Just Rewards: Children and Adults Equate Accidental Inequity with Intentional Unfairness

Elizabeth Donovan* and Deborah Kelemen
Department of Psychology, Boston University, Boston, MA 02215, USA
* Corresponding author, present address: Inflexxion Inc., 320 Needham Street, Suite 100, Newton, MA 02464, USA, e-mail: donovanelizabeth@gmail.com

Abstract
Humans expect resources to be distributed fairly. They also show biases to construe all acts as intentional. This study investigates whether every unequal distribution is initially assumed to be intentional unfairness. Study 1 presents a control group of adults with a movie showing one individual accidentally receiving less reward than expected for a task. The experimental group was shown the same scenario, except that the individual was now in the presence of an additional person who received the full reward. Despite the similarity of the scenarios, as predicted, participants in the control condition responded as if the disappointing reward was accidental, while those in the experimental condition responded as if the act was intentional: Their tendency to avoid the “perpetrator” did not differ from that of participants in another control condition who saw an intentionally unfair reward distribution. In Study 2, 7- and 8-year-old children’s results replicated those of adults. Implications for social and moral cognition are discussed.

Keywords
Intentionality, fairness, social cognition, morality

The world is not fair. Throughout life, some people will be given advantages while others will not. Despite this, we still expect resources to be distributed fairly. Research indicates that this expectation robustly emerges at around 7 years of age (Murnigham and Saxon, 1998; Benenson et al., 2007; Fehr et al., 2008), although it may show signs as early as the preschool years (Nishida et al., 2005). Indeed, the assumption may have deep evolutionary roots: monkeys and chimpanzees both respond negatively to intentional distributional inequity (Brosnan and de Waal, 2003; Brosnan et al., 2005). Given the centrality of fairness to our assumptions of just moral outcomes,
how do we evaluate and act towards others whose actions – whether intentional or not – violate expectations that rewards should be distributed fairly?

One view is that we take the intentions underpinning the actions into consideration and our behavior toward others is guided by this. According to this view, adults should forgive individuals whose actions accidentally produce an unequal distribution of reward and only blame those who intend to engage in unfairness. In his extensive work on moral decision-making Piaget regarded this as typifying an adult response. He plotted a developmental trend in which young children start out ignoring underlying intentions and instead focus on outcomes while older children and adults shift to doing the reverse. Thus, adult differ from 3-year-olds by viewing a person who accidentally breaks 15 cups as less morally culpable than someone who breaks one intentionally (Piaget, 1932; also Zelazo et al., 1996; Helwig et al., 2001; but see Nobes et al., 2009 on children).

Another view, however, is that our belief that people should be treated equally is so deeply rooted that we are hyper-vigilant to any violation of this expectation; as a result, our attributions of blame are far less rational. That is, while reflective reasoning might lead us to claim that intentional unfairness is worse than an accidental inequality, our behavior might reveal a tacit tendency to ascribe blame to anyone whose actions violate the notion of equal distribution of rewards, whether they behaved deliberately or not.

There are two reasons to expect this non-rational response. First, a body of contemporary research suggests that moral judgments are often more driven by emotion-based “gut reactions” than reasoned arguments (e.g., Haidt, 2001; Greene and Haidt, 2002; Mikhail, 2007; Hauser et al., 2007). So, while we might engage in post-hoc rationalizations about our moral evaluations that serve to justify them – particularly when we want to convince others – the judgments themselves are unconscious, automatic, and rooted in a handful of innately prepared, emotionally valenced intuitions (Haidt and Bjorkland, 2008). As a result, people can make moral judgments on an emotionally charged scenario – for example, family members eating their pet dog after it is accidentally killed – almost reflexively. What takes much longer, and often fails, is their ability to rationally justify their automatic moral evaluation especially in cases like the one above which involve no harm (Haidt et al., 1993).

A second reason to expect that adults might blame someone who accidentally causes inequity is research which suggests that we have an underlying bias to initially appraise any action involving an agent as intentional, a bias which adults are particularly prone to reveal under processing pressure. This tendency has long been noted in children who will claim that all manner of
behaviors (e.g., sneezing, dropping and breaking possessions) are deliberate acts (e.g., Shultz et al., 1980). Traditionally, it has been assumed that children make these errors because they are simply less expert than adults at accurately distinguishing when something is intentional versus accidental. However, Rosset (2008) has recently suggested an alternative explanation: children lack the body of knowledge (e.g., knowledge of involuntary physiological behaviors, commonly desired outcomes) that allows older individuals to inhibit an initial assumption that any event involving an agent is intentional. In contrast to the traditional view, this proposal makes the distinctive prediction that when processing resources are limited, even adults will demonstrate tendencies to view ambiguous events as deliberate. Consistent with this, Rosset finds that adults who are asked to make judgments at speed are more likely than unspeeded adults to view ambiguous but prototypically accidental events (e.g., “she popped the balloon”) as intentional rather than accidental. She also finds that when asked to give descriptions of ambiguous but prototypically accidental events (e.g., “she broke the vase”), adults spontaneously offer intentional interpretations. Evidence of a life-long intentionality bias therefore joins a body of research suggesting that a number of explanatory biases evident in early childhood may become less overt over the course of development but never go away (e.g., Subbotsky, 2000; Gelman, 2003; Bloom, 2004; Kelemen and Rosset, 2009; Hood, 2009, for review).

The current investigation draws together these diverse findings on automatic moral evaluations and on the intentionality bias to make a prediction about a cognitive and social response to unequal distribution. Given that maintaining fairness is important to our social lives, that moral decisions are often made before weighing the evidence, and that under pressure, adults default to assumptions of intentionality to explain ambiguous actions, the current study investigates whether unfairness is so objectionable that even unambiguously unintentional acts of inequality are tacitly judged to be intentional. That is, while Rosset (2008) found that asking people to respond at speed reveals people’s underlying propensity to judge ambiguous actions as intentional, we hypothesized that the salience of unequal outcomes might be such that – even absent processing constraints – unambiguously accidental acts resulting in unequal outcomes might be viewed as acts of intentional unfairness. In particular, we were interested in how such tacit assumptions might affect social behavior: That is, if I witness someone receive less than the person next to them – regardless of the actual circumstances – will I automatically assume intentional unfairness and avoid the “perpetrator.”

Specifically, across three minimally different conditions, participants saw one of three movies. In the first movie, Person A, and then Person B, engage
in acts that accidentally inconvenience Lauren. When asked to decide who they would prefer to interact with again in the future, we predicted that participants would have no preference for either individual as both acts were explicitly accidental and, thus, equally likely to occur in future behavior (“accidental baseline condition”). The second video condition was the same – Lauren was accidentally inconvenienced by Person A and then Person B – except for one key difference: An additional person, Michelle, sat alongside Lauren. She was also accidentally inconvenienced by Person A, but not by Person B. With the simple introduction of this social disparity, we predicted that participants’ evaluations of Person B’s accidental actions would flip from accidental to intentional. That is, we predicted that an act previously judged as accidental would now be judged as intentional, solely because the presence of an additional person – Michelle – altered Person B’s accidental act to a case of inequity (“accidental inequity condition”). We also included a third, “intentional unfair control condition” where Person B was intentionally unfair to Lauren. We predicted that here too that participants would choose to avoid Person B so as to avoid experiencing her intentional unfairness again in the future. Study 1 explored these predictions with adults and Study 2 tested them with elementary school-aged children.

**Study 1**

**Method**

**Participants.** 48 undergraduates (30 females, mean age 19 years, range 18–22 years) who were tested individually.

**Materials.** Three movies in which two experimenters (Persons A and B) asked interviewees to complete questionnaires in context of their voluntary participation in a research study and then rewarded them with university-logo mugs. Photographs of the two experimenters from the movie were displayed next to the monitor.

**Design and Procedure.** Participants were randomly assigned to one of three conditions: In the movie for the “accidental baseline condition”, participants saw an experimenter (Person A) introduce herself and ask the interviewee, Lauren, to perform the task of filling out a questionnaire for a research study. In this first scene, Person A was associated with an accidental negative outcome in that she offered a questionnaire that was damp and hard to fill out because she had accidentally left it by an open window. Importantly, she
explicitly made clear that leaving the questionnaires by the window was an accident. After Lauren had completed the questionnaire, Person A rewarded her with two mugs in factory packaging and then left for the day.

In a second scene – shot to convey the passage of time – a second experimenter (Person B) introduced herself and asked Lauren to fill out another questionnaire. On completion, Person B thanked her and rewarded her with two mugs in factory packaging and then packed up and left for the day. After Person B had gone, Lauren unwrapped the mugs; however, she discovered one mug was chipped. Importantly, in this movie, as in the movies of all conditions, both experimenters were associated with accidental negative outcomes (Person A offering damaged, awkward-to-complete surveys; Person B offering a defective reward). Extensive pilot work had established that students viewed these as equivalently negative outcomes, unlike an earlier iteration of the study in which Person A’s time-wasting error of inadvertently delaying the interviewee’s departure was judged to be far more egregious than Person B’s defective mug. Additionally, steps were taken to include practical details in the film to underscore that Person B’s actions were entirely accidental and that she would not have known that she had left a defective reward: For example, all reward mugs arrived in original factory packaging so the chip could not have been previously discovered. Person B also packed up and left for the day before Lauren discovered her chipped mug and was, therefore, unable to witness or rectify any mistake. This detail was important because earlier pilot work revealed that, even in the accidental baseline condition, Person B was judged intentional and blameworthy if she inadvertently gave a chipped mug but was present at its discovery and, therefore, knew her responsibility to make it right.

The “accidental inequity condition” movie was similar to the “accidental baseline condition”, movie but differed only by involving two interviewees, Lauren, and an additional person, Michelle. It, therefore, introduced the social comparison central to assessments of equality. In this version, Person A gave Lauren and Michelle damp questionnaires and rewarded them each with two mugs and left for the day. Person B then rewarded the interviewees with two mugs for completing a further questionnaire and left for the day. Lauren later discovered that one of her mugs was chipped. Lauren and Michelle were able to see each others’ mugs at the time of this discovery (and, thus, make a comparison). Once again, in this movie then, both experimenters were associated with an accidentally negative outcome but this time one of them was also associated with an unequal distribution of reward.

The “intentional unfairness control condition” movie was exactly the same as the “accidental inequity experimental condition”, except that in the second
scene, Person B’s actions constituted deliberate unfairness: Thus, after Person A accidentally gave damp questionnaires and rewarded both interviewees with two mugs, they completed a second set of questionnaires for Person B who then rewarded Michelle with two packaged mugs and Lauren with only one. In consequence, in this movie, both experimenters were associated with a negative outcome but only one of them was associated with an intentionally unfair distribution of reward.

At the conclusion of both these movies, we asked two memory questions: (1) What was the name of the first experimenter?; (2) What was the name of the second experimenter? Performance was at ceiling. We then asked two test questions: (1) If you were participating in a study like that, who would you rather work with?; (2) Who do you think she (the interviewee with the chipped mug) would rather do a follow-up interview with? Again, we chose this indirect measure because we were interested in the social consequences of tacit evaluations and also because directly asking participants to reflectively judge whether an act was intentional would have encouraged rational decision making and the suppression of any intentional interpretation. After the experimental questions, participants were encouraged to comment on the actors and what they had seen in the video. Each participant understood the procedure and no participant was excluded from analyses.

Participants indicated their answers using the headshots on either side of the computer. On test questions, participants were given a score of 0 if they selected Person B and a score of 1 if they selected Person A. Thus, higher scores indicated greater avoidance of the “unfair” experimenter (0–2 range). Table 1 presents a schematic for the 3 conditions. In all conditions, the roles were counterbalanced by creating versions of movies where the actors switched roles. All movies were shot at an angle so that actors’ facial responses to events were obscured.

Results

To recap, we hypothesized that all acts resulting in unequal reward distribution are automatically tacitly construed as intentional unfairness. We predicted that there would be no preference for experimenter in the “accidental baseline condition”, given that the outcomes associated with Person A and Person B were both mildly negative. However, we predicted that avoidance of Person B might be significantly higher in the “accidental inequity experimental condition”, given that the outcome associated with Person B was now also unequal. Avoidance of Person B was expected to be significantly higher in both the accidental inequity and the intentional fairness conditions than in the “accidental baseline condition”.

Results can be seen in Fig. 1. A one-way analysis of variance (ANOVA) on participants’ tendency to favor Person A over Person B across the three conditions was significant, $F(2, 45)=10.91, p<0.0001$. As predicted, a $t$-test indicated that subjects had no preference for Person A or B in the Baseline ($t(14)=-1.29, p=0.22$). A one-way ANOVA indicated that they avoided Person B significantly more in the “accidental inequity condition” ($M=1.53, SD=0.62$) and in the “intentional unfairness condition” ($M=1.75, SD=0.45$) than in Baseline ($M=0.73, SD=0.80$), Fisher’s LSD tests, $p<0.01$. In a result that provides strong support for the notion that accidental acts involving inequity are construed as intentional, avoidance of Person A in the “accidental inequity condition” ($M=1.53, SD=0.62$) and in the “intentional unfairness condition” did not differ, ($M=1.75, SD= 0.45$), Fisher’s LSD, $p=0.32$.

**Discussion**

The findings of Study 1 support the hypothesis that when an accidental outcome results in inequity, people treat the person who was unwittingly responsible as if she were guilty of intentional unfairness. Given these findings, we were interested in the development of the bias to view all unequal outcomes as intentional.
Study 2 explored this question with 7- and 8-year-old elementary school-aged children – an age group selected because, unlike younger children, they are accepted as having a robust grasp of the distinction between accidental and intentional acts (e.g., Astington, 2001) and also recognize, and respond negatively to, unequal distribution (Murnigham and Saxon, 1998; Benenson et al., 2007; Fehr et al., 2008).

**Study 2**

**Method**

**Participants.** Forty-seven 7- and 8-year-old English children participated in the study (21 females; mean age 7 years, 4 months; range 7 years to 8 years, 3 months). Family income ranged from low to moderate. Children received a small gift for participating.

**Materials.** Participants saw one of three child-appropriate movies structured as in Study 1: Two teachers asked children to draw signs for classroom doors

Figure 1. Mean number of times participants chose Person A, by condition.
and then rewarded them with pens. Photographs of the two teachers from the movies were displayed next to the monitor.

**Design and Procedure.** The “accidental baseline condition” involved only one child. Teacher A asked a 7-year-old girl to draw a sign but she accidentally caused a negative outcome by forgetting to tell her that the sign had to be drawn in red: the girl, therefore, had to start all over again. Teacher A rewarded the girl with two new pens for drawing the sign. Teacher B also rewarded her with two pens for drawing the sign but, after Teacher B had departed, one of these was later found not to work.

In the “accidental inequity experimental condition”, Teacher A asked two 7-year-old girls to draw signs and she accidentally caused a negative outcome by forgetting to tell them that the signs had to be drawn in red: they therefore had to start all over again. Afterwards, Teacher A rewarded them with two new pens each. Later, Teacher B also asked the children to draw signs and rewarded them with two new pens each. However, after she had left, one girl discovered that one of the pens does not work.

In the “intentional unfair control condition”, Teacher A behaved as above. However, after the children had drawn signs for Teacher B, she intentionally gave one child two pens and the other only one.

Children were asked two memory questions as in Study 1. One child answered incorrectly and was excluded from analyses. Children were then asked two test questions: (1) “If you had to help one of the teachers, who would you help?” (2) “She (the girl who received the broken pen) has to help out one of these teachers the next day. Who do you think she will want to help?” Participants were given a score of 0 if they selected Teacher B and a score of 1 if they selected Teacher A. A higher score indicated greater avoidance of the “unfair” teacher (0–2 range). Like adults, children were encouraged to comment on what they had seen. They all understood the procedure and no participant was excluded from analyses.

**Results**

Results are shown in Fig. 1. A one-way ANOVA comparing children’s tendency to select Teacher A across conditions was significant, $F(2, 44)=5.47, p<0.01$. As predicted, and like the adults, a $t$-test indicated that children had no preference for Teacher A or B in the Baseline $t(16)=0.44, p=0.67$. A one-way ANOVA indicated that they avoided Teacher B significantly more in the “accidental inequity condition” ($M=1.66$, SD=0.49) and in the “intentional unfairness condition” ($M=1.50$, SD=0.52) than in Baseline ($M=1.06$, SD= 0.57), Fisher’s LSD tests, $p<0.05$. In a result that provides strong support for the notion that
children also construe accidental acts involving inequity as intentional, avoidance of Teacher A in the “accidental inequity condition” ($M=1.66$, $SD=0.49$) and in the “intentional unfairness condition” ($M=1.50$, $SD=0.52$) did not differ, Fisher’s LSD, $p=0.39$.

To explore whether children and adults differed in their tendency to avoid a person associated with an accidentally unequal outcome, a 2 (age) × 3 (condition) ANOVA was conducted on choice of Person A. The analysis found an effect of condition, $F(2,94)=15.55$, $p<0.0001$, no effect of age and no interaction. As can be seen in Fig. 1, regardless of age, participants were as likely to select Person A in the intentional unfairness ($M=1.63$, $SD=0.49$), and accidental inequity experimental conditions ($M=1.59$, $SD=0.56$) and more likely to do so in both of those than in the “accidental baseline condition”, ($M=0.90$, $SD=0.70$).

Discussion

In Study 2, children showed the same pattern as adults in Study 1. Like adults, when an accidental outcome resulted in inequity, 7- and 8-year-olds responded as though it were deliberate and avoided the perpetrator.

General Discussion

The current experiments were designed to investigate humans’ aversion to unfairness by exploring whether people have a cognitive bias to automatically assume unequal distribution is intentional unfairness. We were also interested to investigate the social consequence of such a bias. Based on the literature indicating that moral judgments are often based on “gut reactions”, coupled with findings indicating that, under certain circumstances, adults default to a broad attribution of intentionality to explain ambiguous acts, we predicted that an aversion to unfairness might be so strong that people would judge even explicitly accidental acts (rather than ambiguous acts) to be intentional. Our prediction was supported with both child and adult participants. Judgments flipped from accidental to intentional, simply based on social comparison, and perception of inequality. Our results suggest that the cognitive bias to initially view all events involving an agent as intentional can be found not only under pressurized “speeded” conditions, but also when the moral imperative of “equal rewards for equal effort” is transgressed.

What are the implications of this finding for social life? The results indicate that in everyday human dealings, adults, like children, have a tendency to convict others on the basis of the consequences of their actions rather than
veridical evaluations of the intentions underpinning those actions. The most obvious implication of this finding is for the legal system: it may be more difficult for jurors to weigh intentionality if an outcome involves inequity. However, it is not only in the courtroom that people are faced with making a decision about whether an unequal outcome was intended; people tacitly make these decisions on a daily basis. For example, the current finding implies that if the waitress accidentally forgets to bring bread to my table, I am much more likely to suspect it is an intentional action if I have evidence that diners at the table next to me received theirs.

Why does inequity provoke this level of irrational response? Being habitually prone to blame and shun others who distribute resources unequally—regardless of their motives—would seem to be highly maladaptive. Surely it is better to give those who are able, and potentially willing, to distribute resources the “benefit of the doubt”. One possibility is that inequitable outcomes are particularly prone to evaluative error simply because the stakes are high: Distributional justice may be so central to maintaining stable and beneficial cooperative relations among group members in the long term that it may not ultimately cost us to hyper-actively avoid anyone whose actions, intentional or otherwise, have resulted in an inequitable outcome. This is not only because the agent may actually turn out to be a costly free rider, and so avoidance is prophylactic to potential future ills but also because an avoidance response sends out general social signals about intolerance of unfairness (see Fehr and Schmidt, 1999; Brosnan, 2006).

A number of interesting questions remain for future research. For example, would the present effects moderate if the unwitting perpetrator of inequality was obviously familiar to the victim? Prior human and non-human primate research suggests that they might (see Brosnan, 2006). Furthermore, is the moral valence of distributional inequality unique in invoking the kind of response revealed here? Using an explicit judgment task, Young et al. (2007) found that adults judged the actions of agents who inadvertently caused mortal harm (e.g., by accidentally stirring poison rather than sugar into coffee) as permissible and significantly more acceptable than the actions of someone who intentionally caused someone to die. However, there were indications in their data that the accident-prone protagonist was not entirely exculpated. She was still judged less favorably than someone who neither intended nor caused any harm. Although difficult, it would be interesting to follow-up the work of Young and colleagues with more implicit measures like those of the present study: That is, given two equivalently accident prone agents, is the one who inadvertently caused mortal harm more likely to be avoided? Would she be shunned to the same extent as someone who committed murder? Exploration
of such moral scenarios, in addition to those focused on other moral concerns proposed as the potential primitives of an innate intuitive ethics (e.g., loyalty, respect and purity) could yield informative results (see Haidt and Joseph, 2004).

In closing, the current results suggest that, for both young children and adults, inequity promotes an automatic social avoidance response that is driven by judgment of outcome rather than precursor intentions. It remains an open question how readily participants might be able to modify their gut response if asked to explicitly reflect on the intentionality and permissibility of the relevant actors’ actions (e.g., Young et al., 2007). Prior research certainly suggests, however, that even if reflections like these might assist in suppressing the overt expressions of automatic responses, the underlying intuitions simply go underground and continue to exert tacit influence (e.g., Greenwald and Banaji, 1995; Chaiken and Tropé, 1999; Bargh, 1999; Strack and Deutsch, 2004). It remains to be seen whether this is also true for attitudes to accidental unfairness.

Acknowledgements

The research was supported in part by a Clara Mayo dissertation award. Thank you to those who helped with data collection, including past and present members of the Child Cognition Lab at Boston University and the teachers and children from participating schools. This paper is based on a doctoral dissertation.

References


