Cognition 108 (2008) 771–780

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

It's no accident: Our bias for intentional explanations

Evelyn Rosset

Psychology Department, Boston University, 64 Cummington Street, Boston, MA 02215, United States

ARTICLE INFO

Article history:
Received 12 February 2008
Revised 1 July 2008
Accepted 1 July 2008

Keywords:
Intentionality
Intentional explanation
Explanatory biases
Heuristics and biases
Social cognition
Development
Action perception

ABSTRACT

Three studies tested the idea that our analyses of human behavior are guided by an "intentionality bias," an implicit bias where all actions are judged to be intentional by default. In Study 1 participants read a series of sentences describing actions that can be done either on purpose or by accident (e.g., "He set the house on fire") and had to decide which interpretation best characterized the action. To tap people's initial interpretation, half the participants made their judgments under speeded conditions; this group judged significantly more sentences to be intentional. Study 2 found that when asked for spontaneous descriptions of the ambiguous actions used in Study 1 (and thus not explicitly reminded of the accidental interpretation), participants provided significantly more intentional interpretations, even with prototypically accidental actions (e.g., "She broke the vase"). Study 3 examined whether more processing is involved in deciding that something is unintentional (and thus overriding an initial intentional interpretation) than in deciding that something is unpleasant (where there is presumably no initial "pleasant" interpretation). Participants were asked to judge a series of 12 sentences on one of two dimensions: intentional/unintentional (experimental group) or pleasant/unpleasant (control group). People in the experimental group remembered more unintentional sentences than people in the control group. Findings across the three studies suggest that adults have an implicit bias to infer intention in all behavior. This research has important implications both in terms of theory (e.g., dual-process model for intentional reasoning), and practice (e.g., treating aggression, legal judgments).

1. Introduction

When someone cuts us off in traffic, why are we so quick to assume it is intentional? Why do we praise people for fortuitous accidents (even awarding the Nobel Prize for the accidental discovery of penicillin), and blame them for their misfortune (e.g., Burger, 1981; Lerner, 1980)? After all, although children may claim that accidents and mistakes are done on purpose (see Baldwin & Baird, 2001; Miller & Aloise, 1989 for reviews), adults presumably understand the difference between intentional and unintentional acts; where a child may state that she meant to mess up a tongue twister, or jerk her knee reflexively (Shultz, Wells, & Sarda, 1980), an adult would acknowledge the unintentional nature of these actions.

The above examples, however, along with findings in moral philosophy (e.g., Knobe, 2003), social psychology (e.g., Langer, 1975; Lerner, 1980; Wegner, 2002), judgment & decision making (e.g., Baron & Hershey, 1988; Kordes-de Vaal, 1996), and recent work in the cognitive science of religion (e.g., Barrett, 2000; Boyer, 2003; Guthrie, 2001), hint that the overattributions of intention found in childhood may not be completely outgrown. Coupled with research in cognitive development (e.g., Kelemen, 2004; Preissler & Bloom, 2008) these findings helped prompt the idea explored in this paper: our explanations of behavior – those of both children and adults – are guided by an implicit bias to interpret all actions as intentional.

It is noteworthy that although scores of studies have looked at how children distinguish intentional acts (like kicking a ball) from accidental ones (like tripping over one), comparatively little research has examined the same capacity in adults. This is perhaps due to the assumption

E-mail address: erosset@gmail.com

0010-0277/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2008.07.001

that adults master the ability to distinguish intentional and unintentional acts and are therefore an uninformative population to study. This is a legitimate assumption to some degree; on an explicit level, adults recognize that kicking is intended, and that tripping is not. Indeed, when asked what makes an act intentional, adults show considerable consensus and identify several key elements (Malle & Knobe, 1997a).

What happens, however, when the task depends on more subtle behavioral measures than those used in verbal report? Although adults may recognize the unintentional nature of many actions on an explicit level, the fragility of this understanding could be revealed using tasks that involve more implicit processing. Intentional explanation, in other words, may best be understood using a dual-process model, where intention is inferred both by being activated on an automatic, nonconscious level, as well as by more deliberate reasoning. Indeed, considering the importance of dual-process models in advancing the research in so many domains of cognition, formulating such a model for intentional explanation may be of significant value.

The cognitive mechanism at the core of this proposal is an "intentionality bias:" when evaluating an agent engaged in an action, an intentional interpretation is automatically activated (Rosset, 2007). In short, by default we view *everything anyone ever does as intentional*. It is the ability to override, and thus inhibit, this bias that develops with age, not the bias itself. This is not to say, of course, that adults labor under the illusion that sneezes, or stumbles, or slips of the tongue, are intentional. It is to argue, however, that adults come to this interpretation secondarily after overriding an initial interpretative impulse, the inhibition of which is only possible because they have additional knowledge. Such knowledge includes an understanding of alternative causes for the action (e.g., sneezes are due to allergies), an understanding of social norms and people's beliefs, goals, and desires (e.g., people do not generally want to trip), and an understanding of behavioral cues (e.g., a look of embarrassment indicates an undesired outcome).

Models based on implicit, automatic intentional inference have, of course, been proposed before (e.g., Baron-Cohen, 1995; Premack, 1990). The present proposal is meant to complement such proposals, but carries an important difference: instead of focusing on how intention is inferred, this research focuses on how it is inhibited. Consequently, instead of focusing on infants, this research focuses on adults. Underlying this research, therefore, is the assumption that a mature understanding of behavior, what separates children from adults, depends not on differences of intentional *inference*, but rather on differences of intentional *inhibition*.

According to the intentionality bias every action is judged to be intentional until proven otherwise. An accident is deemed to be an accident only by replacing the initial intentional interpretation with a non-intentional one. For instance, if asked whether an act like "setting a house on fire" is intentional or accidental, according to the intentionality bias our initial reaction is that it is intentional. This automatic inference of intention, however, may subsequently be inhibited by additional knowledge such as an understanding of human fallibility (e.g., forgetting a burning cigarette), social norms (e.g., burning houses are generally undesired), and flammability (e.g., of curtains and wood). Therefore, although an intentional interpretation is initially activated, after a beat of consideration we may decide that such an act is accidental. Importantly, however, if time constraints limited our access to our knowledge of fallibility and flammability, the initial interpretation of this act – and any act for that matter – would be that is was intended. The motivation of Study 1, therefore, was to find a way to tap into this initial interpretation using a speeded judgment task.

2. Study 1

Study 1 asked people to decide whether a series of actions were best characterized as done on purpose or by accident, and compared the responses of people who had to do so quickly – and therefore had less time to override the intentionality bias – to the responses of people who were not so rushed. The judgments of interest were those concerning *ambiguous* actions, actions that could be done either on purpose or by accident (e.g., "He set the house on fire").

Importantly, however, although these ambiguous actions *could* be done on purpose, we were particularly interested in actions that are generally done by accident. For instance, an action like "He deleted the email" could be performed both intentionally or accidentally, but it is frequently done on purpose. The finding that when speeded, people are more likely to judge such actions to be intentional is of limited interest as such differential responding could be due to simple associative mechanisms, the effects of which are more apparent in a speeded task.

The same cannot be said, however, for prototypically accidental actions such as "He hit the man with his car" where, if intentional inference depended exclusively on a probability analysis, one would expect increased "by accident" judgments when speeded. Responses on typically accidental actions therefore served as an especially stringent test of the intentionality bias; if intention were not inferred via such a mechanism, one would expect that deciding that a prototypically accidental action is done on purpose would require *more* time than judging it to be done by accident.

2.1. Method

2.1.1. Participants

Participants were 90 undergraduate students taking psychology classes at a large northeastern university. Participants were randomly divided into either the experimental condition (i.e., speeded; $n = 48$) or the control condition (i.e., unspeeded; $n = 42$).

2.1.2. Procedure

Participants sat around a large table, in groups of 5–10, in a classroom setting and read through the instructions with the experimenter. The instructions indicated that par-

ticipants would see a series of sentences appear one at a time on an overhead screen, and that they were to decide whether the action described in the sentence was generally done "on purpose" or "by accident" by checking the appropriate box on an answer sheet in front of them. Examples of sentences done on purpose ("She set the table") and by accident ("She broke her leg skating") were provided, along with the note that some sentences could fit both descriptors ("She scratched herself"). In these cases, participants were instructed to decide which description *best* characterized the action.

The sentences were presented consecutively, in one of two orders, using PsyScope programming software (Cohen, MacWhinney, Flatt, & Provost, 1993) that ensured accurate timing between each sentence. Each sentence remained on the screen for either 2400 ms (speeded condition) or 5000 ms (control condition) after which the next sentence appeared automatically. A pause, indicated by an "∗" and ended by the experimenter's keypress, was inserted every eight sentences; this was designed to prevent people from losing their place due to a single missed item, as well as to give people time to turn the pages of their answer sheets without missing items. The stimuli were divided into 10 blocks of eight sentences each. Within each block were four control sentences (two intentional and two accidental, explained below) and four test sentences. The order of the sentences within each block was determined randomly by PsyScope and was then presented in a fixed order to all participants. Three blocks of practice items preceded the test items to give participants a sense of how fast judgments were to be made. Everyone was debriefed at the end of the experiment.

2.1.3. Stimuli

The stimuli consisted of short sentences describing simple actions: 34 test sentences and 40 control sentences.[1] The test sentences described actions where the intentions of the actor are ambiguous, and which could be done either on purpose or by accident (e.g., "She stepped in the puddle"). The control sentences included two types: *intentional control sentences* and*accidental control sentences.* The intentional control sentences described unambiguously intentional actions (i.e., actions that are always done on purpose, such as "She proofread the paper"), whereas the accidental control sentences described unambiguously accidental actions (i.e., actions that are always done by accident, such as "He fell down the stairs"). The control sentences were included to track reading ability, as well as flag any indiscriminate responders.

It should be noted that although the intentionality bias makes no prediction of differential responding with the intentional control sentences because both groups' initial interpretation would be that it is intentional, it does, *in theory*, make the prediction that the people in the speeded group may be more likely to judge the accidental control sentences as intentional. Such differential responding, however, was not expected for these sentences; given the obviousness of the sentences, coupled with the participants' age and experience, it is likely that the amount of additional time needed to override the intentionality bias would be of a degree too small to be measured by the present paradigm.

To ensure that the majority of the sentences described prototypically accidental actions, a pilot study asked a separate group of participants ($n = 26$) to make intentionality judgments about each of the sentences by reading through the list of 74 sentences directly on the answer sheet and checking the appropriate box. The percentage of people who checked "on purpose" was calculated for each sentence, yielding an "intentionality likelihood rating" for each sentence. The number of test sentences judged to be accidental by at least 60% of participants (i.e., with an intentionality likelihood rating of 40 or below) was calculated. Using this criterion, two-thirds of the sentences (i.e., 22/34) were categorized as *Accidental Test Sentences.* This preliminary study also allowed verification that the control items were judged to be unequivocally on purpose or by accident, with over 99% of participants judging the intentional control sentences to be generally done on purpose, and the accidental control sentences to be generally done by accident. See Table 1 for a full list of stimuli, including the pilot judgments on the test sentences.

The sentences were designed to be roughly the same length, and simple enough to be read within the allotted time. Additionally, the intention of the action never depended on the final word of the sentence (e.g., "She buttoned her shirt unevenly"), a precaution taken to ensure that any difference between groups was not due to being unable to read the entire sentence. In the control sentences, the intention of the action was often coded in the verb itself (e.g., *fell* can only be done by accident, and *proofread* can only be done on purpose); this was, of course, never the case for the target sentences (e.g., *broke, ripped, stepped* can all be done both on purpose or by accident).

2.2. Results

Data from five participants were not included in the analysis for failure to follow instructions (i.e., marking both boxes; $n = 3$), and for leaving more than 25% of the items blank ($n = 2$). The data from 85 participants were therefore included in the analysis.

An "intentionality endorsement score," defined as the number of "on purpose" responses divided by the number of possible responses, was calculated for each participant for each item type (test sentences, intentional control sentences, accidental control sentences). The intentionality endorsement score therefore represented each person's tendency to choose an intentional interpretation. An overall high rate of accuracy on the control items (97% overall) showed that people understood the task and were not subject to response bias or limited by reading ability. Not surprisingly, the responses were not normally distributed; they were skewed to the right for the intentional control

[1] Six additional sentences were included as pilot sentences for a separate study, bringing the total number of sentences to 80. These described unintentional actions that are due to chance (e.g., "She ran into a childhood friend"). These sentences, determined prior to testing, were not included in this analysis because they are not ambiguous in the same way the test sentences are; one does not experience good or bad luck either on purpose or by accident.

Table 1
Study 1 stimuli, with intentionality likelihood ratings (ILR) from pilot study

Control sentences		Test sentences			
Unambiguously accidental	Unambiguously intentional	Prototypically accidental	ILR	Neutral/prototypically intentional	ILR
He poked himself in the eye	He buttoned his jacket	He hit the man with his car	0	She cut him off driving	46
He sneezed from allergies	He drew a picture of the beach	He gave her the wrong change	0	The boy knocked over the sand castle	46
He stubbed his toe	He erased the scribbles	She burnt the meal	0	She walked by without saying hello	46
He failed the driving test	He folded the letter carefully	She broke the vase	0	He took an illegal left turn	50
He fell down the stairs	He listened attentively	He tracked mud inside	0	He ripped the piece of paper	57
He fell off the skateboard	He shaved in front of the mirror	He forgot his homework	4	She sprayed him with water	69
He missed the hoop with the ball	He threw the football	He arrived 5 min late for class	4	The man left without leaving a tip	73
He pinched his fingers in the door	He typed the email	He bumped into a classmate in the hall	4	She made a mark on the paper	77
He broke his tooth playing hockey	He vacuumed the carpet	He broke the window	4	She drove over the speed limit	81
She tripped on the curb	She addressed the letter	The painter inhaled the fumes	8	He deleted the email	88
She caught a cold	She baked a cake	He drank the spoiled milk	8	She ignored the question	88
She lost her keys	She changed the flat tire	She woke the baby up	15	She averted her eyes	92
She broke her cell phone	She followed the recipe	He stepped in the puddle	15		
She slipped on the ice	She looked for her keys	He set off the alarm	15		
She blushed from embarrassment	She painted her toe nails	He jumped when the bell rang	15		
She burned her hand on the stove	She proofread her paper	He dripped paint on the canvas	23		
She jumped back in surprise	She studied for the exam	She kicked her dog	27		
She tripped on the jump rope	She threaded the needle	She left the water running	27		
The boy hiccupped	The boy smiled for the picture	He set the house on fire	31		
The girl had a seizure	The girl lit the candle	He ate the bruised part of the apple	31		
		She told the same joke twice	35		
		The girl popped the balloon	38		

sentences (towards a proportion score of 1), and to the left for the accidental control sentences (towards a proportion score of 0). They were also slightly skewed to the left for the test sentences, prompting the use of Mann–Whitney U comparisons for all analyses. Mann–Whitney U comparisons yielded no significant differences across the two different orders, and the lists were therefore collapsed.

To examine whether the speeded and control conditions differed on their tendency to judge whether actions were generally done on purpose, comparisons were made on each of the item types. Man–Whitney U comparisons revealed no significant differences across conditions for the intentional control sentences ($U = 772$, ns), with both conditions judging an average of 98% of them to be generally done on purpose. There was, however, a significant difference for the accidental control sentences ($U = 652$, $p < .02$); 2% were judged to be generally done on purpose in the unspeeded group, whereas that number rose to 5% for the speeded condition. Most relevant to the present analysis, however, was the comparison between the two conditions for the test sentences, particularly the Accidental Test Sentences. Comparing all the test sentences revealed a significant difference ($U = 605$, $p < .01$); in the unspeeded condition, 33% of the items were judged to be on purpose, whereas 39% were judged to be on purpose by people in the speeded condition. When focusing exclusively on the Accidental Test Sentences, the effect was more pronounced ($U = 527$, $p < .001$), with 15% of these actions being judged to be generally on purpose in the control group, vs. 22% in the speeded group. As mentioned earlier, this is the opposite pattern of results one would expect if people inferred intention through a simple probability analysis. In fact, further analyis found that although the pattern of responding on the Neutral/Intentional Test Sentences was similar to the Accidental Test Sentences, with the speeded condition judging more of them to be intentional (69% vs. 66%), the difference between groups did not reach significance for these items ($U = 832$, ns). See Table 2.

2.3. Discussion

The present results support the predictions based on the intentionality bias. When deciding whether a given action is done on purpose or by accident, people were more likely to judge that it is done on purpose when they had

Table 2
Results for Study 1, comparing % of sentences in each condition judged to be on purpose

	Control sentences		Test sentences		
	Intentional	Accidental	All	Accidental	Neutral/Int.
Unspeeded	98	2	33	15	66
Speeded	98	5	39	22	69
	ns	$p < .02$	$p < .01$	$p < .001$	ns

to make their decision quickly, suggesting that the initial interpretation of the action is that it is intentional. These results are particularly strong considering that the difference across conditions was greatest for the test items that described the most prototypically accidental actions.

Related to this, a significant difference was found for the accidental control sentences, despite the fact that these sentences described actions that are *always* done by accident. Although both conditions judged these to be overwhelmingly done by accident, there were, nonetheless, significantly more "on purpose" judgments in the speeded condition. Evidence of such differential responding in a task of this nature, which does not rely on particularly sensitive measures such as response latency, attests to the strength of the bias to interpret all actions as intentional by default. These findings were explored more in Study 3.

The pattern of responses on the accidental control sentences also highlights the importance of group differences, rather than absolute numbers, in this study. One could argue that an overall hit rate of 36%[2] for "on purpose" responses for the target items is rather low in light of the claim that people interpret all acts as intentional. However, as described earlier, the test sentences were selected to be prototypically accidental, as these sentences may provide stronger evidence for an intentionality bias. For instance, if the stimuli were selected to include exclusively highly prototypically accidental actions, like "He hit the man with the car", we would expect very low "on purpose" hit rates for both conditions, but nonetheless find that people in the speeded group were significantly more likely to endorse such an interpretation than people in the unspeeded group. Indeed, this is what was found for the accidental control sentences. Therefore, the actual number of "on purpose" responses is less relevant in this task than the differential responding between conditions.

Although the Accidental Test Sentences provided the strongest evidence of the intentionality bias, it is odd that the difference between groups on the other test sentences – those that were prototypically intentional or neutral (e.g., "She sprayed him with water") – was not stronger. Although participants in both conditions were presumably processing the sentences in the same way (i.e., accepting their initial intentional interpretation), this explanation holds for only two-thirds of the sentences (on average 66% and 69%). This finding merits further research with more sensitive measures; for example, would reaction time data time reveal stronger effects with more prototypically intentional sentences, or with more prototypically accidental sentences?

Anecdotally, after debriefing, several people made comments along the lines of: "Funny, because when I read the sentence 'She kicked her dog' I pictured it being done on purpose, but then realized that it's probably *generally* done by accident, so I checked that box." This prompts the question: what if people had simply been asked what image comes to mind when they read a given sentence? If people had not been asked whether the action was generally done on purpose or by accident, would the same actions have generated more intentional interpretations?

In other words, the nature of the task itself may have primed people to think of an accidental interpretation that they may not have otherwise considered. The intentionality bias maintains that people do not consider accidental interpretations unless presented with enough cues to override their initial intentional interpretation. In Study 1, however, participants *had* to think about the accidental interpretation of each action; after all, for each item they were asked whether it was best characterized as generally done on purpose or by accident!

Changing the instructions, therefore, provides another way to test the intentionality bias. If we did not have a bias to interpret everything as intentional, there would be no reason to expect that reminding people of both interpretations would influence their judgments; if our judgments were simply reflections of the likelihood with which a given act is performed on purpose or by accident, we would not expect spontaneous descriptions to differ from forced-choice judgments. If, on the other hand, we have a bias to interpret acts as intentional, we would expect people to provide accidental interpretations less frequently if not instructed to consider both interpretations. This was further explored in Study 2.

3. Study 2

When people are simply asked what image comes to mind when presented with an ambiguous action – even an action that is typically accidental – do they nonetheless first think of the intentional interpretation? Would they, in other words, be less likely to provide an accidental interpretation if not explicitly reminded of the possibility of one? The prediction based on the intentionality bias is that there would be such a difference; unless specifically directed to consider an accidental interpretation, people would be less likely to.

3.1. Method

3.1.1. Participants

Participants were 92 undergraduate students taking psychology classes at two northeastern universities, divided into either the "Open-ended" group ($n = 40$), or the "Reminded" group ($n = 52$).

3.1.2. Procedure

This task was completed online. Participants were given access to an online form where they found the instructions, the stimuli, and the debriefing. For the Open-ended group, the instructions indicated that they were going to read a list of sentences, and that after each sentence they were to "write a brief description of the image that comes to mind" when reading the sentence. Two examples were provided, one of a sentence describing an accidental act (*She tripped on the curb*; "I see a girl on Comm. Ave. talking on her cell phone and tripping over the curb as she crosses the street."), and one of a sentence describing an

[2] This is the percentage for both conditions, for all 34 target sentences. People in the speeded group endorsed 39% of these items as intentional, on average, and people in the unspeeded group endorsed 33% of these items as intentional, on average.

intentional act (*He looked for his key*; "A guy is in front of his car and he's feeling his pockets for his keys, looking confused."). Each sentence was followed by a textbox where participants could write their responses.

After completing the descriptions for the full set of sentences, participants were instructed : "Important! Don't hit submit yet!?For the next part, we'd like you to go back to each of your responses and clarify whether the event you described was done on purpose or by accident. Please simply write the words "on purpose" or "by accident" after your description." This step was included to clarify the coding of any descriptions that may have been ambiguous as to the intentions of the actor. For instance, given the sentence "He dripped paint on the canvas," and a response like: "I see a guy in overalls holding a paint brush and looking down at a large canvas on the floor in a loft like building," it is not clear whether the writer interpreted the act as intentional or not. Importantly, to avoid priming participants to consider an accidental interpretation they may have not considered, people were not asked for this clarification until after they had completed their Open-ended responses for all of the sentences.

The procedure was the same for the Reminder group, except the instructions were reversed: instead of first being asked to describe what image comes to mind, and then indicating their intentionality judgments, this group was asked first to indicate whether each sentence described an action that was generally done on purpose or by accident, and then write a brief description.

3.1.3. Stimuli

The stimuli consisted of the 34 ambiguous test sentences from Study 1, presented in one of the two fixed orders used in Study 1.

3.2. Results

Data from one participant was discarded for failure to follow task instructions. For each of the remaining 91 participants, there were 34 questions, each of which included a description and an intentionality judgment. Each response was scored a "1" if the interpretation was that it was done on purpose, and a "0" if the interpretation was that it was done by accident. Of the 3094 possible descriptions (91 participants $\times$ 34 items), 5.5% (171/3094) were uncodable due to vagueness of description or misinterpretation of the sentence.

As in Study 1, each participant was given an "intentionality endorsement score," defined as the number of intentional interpretations divided by the number of possible responses. Again, as in Study 1, the responses were not normally distributed; Mann–Whitney U comparisons yielded no significant differences across the two different orders, and the lists were therefore collapsed.

To see whether people offered fewer accidental interpretations if not explicitly reminded of the accidental possibility of the action, Mann–Whitney U comparisons were performed, yielding a significant difference ($U = 479$, $p < .0001$). In the Reminded group, 36% of the items were judged to be done on purpose, whereas in the Open-ended group, 45% of the actions were judged to be done on purpose. Furthermore, comparing only the prototypically accidental actions (the Accidental Test Sentences) strengthens this effect ($U = 504$, $p < .0001$), with two-thirds as many sentences judged to be intentional when people were not reminded of the accidental possibility of the action (25% vs. 15%). See Table 3.

3.3. Discussion

The present results show that when not explicitly reminded of the accidental possibility of a given act – even those that are usually done by accident – people are significantly more likely to provide an intentional interpretation. When asked whether a given action such as "He broke the window" is best characterized as being done on purpose or by accident, most respondents (92%) said it was generally done by accident. However, when asked what image comes to mind when reading the same sentence, almost half the respondents (46%) provided an intentional interpretation (e.g., "I see an angry man smashing a window with his fist").

The results of these two studies suggest that our initial interpretation of any action is that it is intentional. Adults manage to override this automatic inference with additional information related to the accidental possibility of the act (e.g., social norms, alternative causes). The additional processing required to override the intentionality bias was demonstrated in Study 1 by the increased tendency of people in the speeded condition to endorse intentional interpretations, and in Study 2 by the increased tendency of people to provide intentional interpretations when not reminded of the possibility – indeed in the majority of cases, of the *probability* – that the act may be performed by accident. Although the items of interest in the first two studies described ambiguous actions that were generally done by accident, in its strongest form, the intentionality bias maintains that *all* acts are initially judged to be intentional, even those that are *never* done on purpose. This was the motivation for the third study.

4. Study 3

The aim of the third study was to investigate whether people require additional processing to decide that obviously accidental acts are indeed accidental. Such a question may even sound silly; why would establishing that an accident is accidental require more processing than establishing that an accident is intentional? Although it sounds counterintuitive, such is the prediction of the intentional-

Table 3
Results for Study 2, comparing % of sentences in each condition judged to be on purpose

	Test sentences	
	All	Accidental
Reminded	36	15
Open-ended	45	25
	$p < .0001$	$p < .0001$

ity bias; *all* acts are judged to be initially intentional, even those that are *always* accidental. The results of Study 1 provided some preliminary support for this; when speeded, people judged significantly more of the accidental control sentences (e.g., "He tripped over the curb") to be done on purpose. These were unexpected results; we did not expect to capture the additional processing needed to override the intentionality bias using the methodology of Study 1. Furthermore, given the small difference in actual means between conditions, this finding deserved replication. The purpose of the final study was to demonstrate through more sensitive measures that even actions that are never done on purpose require additional processing to override the initial intentional interpretation activated by the intentionality bias.

It should be emphasized that it is the processing involved in *overriding* the intentional inference that increases computational load, not simply processing acts which happen to be accidental. In other words, it is evaluating the intentions of the actor, and inhibiting an intentional inference when necessary, that requires additional processing; the intentionality bias makes no prediction regarding the processing involved in evaluating intentional and unintentional actions on any other dimension (e.g., how pleasant it is, or whether the actor is male or female).

The additional processing required for overriding the automatic intentional inference was tested in Study 3 using a sentence recall paradigm. Memory research has consistently shown a strong relationship between processing load and recall; specifically, increased processing leads to increased recall (e.g., Shiffrin & Atkinson, 1969; Lockhart & Craik, 1990). If judging an act to be accidental requires more processing than judging an act to be intentional, this makes the prediction that there would be increased recall for the accidental acts. Importantly, however, this increased recall is expected only when making an intentionality judgment, but not when making a judgment on a different dimension such as – in this case – how pleasant the act is.

Including a comparison with people who made a different kind of judgment served as a control group, and was important for two reasons. First, if we found that people do remember more unintentional sentences, the argument could be made that these sentences were better remembered simply because they were more salient – perhaps they were more unusual, or more negatively valenced – but not because they required an additional processing step of overriding the default interpretation. If this was the case, however, the unintentional sentences would be better remembered by people making both types of judgments. Second, if participants showed increased recall for the unintentional sentences, one could argue that the unintentional sentences were better remembered simply because it takes more processing to judge a sentence to be *not x* (e.g., not intentional) than it does to judge something to be *x* (e.g., intentional). The prediction based on the intentionality bias is not that the unintentional sentences are better remembered; rather it is that deciding that a given sentence is unintentional increases recall more than deciding that the same sentence is unpleasant.

4.1. Method

4.1.1. Participants

Participants were 76 undergraduate students taking psychology classes at a large northeastern university. This included 37 participants in the experimental group (intentionality judgment) and 39 participants in the control group (pleasancy judgment).

4.1.2. Procedure

Participants performed the task as a large group in a classroom setting, where they were given a list of 12 sentences and were asked to make a judgment about each sentence by checking one of two boxes next to each sentence. For the experimental group, participants were instructed to decide whether the action described in each sentence was best characterized as "intentional" or "unintentional." For the control group, participants were instructed to decide whether the action described in each sentence was best characterized as "pleasant" or "unpleasant." We chose to use the term "unintentional" instead of "accidental" simply because it is difficult to think of actions that are both pleasant and accidental, whereas actions that are both pleasant and unintentional are more plausible.[3] After making the judgments, participants were told to turn their sheet of paper over, and were asked a distracter question regarding a recent film. After the brief distraction, participants were asked to write down all the sentences they could remember from the original list. Everyone was debriefed at the end of the experiment.

4.1.3. Stimuli

The stimuli consisted of 12 sentences: three sentences from each of four categories: intentional pleasant (e.g., "He ate an ice cream cone"), unintentional pleasant (e.g., "She found a penny"), intentional unpleasant (e.g., "She changed the flat tire"), unintentional unpleasant (e.g., "He dropped his glass of milk"). See Table 4.

4.2. Results

The number of intentional and unintentional sentences (both full sentences and sentence fragments) was calculated for each participant. Responses were coded as a full sentence if both verb and object were correct, even if the gender of the pronoun was changed (e.g., "She (sic) ate an ice cream cone" was coded as a full sentence). Responses were coded as a sentence fragment if the verb or object was missing or changed (e.g., "He dropped the milk"

[3] The fact that there are so few "pleasant" accidents is telling: one particularly salient way to override the intentionality heuristic is by noticing that an action produced an undesired outcome. Indeed, this may be how children understand that not all actions are intentional; they may initially interpret an undesired outcome as intentional ("Okay, I guess I wanted to knock over the juice"), but over time realize that an outcome and an intention are not identical. As mentioned in the introduction, fortuitous accidents like the discovery of penicillin are rewarded as if they were intentional.

Table 4
Stimuli used in Study 3

	Pleasant	Unpleasant
Intentional	She baked a chocolate cake	She took the exam
	He drew a picture	He changed the flat tire
	He ate an ice cream cone	She washed the dishes
Unintentional	He won the lottery	She caught a cold
	She bumped into an old friend	He lost five dollars
	She found a penny	He dropped his glass of milk

instead of "He dropped his glass of milk" or "eating ice cream" instead of "He ate an ice cream cone" were coded as fragments). Full sentences were given a score of 1, and sentence fragments were given a score of .33. The sum of full sentences and sentence fragments was calculated for each participant for each sentence type, resulting in a "total recall score" for intentional sentences and unintentional sentences.

To examine whether people in the experimental group recalled more unintentional sentences than did people in the control group, a repeated-measures analysis of variance, with item type as the within-subject factor, and judgment type as the between subject factor was performed, yielding a significant interaction, F (1,71) = 11.377, p = .001.[4] Paired-samples t-tests explored the means further and found that people in the experimental group recalled more unintentional sentences (mean = 3.5, sd = 1.2) than intentional ones (mean = 2.8, sd = 1.1), $p < .01$, whereas people in the control group remembered more intentional sentences (mean = 3.8, sd = 1.0) than unintentional ones (mean = 3.3, sd = 1.1.), $p < .05$. These results suggest that the increased recall demonstrated by the experimental group was not due to anything inherent in the unintentional sentences themselves, but rather due to the judgment required.

To explore the possibility that people in the experimental group remembered more unintentional sentences simply because judging something to be *un-x* (i.e., unintentional) required more processing than judging something to be *x* (i.e., intentional), a paired-samples t-test was conducted between pleasant and unpleasant sentences for people in the pleasancy judgment group. This group, however, did not remember significantly more unpleasant sentences (mean = 3.7; sd = 1.0) than pleasant ones (mean = 3.5; sd = 1.0), p = .52.

4.3. Discussion

The results of Study 3 showed that people remembered more unintentional sentences than intentional sentences *only* when the task at hand was to judge whether or not an act was intentional. This suggests that the processing involved in deciding that a sentence is unintentional is greater than the processing involved in deciding that a sentence is intentional. Importantly, this effect does not seem due to deciding that a given action is *un-x*, as people in the control group did not remember significantly more unpleasant sentences than pleasant ones.

Interestingly, the pleasancy judgment group also showed a difference in their recall for the unintentional sentences as compared to the intentional sentences, although this group showed the opposite trend than the experimental group, remembering more intentional sentences than unintentional ones. It is possible that this group demonstrated increased recall for the intentional sentences because *these* sentences were in fact more salient; although such a conclusion cannot be drawn from the present study, if true, the patterns of results demonstrated by the experimental group are all the more striking. This would suggest that the increased processing involved in overriding the intentionality bias is strong enough to counteract the increased recall due to increased salience evidenced by the control group.

Only people in the intentionality judgment group were expected to remember more of the unintentional sentences because only people in this group had to override the intentionality bias. People in the pleasancy judgment group did not override the intentional inference because they were not asked to judge the actions on this dimension, and people do not spontaneously override the heuristic under such conditions. In other words, for this group, these actions presumably maintained their "intentional status" and therefore did not require additional processing.

5. General discussion

The results of three studies provide support for the idea that adults have a bias to interpret all behavior as intentional until proven otherwise. Study 1 demonstrated that when time constraints limited processing, people were more likely to endorse intentional explanations for ambiguous actions. The results of Study 2 showed that without explicitly prompting people to consider the accidental interpretation of a given ambiguous action, they were more likely to consider that the action was intentional. Finally, Study 3 suggested that people require additional processing to decide that an action is unintentional, even for actions that are *always* done unintentionally!

The present proposal describes these phenomena as demonstrating a perceptual bias: when evaluating an agent and an action, an intentional inference is automatically activated as the cause of the action. It is important to note, however, that the procedure used in all three studies involved people reading sentences. The finding that less processing was involved in judging such sentences to be intentional could, therefore, be the result of a linguistic bias rather than a perceptual one. For instance, a sentence like "She woke the baby up" may imply an intentional cause due to certain linguistic cues such as an agent performing an action; when an action is performed accidentally, it may be worded differently ("The baby was woken up"), or explicitly marked as accidental ("She woke the baby up by accident"). It is for future research to explore different possible sources for this intentionality bias, for example by using both linguisitic and visual stimuli.

[4] The analysis with only full sentences (instead of including full sentences and sentence fragments) does not change the level of significance, F (1,71) = 11.143, p = .001.

Would people be as likely to overattribute intention when evaluating *images* of actors engaged in ambiguous actions (e.g., stepping in a puddle, cutting someone off in traffic)?

These findings have important theoretical and practical implications. On the theoretical level, they indicate that research on intentional understanding would benefit from enlarging its subject pool beyond the early years of childhood; these studies focused on adults, but adolescents and school-aged children may also be informative populations of study. As the ability to override the intentionality bias develops gradually as a function of one's experience with non-intentional explanations for behavior, the prediction is that adolescents would make fewer intentional overattributions than children, but that they would still make more than adults.

Although children past a certain age will easily state that tripping over a ball is different from kicking it, relying on measures beyond verbal report may yield more nuanced results. Although such measures are commonly used with preverbal infants, and with adults in other subdisciplines of psychology, they have not been fully exploited with older children within the field of cognitive development. And yet, if we are interested in understanding the developmental trajectory of intentional explanation from infancy to adulthood, older children and adolescents are fertile ground for future study.

This of course raises another important theoretical issue, namely that of multiple levels of intentional understanding. As with many forms of reasoning, intentional inference may be best understood using a dual process model, where intention is automatically activated on an implicit level via the intentionality bias, and later overridden with more deliberate reasoning. Indeed, it appears as if it is the process of overriding, and inhibiting, the initial intentional inference that characterizes a mature understanding of behavior. As demonstrated with the present studies, measures that tap into this more immediate level of activation may provide important information regarding not only the developmental trajectory of intentional explanation but also in terms of individual differences in inhibiting the initial intentional judgment.

The ways in which individuals, and to another degree cultures, differ in their readiness to attribute intention to various acts leads to a host of more practical implications raised by the present proposal. For instance, studying the behavior of individuals with autism in terms of the intentionality bias may prove informative in understanding both normal and atypical behavior. As people with autism are believed to have deficits in reading intention, one possibility is that intention is not activated automatically upon evaluating an agent and an action as it is with typically developing individuals. If this is the case, one would predict that these individuals would be less susceptible to the manipulations used in the present studies; time constraints and priming would not influence their explanations one way or the other. Individuals with hostile attributional bias, on the other hand, may have an "overactive" intentionality bias, in that they infer intentions in ambiguous situations. They may need more cues than the average person to inhibit the automatic intentional interpretation that, say, a brush on the shoulder is due to malice instead of clumsiness. This has particularly important implications given that a recent meta-analysis investigating the causes of aggression found hostile attributional bias to be an integral component of aggressive behavior (Orobio de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002).

And of course, distinguishing levels of intent forms the crux of our moral judgments and our legal system. Perhaps the most extreme illustration is that the difference between murder and manslaughter rests on intentional inference; yet all judgments of guilt or innocence rely on how much the agent intended the action in which he was involved. Interestingly, there is a class of injury referred to as "intentional injury;" yet they are designated as such not by knowledge of the perpetrator's intentions, but by the fact that they were inflicted, in other words, produced by an action. A recent report by Mawson (2005, p. 377) states that "Injurious intent is taken for granted, with little or no attempt at verification... Intentions are rarely known and seldom investigated." This is consistent with the intentionality bias where, all else being equal, actions will be judged to be intentional until proven otherwise. This bias thus highlights the importance of the reminder, so crucial to a fair trial, that people are innocent until proven guilty. Indeed, the fact that we need such a reminder is telling, as it seems we often neglect to see that many actions are caused *un*intentionally. Rarely, however, do we have to explicitly persuade ourselves that people act intentionally – the intentionality bias does it for us.

Acknowledgments

This work was supported in part by a Clara Mayo dissertation award. Thanks also go to Deb Kelemen, Krista Casler, Cara DiYanni, Becca Seston, and three anonymous reviewers for comments on a previous version of the manuscript.

References

Baldwin, D., & Baird, J. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Sciences, 5*, 171–178.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge: MIT Press.

Baron, J., & Hershey, J. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54*, 569–579.

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences, 4*, 29–34.

Boyer, P. (2003). Religious thought and behavior as by-products of brain function. *Trends in Cognitive Sciences, 7*, 119–124.

Burger, J. (1981). Motivational biases in the attribution of responsibility for an accident: A meta-analysis of the defensive-attribution hypothesis. *Psychological Bulletin, 90*, 496–512.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257–271.

Guthrie, S. (2001). Why gods? A cognitive theory. In J. Andresen (Ed.), *Religion in mind: Cognitive perspectives on religious belief, ritual, and experience* (pp. 94–111). New York, NY: Cambridge University Press.

Kelemen, D. (2004). Are children 'intuitive theists'? Reasoning about purpose and design in nature. *Psychological Science, 15*, 295–301.

Knobe, J. (2003). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology, 16*, 309–324.

Kordes-de Vaal, J. (1996). Intention and the omission bias: Omissions perceived as nondecisions. *Acta Psychologica, 93*, 161–172.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*, 311–328.

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum.

Lockhart, R. S., & Craik, F. I. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology, 44*, 87–112.

Malle, B. F., & Knobe, J. (1997a). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33*, 101–121.

Mawson, A. (2005). Intentional injury and the behavioral syndrome. *Aggression and Violent Behavior, 10*, 375–405.

Miller, P. H., & Aloise, P. A. (1989). Young children's understanding of the psychological causes of behavior: A review. *Child Development, 60*, 257–285.

Orobio de Castro, B., Veerman, J., Koops, W., Bosch, J., & Monshouwer, H. (2002). Hostile attributional intent and aggressive behavior: A meta-analysis. *Child Development, 74*, 329–345.

Preissler, M., & Bloom, P. (2008). Two-year-olds use artist intention to understand drawings. *Cognition, 106*, 512–518.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition, 36*, 1–16.

Rosset, E. (2007). Intentional until proven otherwise: Evidence of an explanatory bias in children and adults. *Dissertation Abstracts International, 68*, 2689.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review, 76*, 179–193.

Shultz, T. R., Wells, D., & Sarda, M. (1980). Development of the ability to distinguish intended actions from mistakes, reflexes, and passive movements. *British Journal of Social and Clinical Psychology, 19*, 301–310.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.