appreciate how *unique features* of different disabilities are implicated in *specific* non-normative behaviors, thus rendering some behaviors more excusable than others. For example, children and adults may believe that a child with a physical disability has a more valid reason for seeing a marker fall on the floor across the room

and not picking it up (they cannot easily move to get it) than a child with a hearing disability (who could see the marker fall *and* walk toward it).

Across the age-range sampled, children judged the "conventional" violations as naughtier than the "moral" violations, contrary to some past work (e.g., Hardecker et al., 2016). Perhaps our conventional violations were interpreted by children as more active (e.g., talking too loudly) than our moral violations (e.g., passively not helping someone). Children and adults evaluate people more negatively when an outcome is a function of action vs. inaction (Greene et al., 2009), although it is unknown whether this pattern generalizes to children's judgments of persons with disabilities. Past research has also revealed that violations involving psychological hurt may be evaluated more leniently than violations involving physical hurt, perhaps because the former demands additional theory-of-mind processing (Ball et al., 2016). Our use of a "moral" violation chosen to capture psychological hurt (see Appendix A: "Rejecting Classmate's Request to Play") may have contributed to children's lower naughtiness judgments averaged across our "moral" violations. Future studies can address these concerns more directly.

Moreover, additional studies are needed to evaluate how children in different cultures reason about these scenarios, because societies might have different norms for the inclusion and accommodation of folks with disabilities, and because concepts of what constitute different violation categories (e.g., moral vs. conventional) varies between societies (e.g., Levine et al., 2021). Indeed, cultural norms have a significant impact on moral judgment (Haidt, Koller, & Dias, 1993; Turiel, 2007). For example, speaking loudly during quiet classroom time (one of our conventional violation scenarios) may be interpreted as more "moral" than "conventional" by children in more collectivist cultures that prioritize the harmony of the group over the specific needs or behaviors of individuals. However, in more individualist cultures, such as the United

States, similar scenarios are often conceptualized as conventional (Nucci, 1981; Tisak & Turiel, 1988; Smetana, 1981). Importantly, in advance of future research on this topic, the current study identified several patterns that *replicated across* violations that we have categorized as "moral" and "conventional", demonstrating the robustness of the findings: beginning at around 4.5-years of age, children's evaluations of and reasoning about persons who committed norm violations differed depending upon whether those persons had disabilities that might account for those violations.

Thus, even before entering formal schooling, children can understand the implications of certain unseen physical and perceptual disabilities, and may begin to flexibly adjust their evaluations of persons who commit violations when those persons have disabilities that account for their behavior. These evaluative adjustments are more marked in middle childhood, likely reflecting developmental advances in children's recognition that others' abilities and disabilities may yield non-normative behavior. Given this early aptitude, it is unlikely that children's ostracism and victimization of peers with disabilities (Koller et al., 2018; Rose & Gage, 2016) can be credited to their merely not understanding those disabilities and their accompanying behaviors. An important next step is identifying how to harness children's early understandings of disability and their early-emerging moral inclinations in order to mitigate the development of biases against persons with disabilities and to promote more positive interactions among typically-developing children and children with disabilities.

References

Baird, J. A. & Astington, J. W. (2004). The role of mental state understanding in the development of moral cognition and moral action. *New Directions for Child and Adolescent Development, 103,* 37-49. doi:10.1002/cd.96

Ball, C. L., Smentana, J. G., & Sturge-Apple, M. L. (2016). Following my head and heart: Integrating preschooler's empathy, theory of mind, and moral judgments. *Child Development, 88*, 1-15. doi:10.1111/cdev.12605

Bastian, B. & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology, 42*(2), 228-235. doi:10.1016/j/jesp.2005.03.003

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159. doi: 10.1037//0033-2909.112.1.155

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition, 127*, 6-21. doi:10.1016/j.cognition.2012.11.008

Dahl, A., & Kim, L. (2014). Why is it bad to make a mess? Preschoolers conceptions of pragmatic norms. *Cognitive Development, 32,* 12-22. doi:10.1016/j.cogdev.2014.05.004

Diamond, K. E., & Hestenes, L. L. (1996). Preschool children's conceptions of disabilities. *Topics in Early Childhood Special Education, 16*(, 458-475. doi:10.1177/027112149601600406

Diamond, K. E., Hong, S-Y., & Tu, H. (2008). Context influences preschool children's decisions to include a peer with a physical disability in play. *Exceptionality*, *16*, 141-155. doi:10.1080/09362830802198328

Diamond, K. E. & Hong, S-Y. (2010). Young children's decisions to include peers with disability in play. *Journal of Early Intervention, 32*, 163-177. doi:10.1177/1053815110371332

Diamond, K. E., Le Furgy, W., & Blass, S. (1993). Attitudes of preschool children toward their peers with disabilities: a year-long investigation in integrated classrooms. *The Journal of Genetic Psychology, 154*, 215-221. doi:10.1080/00221325.1993.9914735

Diamond, K. E., & Kensinger, K. R. (2002). Vignettes from sesame street: Preschooler's ideas about children with down syndrome and physical disability. *Early Education and Development, 13*, 409-422. doi:10.1207/s15566935eed1304_5

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development, 51*, 162-170. doi:10.2307/1129603

Dodge, K. A., Pettit, G. S., McClaskey, C. L., Brown, M. M., & Gottman, J. M. (1986). Social competence in children. *Monographs of the Society for Research in Child Development, 51*, I-85. doi:10.2307/1165906

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. doi:10.3758/BF03193146

Gelman, S.A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*(9), 404-409. doi:10.1016/j.tics.2004.07.001

Göckeritz, S., Schmidt, M. F. H., & Tomasello, M. (2014). Young children's creation and transmission of social norms. *Cognitive Development, 30,* 81-95. doi:10.1016/j.cogdev.2014.01.003

Greene, J.D., et al. (2009). Pushing moral buttons: the interaction between personal force and intention in moral judgement. *Cognition, 111*, 364-371. doi:10.1016/j.cognition.2009.02.001

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*(4), 613-628. doi:10.1037/0022-3514.65.4.613

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. doi:10.1016/j.jecp.2016.06.012

Hebble, P. W. (1971). The development of elementary school children's judgment of intent. *Child Development, 42*(4), 1203-1215. doi:10.2307/1127804

Huckstadt, L. K., & Shutts, K. (2014). How young children evaluate people with and without disabilities. *Journal of Social Issues, 70*, 99-114. doi:10.1111/josi.12049

Josephs, M., Kushnir, T., Gräfenhain, M., & Rakoczy, H. (2016). Children protest moral and conventional violations more when they believe actions are freely chosen. *Journal of Experimental Child Psychology, 141,* 247-255. doi:10.1016/j.jecp.2015.08.002

Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition, 119*, 197-215. doi: 10.1016/j.cognition.2011.01.006

Koller, D., Le Pouesard, M., & Anneke Rummens, J. (2018). Defining social inclusion for children with disabilities: A critical literature review. *Children & Society, 32*, 1-13. doi:10.1111/chso.12223

Koller, D., & San Juan, V. (2015). Play-based interview methods for exploring young children's perspectives on inclusion. *International Journal of Qualitative Studies in Education, 28*, 610-631. doi:10.1080/09518398.2014.916434

Lane, J. D., Evans, E. M., Brink, K. A., & Wellman, H. M. (2016). Developing concepts of ordinary and extraordinary communication. *Developmental Psychology, 52*, 19-30. doi:10.1037/dev0000061

Lane, J. D., Wellman, H. M., & Evans, E. M. (2010). Children's understanding of ordinary and extraordinary minds. *Child Development, 81*, 1475-1489. doi:10.1111/j.1467-8624.2010.01486.x

Levine, S., Rottman, J., Davis, T., O'Neill, E., Stich, S., & Machery, E. (2021). Religious affiliation and concepts of the moral domain. *Social Cognition.*

Moll, H., Carpenter, M., & Tomasello, M. (2014). Two- and three-year-olds know what others have and have not heard. *Journal of Cognition and Development, 15*, 12-21. doi: 10.1080/15248372.2012.710865

Nucci, L. (1981). Conceptions of personao issues: a domain distinct from moral or societal concepts. *Child Development, 52*(1), 114-121. doi:10.2307/1129220

Orobio De Castro, B., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). Hostile attribution of intent and aggressive behavior: a meta-analysis. *Child Development, 73*, 916-934. doi:10.1111/1467-8624.00447

Piaget, J. (1932/1965). *The moral judgment of the child.* The Free Press.

Pillow, B. H. (1989) Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology, 47*, 116-129. doi:10.1016/0022-0965(89)90066-0

Rose, C. A., & Gage, N. A. (2016). Exploring the involvement of bullying among students with disabilities over time. *Exceptional Children, 83*, 298-314. doi:10.1177/0014402916667587

Rosen, A. B., & Rozin, P. (1993). Now you see it, now you don't: The preschool child's conception of invisible particles in the context of dissolving. *Developmental Psychology*, 29, 300–311. doi:10.1037/0012-1649.29.2.300

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52*(4), 1333-1336. doi:10.2307/1129527

Stata Corporation (2019). *Stata statistical software: release 16*. College Station, TX: StataCorp LLC.

Tisak, M. S., & Turiel, E. (1988). Variation in seriousness of transgressions and children's moral and conventional concepts. *Developmental Psychology, 24,* 352-357. doi:10.1037/0012-1649.24.3.352

Turiel, E. (2007). The development of morality. *Handbook of Child Psychology, 3,* 473-514. doi: 10.1002/9780470147658.chpsy0313

Williamson, R. A., Brooks, R., & Meltzoff, A. N. (2013). The sound of social cognition: Toddler's understanding of how sound influences others. *Journal of Cognition and Development, 16*, 252-260. doi:10.1080/15248372.2013.824884

Woolley, J. D., & Brown, M. M. (2015). The development of children's concepts of invisibility. *Cognitive Development, 34*, 63-75. doi:10.1016/j.cogdev.2014.12.009

Table 1

Coding Scheme for Open-ended Responses

Category	Description	Examples
Physical or Auditory Limitation	Mention of physical or auditory capabilities	*"Because he **can't**."* *"Because her **ears don't work**."* *"Because he **couldn't walk**."*
Negative Person/Action	Mentions negative aspect of the person's character or behavior	*"He's **naughty**."* *"Because maybe **she was a bully** to her."* *"Because she was **lazy**."*
Negative Motive	Protagonist has a negative intention or desire	*"Probably because he **doesn't like that boy**."* *"Because well she is **doing it on purpose**."* *"Maybe she might have **ignored her** or something."*
Conflicting Desire	Protagonist wants to do something else	*"Maybe she **was already playing with someone**."* *"Maybe she **was having fun**."* *"Because he **wanted to swing** on that swing."*
Apathy	Protagonist doesn't care	*"Doesn't **care**."* *"Doesn't **want to**."*
Unaware	Protagonist didn't know or see. Responses are *unrelated* to the character's disability.	*"She **wasn't listening**."* *"Because she **didn't notice**."* *"Everybody was yelling."*
Diffusion of Responsibility	Denial of obligation	*"Because **he's not the teacher**."* *"Because **he's not his friend**."* *"Because she **thinks someone else will do it**."*
Other	Responses that do not fit other categories, as well as "I don't know" responses.	*"Maybe the girl was just mean to her sometimes."* *"He wanted everyone to hear it."* *"Because he's a boy."*

Table 2
Reasoning about the Behavior of Characters with No Disability, Perceptual Disabilities, and Physical Disabilities

No Disability

Age (years)	Physical/Auditory Limit	Negative Attributions	Apathy	Unaware	Conflicting Desire	Diffusion of Responsibility	Other
3.00-4.99	13%	11%	26%	3%	8%	4%	37%
5.00-6.99	11%	21%	22%	7%	21%	2%	21%
7.00-8.99	3%	29%	33%	5%	11%	5%	23%
Total	9%	20%	27%	5%	15%	3%	26%

Perceptual Disability

Age (years)	Physical/Auditory Limit	Negative Attributions	Apathy	Unaware	Conflicting Desire	Diffusion of Responsibility	Other
3.00-4.99	43%	20%	9%	0%	7%	3%	20%
5.00-6.99	68%	10%	3%	6%	2%	2%	11%
7.00-8.99	88%	0%	5%	2%	2%	2%	7%
Total	67%	10%	5%	2%	5%	2%	12%

Physical Disability

Age (years)	Physical/Auditory Limit	Negative Attributions	Apathy	Unaware	Conflicting Desire	Diffusion of Responsibility	Other
3.00-4.99	48%	16%	6%	0%	8%	3%	21%
5.00-6.99	69%	9%	3%	1%	6%	2%	11%
7.00-8.99	76%	5%	1%	0%	10%	2%	9%
Total	65%	9%	3%	0%	8%	2%	13%

Note. n = 104

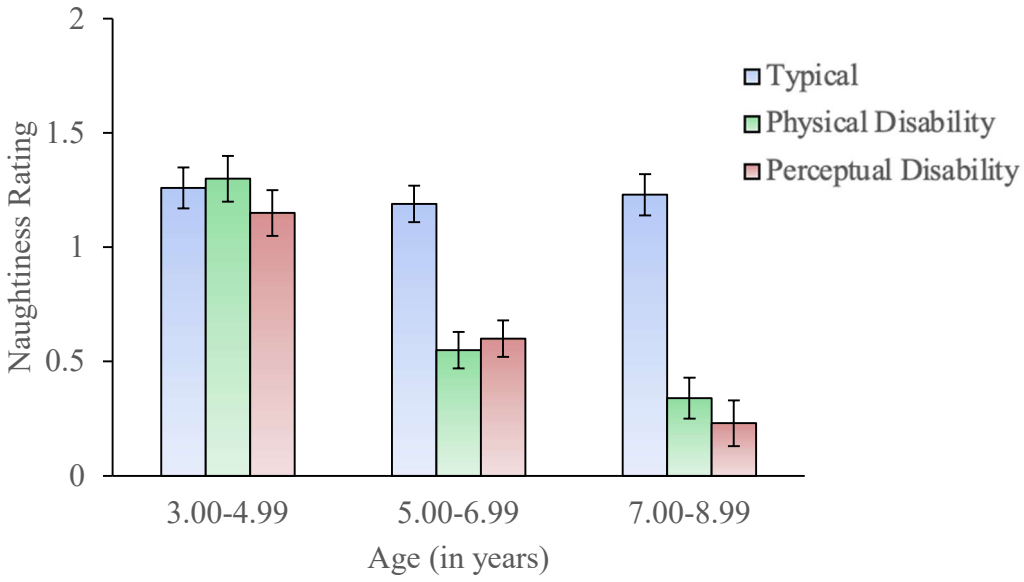


Figure 1. Naughtiness ratings of characters who committed violations, among children in the Youngest (3.00-4.99 years), Middle (5.00-6.99 years), and Oldest (7.00-8.99 years) age groups. Characters either possessed no disability (Typical), a physical disability, or a perceptual disability. Individual naughtiness ratings can range from 0-2. Error bars represent +/- 1 standard error of the mean.

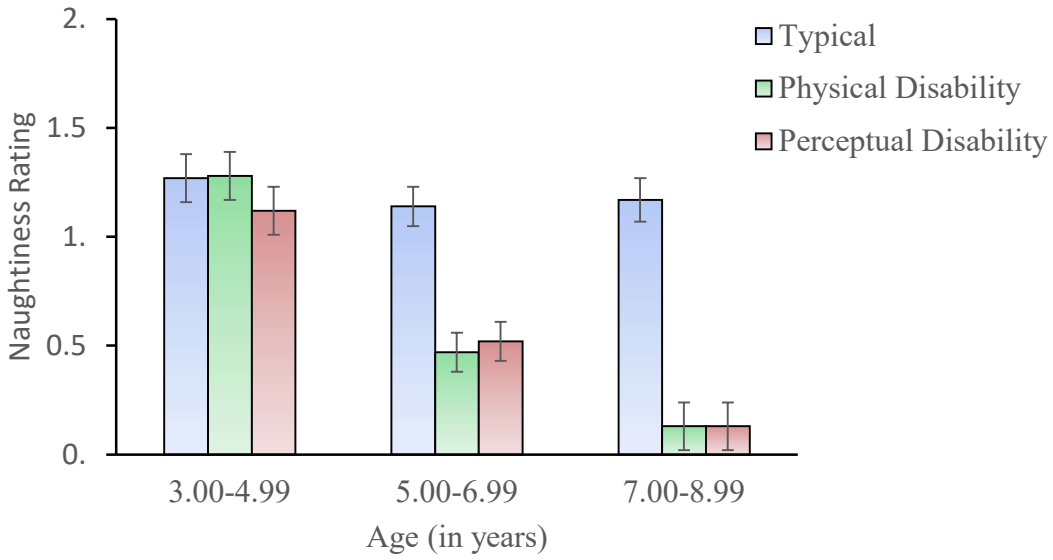


Figure 2. Naughtiness ratings of characters who committed **moral** violations, among children in Youngest (3.00-4.99 years), Middle (5.00-6.99 years), and Oldest (7.00-8.99 years) age groups. Characters either possessed no disability (Typical), a physical disability, or a perceptual disability. Individual scores can range from 0-2. Error bars represent +/- 1 standard error of the mean.

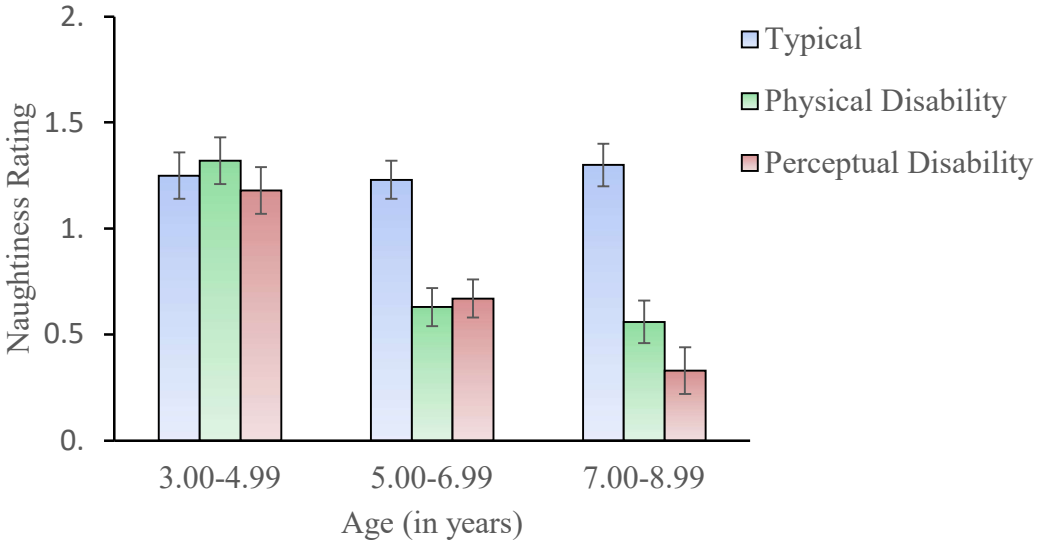


Figure 3. Naughtiness ratings of characters who committed **conventional** violations, among children in Youngest (3.00-4.99 years), Middle (5.00-6.99 years), and Oldest (7.00-8.99 years) age groups. Characters either possessed no disability (Typical), a physical disability, or a perceptual disability. Individual scores can range from 0-2. Error bars represent +/- 1 standard error of the mean.

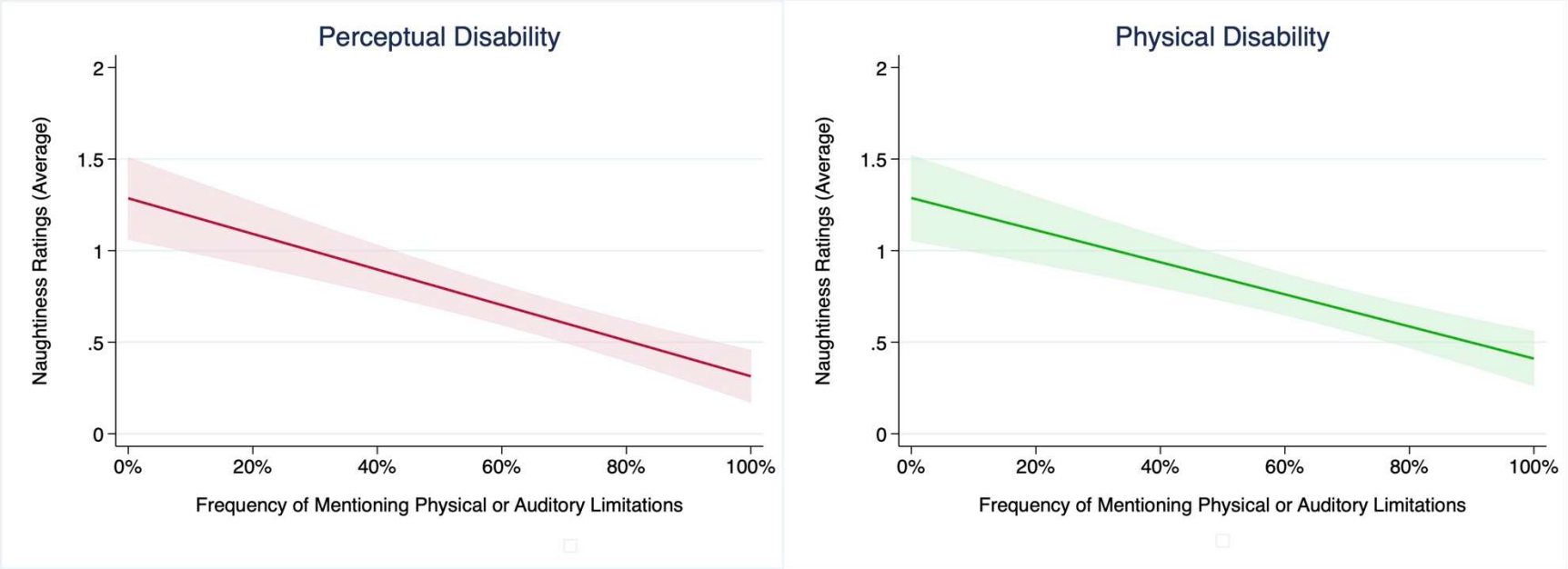


Figure 4. Fitted regression lines depicting relations between children's frequency of mentioning characters' physical or auditory limitations when justifying characters' violations and children's ratings of those characters' naughtiness. Characters had a perceptual disability (left panel) or a physical disability (right panel). Individual *naughtiness ratings* can range from 0-2. Shaded regions depict 95% confidence intervals.

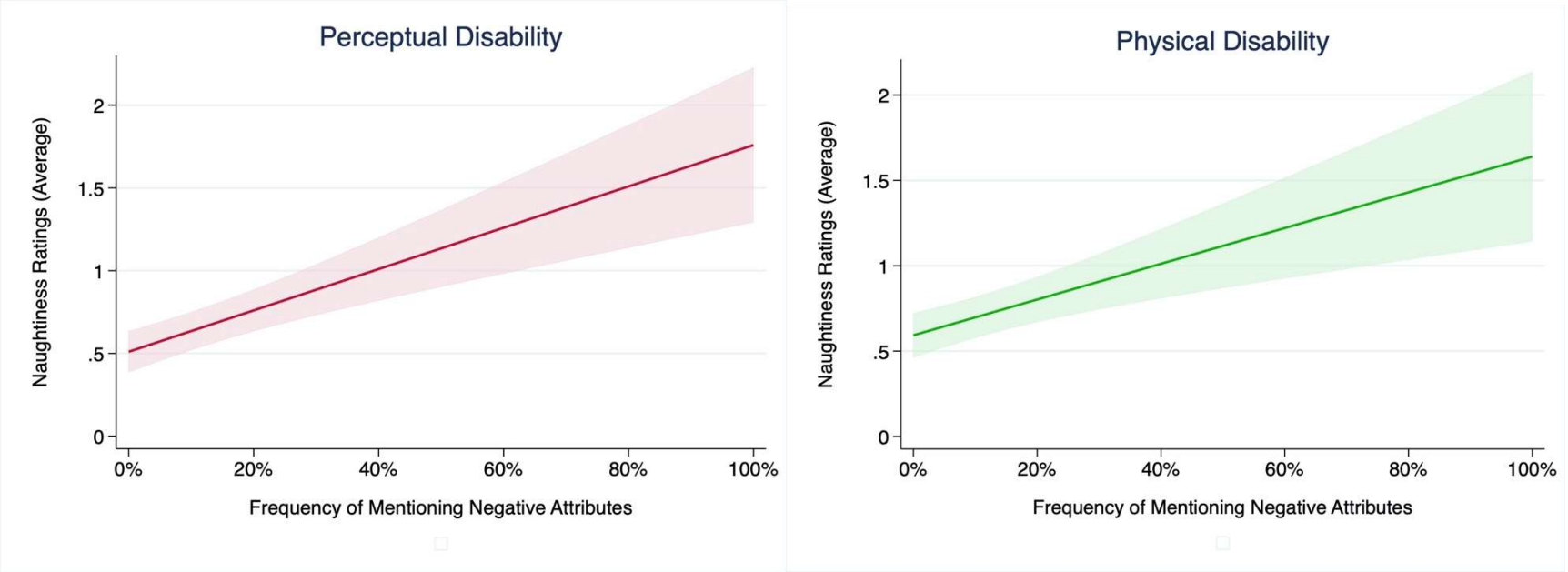
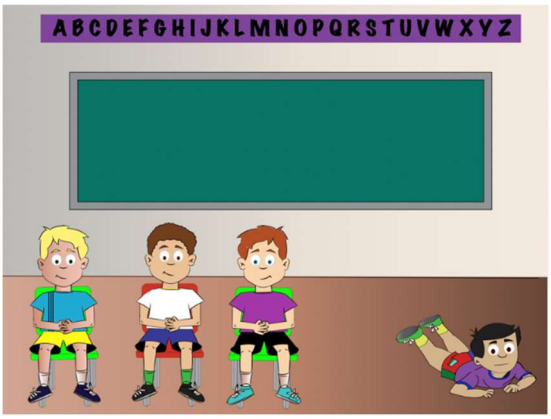


Figure 5. Fitted regression lines depicting relations between children's frequency of mentioning characters' negative attributes when justifying characters' violations and children's ratings of those characters' naughtiness. Characters had a perceptual disability (left panel) or a physical disability (right panel). Individual *naughtiness ratings* can range from 0-2; graphed values above 2 are an artifact of fitted regression lines. Shaded regions depict 95% confidence intervals.

Appendix A

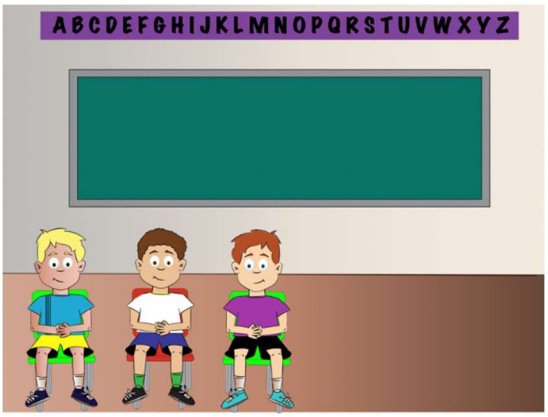
Scenario Scripts and Example Graphics

Not Helping Fallen Classmate - Moral Violation



"These three boys are in the same classroom. "Look!" Another boy in the class trips and falls down. He screams for someone to help him.
[*Character*] doesn't help the boy who fell."

Talking Too Loud in Class - Conventional Violation



"These are the same kids we just talked about. The kids are supposed to be working quietly in the classroom.
[*Character*] is talking too loud in the classroom."

Rejecting Classmate's Request to Play - Moral Violation

"These four girls are in the same class. Look! One girl asks the other girls to run and play with her during recess!
[*Character*] doesn't run and play with the girl."

Not Completing Classroom Activity - Conventional Violation

"These are the same kids we just talked about. Look! It's reading time, and the teacher asks the class to walk to the bookshelf and pick out a book to read." [*note. The teacher was never depicted*].
[*Character*] doesn't walk over to pick out a book for reading time like the teacher asked."