The Big ‘Whoops!’ in the Study of Intentional Behavior: An Appeal for a New Framework in Understanding Human Actions

Evelyn Rosset*
Department of Psychology, Boston University, Cummington Hall, Boston, MA 02215, USA
*Corresponding author, e-mail: erosset@gmail.com

Joshua Rottman
Department of Psychology, Boston University, Cummington Hall, Boston, MA 02215, USA

Abstract

Distinguishing intentional behavior from accidental behavior is a crucial component of social cognition and a major developmental achievement. It has often been assumed that developmental changes in intentional reasoning result from a gradual sophistication in the ability to discern intentions in action. We take issue with this notion, demonstrating that data from cognitive, developmental, and social psychology are more consistent with the hypothesis that it is instead a gradual sophistication in the ability to understand accidents that drives developmental change.

Keywords

Intentions – accidents – cognitive development – interpretive mind understanding

Introduction

Understanding intentions has been heralded as the foundational, and perhaps most essential, ability in understanding human behavior (e.g., Premack, 1990; Feinfield et al., 1998; Baldwin and Baird, 2001; Tomasello et al., 2005). Once one comes to appreciate the mental causes of intended actions, the task of predicting and interpreting human behavior becomes infinitely easier. This
ability to understand intentions is undeniably crucial, as it allows for the knowledge that people often move toward things because they plan to, that outcomes often occur because other agents tried to make them take place, and that coherent patterns can be found in the world by assuming mental causes of human-initiated events. However, as important as intention reading is, it is equally important that people understand that some human behaviors may result from other (i.e., non-intentional) causes (e.g., Provine, 2005). After all, people yawn, trip over their own feet, fortuitously end up in the right place at the right time, and have unfortunate slips of the tongue. Interpreting such actions as intentionally caused would be both troublesome and confusing. A mature and healthy understanding of behavior therefore involves not only identifying actions as intentional, but also recognizing when they are not. Work in developmental psychology has largely focused on how children identify actions as intentional; the purpose of the present paper is to underline the importance of understanding how (and when) we recognize when they are not. Specifically, this paper will present two possible developmental frameworks, and then proceed to give evidence from infancy, childhood, and adulthood suggesting that a focus on non-intentional interpretations is the key to understanding the development of the ability to explain and predict behavior.

The Development of Intentional Understanding

Over the past three decades, many developmental psychologists have explored how infants and children ‘develop intentional understanding’ (e.g., Smith, 1978; Miller and Aloise, 1989; Meltzoff, 1995; Baird and Astington, 2005; Tomasello et al., 2005; Olineck and Poulin-Dubois, 2009; Mull and Evans, 2010). This has been extremely important in discovering how children make sense of the world in general and of other people’s behavior in particular. The methods used to address this issue typically involve presenting children with different kinds of actions – some intentional and others non-intentional1 – and examining when and under what circumstances the two are distinguished. It is, therefore, not simply the ability to discern intentions that is being tapped in these studies but, more precisely, the ability to distinguish intentional actions from non-intentional ones. These numerous studies could be framed as investigating

---

1 For ease of reading, throughout this article we will at times replace “non-intentional action” with ‘accident.’ We recognize, however, that accidents are not the only kind of non-intentional action, as this latter category also includes biologically-caused involuntary actions like sneezing, physically caused actions like being pushed by the wind, and chance.
when children understand intentions (and they generally are framed as such), but they could also be seen as investigations of when children understand accidents.

This is not purely a semantic distinction. Framing the question in terms of understanding accidents makes different predictions in terms of the developmental trajectory of behavioral understanding, and has fundamentally different implications for ascertaining the precise nature of the mature competence for interpreting other people's actions.

The competence with which people understand non-intentional actions has not been a major concern of the philosophers and developmental psychologists who have studied folk psychology, for whom it seems that an understanding of intentions is the *sine qua non* of understanding agentive behaviors. This theoretical position is possibly a result of the intuitive assumption that intentional acts are cognitively harder to understand than non-intentional acts, and that once intentional knowledge has developed, children should be able to reason fluently about the causal underpinnings of the full gamut of human behavior.

Counterintuitively, however, it may be that judging an act to be intentional is cognitively easier than judging it to be non-intentional. Premack (1990), for instance, has argued that whenever an agent is the focus of attention, and this agent is involved in an action that is not directly induced by another agent or object, the impetus of the action is judged to be the intention of the agent by default. Similarly, Rosset (2008) proposed that adults and children are susceptible to an ‘intentionality bias’: when evaluating agents and actions, intention is assumed to be the causal mechanism by default, and it is only through experience with non-intentional actions that we learn to override this default interpretation. It is entirely feasible, therefore, that judging a given behavior as accidental requires more cognitive resources than judging it to be intentional. To stress this point, once a basic understanding of intentional states comes online in infancy, it may actually be *more difficult to understand accidents than intentions*.

**One Corpus, Two Possible Frameworks**

The framework that currently informs the basis of much of the contemporary work in developmental psychology suggests that children develop a nuanced understanding of behavior by honing their ability to understand intentions as the cause of human action (e.g., Miller and Aloise, 1989; Baird and Astington, 2005; Tomasello et al., 2005; Mull and Evans, 2010). According to this view,
children do not initially grasp that actions may be caused by mental states such as intentions; only as this ability becomes more sophisticated do they arrive at a mature understanding of human action. As the emphasis of this framework is placed on discovering when and how a grasp of intention as a causal force develops, we will refer to this model as the ICED framework (Intention as Causal Explanation Develops).

There is, however, an alternative framework that has not yet been given an adequate voice. According to this framework, children quickly develop a default explanatory bias that all human action is intentionally driven. Through repeated experience with cues that many actions may not be intentional (e.g., noticing that outcomes do not always match desires), children develop a mature understanding of behavior. In this model, then, it is the understanding of non-intentional causes of behavior that develops; we refer to this as the ‘NICED’ (Non-Intention as Causal Explanation Develops) model.

In the ICED model, children come to rely increasingly on intentional explanations (with an accompanying decrease in non-intentional explanations). In the NICED model, they rely increasingly on non-intentional explanations (with an accompanying decrease in intentional explanations). The contrasting developmental predictions of the ICED and NICED frameworks are described and depicted graphically in Figure 1.

Review of the Developmental Literature in Light of the Two Models

Looking at the empirical data in light of these two models, there is more consistent support for the predictions generated by the NICED model. This is especially apparent when taking a developmental perspective, and therefore we will sequentially examine data from infancy, childhood, and adulthood. First, research presented in the next section will demonstrate the existence of an early-developing and robust sensitivity to goals and intentions in infants. This work establishes that a mastery of intentional reasoning is not a feat that takes a great deal of time to achieve, but instead, something that matures rather quickly in the first two years of life. Second, another broad line of research has demonstrated that toddlers have difficulty distinguishing accidental from intentional acts, a finding consistent with both models. Contrary to what might be predicted from the ICED model, however, these children tend to err on the side of attributing intentions when none exist. Third, work with adults demonstrates a tendency – not unlike children in comparable situations – to over-attribute intention, and this is especially pronounced in situations of cognitive load. It is noteworthy that it is the later stages of development that reveal
the more consequential differences of the two models; indeed, this may be why the ICED model has remained the dominant model, at least implicitly, in developmental psychology.

**Infancy**

Actions are readily understood as intentional in early cognitive development. Infants as young as three months of age can attribute goals to human and non-human agents (Sommerville et al., 2005; Luo, 2011). By the second year of life, an even more sophisticated understanding of intentions is evident in a variety of domains. Children as young as 18 months perceive and infer the intentions of adults when learning new words (Tomasello et al., 1996; see also Tomasello, 1999; Bloom, 2000 for reviews). Additionally, toddlers can use their knowledge...
of intentions to guide their behavior: 14- and 18-month-olds copy intentional acts much more readily than mistakes (Carpenter et al., 1998). 18-month-olds do not directly imitate failed actions but rather re-enact behaviors consistent with inferred goals (Meltzoff, 1995), and two-year-olds use intentional cues when figuring out a problem-solving task (Carpenter et al., 2002) and when making judgments about how a novel object should be used (DiYanni and Kelemen, 2005). Therefore, while the ability to explain a broad range of intentional acts may become slightly more nuanced through early and middle childhood (Mull and Evans, 2010), research suggests that by two years of age, children can perceive and infer the intentions of agents to learn new words, selectively imitate, problem-solve, understand an object’s function, and make inferences about causality.

An ongoing debate is whether this early intentional understanding is the result of an innate modular system (Leslie, 1994; Premack, 1990) or whether it is constructed (Gopnik and Wellman, 1994; Piaget, 1952) from first-person experience (Meltzoff, 1995; Tomasello, 1999) or from statistical regularities in the environment (Baldwin and Baird, 2001). The jury is still out on which of these theories best characterizes the data – indeed, it is very likely that both of them have something important to contribute. It is noteworthy, however, that initial over-attributions of intentionality could result from either process.

Regardless of how intentional understanding in infancy arises, the acquisition of this initial competence sets the stage for two possible developmental trends from childhood through adulthood. According to the ICED framework, the initial understanding of intention should gradually become less fragile and more nuanced, and should be increasingly used in appropriate situations. The prediction is that young children will have more difficulty with intentional explanations than adults. Conversely, the NICED framework asserts that intentions will be over-attributed as soon as this basic understanding arises in infancy. Here the prediction is that young children should habitually use this explanatory framework to attribute intentional causes to all actions, and this should gradually taper off until these explanations are inhibited when appropriate. The remainder of this paper will demonstrate how the evidence points to the existence of the latter developmental trend.

**Childhood**

Infancy research relies on behavioral measures that tap an implicit, non-verbal level of understanding. Using such measures, the studies described in the previous section suggest an early-developing and robust sensitivity to distinguishing goal-directed and intentional actions. As the subject population becomes older, however, and the behavioral measures rely on the more explicit level of processing involved in verbal report, the ability to distinguish intentional acts
from accidental ones appears more fragile. This is not likely a reflection of decreasing competence in children’s explanations; rather it is an issue of performance. Toddlers are no less able to detect intentions than are infants; they just have a harder task at hand when asked to respond on a more explicit level. According to the NICED framework, the difficulty shown by toddlers and older children in distinguishing intentional acts from other kinds is due to children’s tendency to judge all acts as intentional (including those brought on by physical and biological causes, accidents and mistakes).

In one of the first studies looking at intentional explanation in children (Smith, 1978), four- to six-year-old children watched videos of a woman performing voluntary movements (e.g., chewing something), involuntary movements (e.g., sneezing), and ‘object-like movements’ (e.g., her arm being hooked by an umbrella), and were asked a series of questions such as whether the woman ‘wanted’ and ‘was trying to do’ the movements. The 4-year-olds judged all the acts as intentional; their evaluations of the intentions behind acts as different as sneezing and chewing were identical. Although 5-year-olds were better able to distinguish the three types of acts, only by age six were children performing like adults and giving higher ratings of how much the woman desired and intended to perform the voluntary acts. This pattern of results demonstrates a progressive ability to distinguish actions in terms of their intentionality, and thus could be due to what the ICED framework construes as a developing capacity to recognize intentional action. However, because intentionality ratings decreased as the children got older, it may be more appropriate to interpret such results as showing a developing capacity to recognize non-intentional action. In other words, 6-year-olds are not better at understanding that chewing is intentional; rather, they are better at recognizing that sneezing is not.

Along similar lines, studies such as those by Shultz et al. (1980), Lang and Perner (2002), and Montgomery and Lightner (2004) have shown that children will over-attribute intention even to acts that they themselves intentionally perform. In these studies, three- and four-year-olds claimed that they had intentionally caused various unintentional actions, such as knee-jerk reflexes and drawing pictures during which their hands were guided by an experimenter. Furthermore, other research has found that, while children are able to acknowledge that behaviors may indeed result from physical or biological causes when their responses are sufficiently scaffolded, their initial judgments tend towards explaining all behaviors in intentional terms. This is particularly apparent at younger ages (Shult and Wellman, 1997).

In sum, cognitive developmental studies have demonstrated that, even at a young age, people are exquisite intentionality detectors – so good, in fact, that they tend to over-attribute intentions. The assumptions of the ICED framework – e.g., that intentions are only assigned when an action can reasonably be said to

THE BIG ‘WHOOPS!’ IN THE STUDY OF INTENTIONAL BEHAVIOR

have been directly caused by a mental state – are somewhat puzzling given the findings that intentions are over-attributed and invoked more readily than accidents in children's causal explanations of human behavior. Indeed, although the finding that children over-attribute intentions is well established, this pattern of results is entirely unpredicted by the standard ICED framework. However, these data are easily explained by the NICED model, which not only accounts for over-attributions of intention in childhood, but actually predicts them.

Adulthood

It is noteworthy that the bulk of the empirical work on intentional explanation has been conducted with children. There is comparatively little research focused on understanding intentional explanation in adults, an omission that is not surprising in light of the assumptions of the ICED framework. There are, however, a handful of exceptions. Malle and Knobe (1997), for example, examined adults' folk concepts of intentions and intentionality and found strong agreement among adults, suggestive of a common underlying ‘intentionality concept' made up of five elements (i.e., the presence of desire, belief, intention, awareness, and skill). Such findings do indeed demonstrate a certain level of mastery among adults as to what makes an act intentional. Importantly, however, this study assessed an entirely explicit level of understanding.

The present proposal is not arguing that adults believe on an explicit and conscious level that all acts are intentional. After all, adults, unlike children, do not claim that people intend to sneeze, or trip, or jerk their knee when it is hit by a doctor's mallet. According to the NICED framework, this competence does not stem from a more fully developed ability to attend to the relevant intentional cues of an action. Rather, adults possess a more robust knowledge base regarding the non-intentional (e.g., biological) causes behind sneezes, stumbles, and reflexes. However, the present proposal does make the unique prediction that, if this knowledge was compromised in some way, then adults would also be more likely to over-attribute intentional causes to behaviors brought on by other means. This prediction has been recently confirmed. When adults need to make snap judgments (Rosset, 2008), or are asked questions about behaviors when their executive functioning is diminished by alcohol consumption (Bègue et al., 2010), they become much more prone to judge actions as intentional – presumably due to their weakened inhibitory control.

Moreover, research in social psychology is filled with examples of adults over-attributing intentional causes to behavior. This is present in many everyday epistemological ‘mistakes' such as the fundamental attribution error (Jones and Harris, 1967), the illusion of control (Rotter, 1954), the just-world phenomenon (Lerner and Miller, 1977), the outcome bias (Baron and Hershey, 1988), and the omission bias (Kordes-de Vaal, 1996). This tendency to
over-attribute intention is likewise consistent with recent work in experimental philosophy that has found that people will attribute morally questionable intentions even when they are explicitly told that a bad outcome was only a byproduct of an actor’s true intentions (Knobe, 2003).

A Dual Process Model of Intentional Inference

Fundamental to the present proposal of an intentionality bias is the idea that there are different levels of processing of intentional inference. As with many forms of reasoning, our inferences about intentions may best be understood with a dual process model, where intention is automatically activated via an intentionality bias when evaluating an agent and an action (Premack, 1990; Rosset, 2008), and may be overridden by more deliberate processing. If this is true, then every act, even for adults, is initially interpreted as intentional at some level. Adults (and children for that matter) may say that they recognize the accidental nature of falling down the stairs, or the biological cause of sneezing; their behavior, however, may show otherwise, revealing traces of an intentional interpretation. Indeed, there is evidence that all acts are initially construed as intentional; this is true even for actions that are typically accidental, like breaking a vase or setting a house on fire (Rosset, 2008). While different levels of intentional inference are integral to the present proposal, traditional frameworks do not predict a dissociation between explicit beliefs and implicit behavioral measures. Indeed, some of the inconsistencies that have been found in current approaches may be due to a failure to note the importance of this distinction.

The developmental trend that is observed in intentional understanding is therefore not due to a gradually developing ability to attribute intentions. Rather, it is likely due to the strengthening of executive processing (and especially inhibitory control), which allows older children and adults to override their initial impulses to interpret actions intentionally. Additionally, both the familiarity (and thus accessibility) of alternative explanations and an understanding of the agent’s behavioral cues increase with age and experience, in part due to Bayesian learning. The source of the developmental shift in behavioral interpretations is therefore rooted in changes in domain-general processes rather than the progressive specialization of a domain-specific competence.

Conclusions and Implications for Future Research

This paper serves as an appeal for a new framework by gathering evidence that people become increasingly skilled at understanding various causes of behavior
not by learning how to apply intentional explanations, but rather by becoming increasingly adept at inhibiting them in favor of non-intentional explanations.

These findings have important implications for future research. First, they challenge the implicit assumptions of current (ICED) approaches to intentional understanding in developmental psychology. More precisely, the results reviewed above suggest that a mature understanding of behavior does not depend on one's ability to detect intentional cues, but rather on the ability to override a default intentional interpretation when necessary. As this approach puts adults and children on a continuum, it suggests that studying adults – not to mention adolescents and school-aged children – may yield important information regarding the developmental trajectory of intentional understanding. Contrary to predictions of the ICED framework, the NICED proposal argues that adults do not master intentional explanation; they simply have more experience, more knowledge, and greater inhibitory control that they can employ to override the intentionality bias. Adults may inhibit their initial intentional interpretations faster than children, but they cannot erase them completely; intentional explanations will still be activated initially. Adolescents in particular present an age ripe for research on this topic. On the one hand, they know that sneezes and stumbles are involuntary, that things happen by chance, and that people make mistakes. On the other hand, this explicit knowledge is fragile and thus may result in an increased sensitivity to potential slights.

This brings us to the second implication of the present proposal in terms of future research: the importance of a dual-process model of intentional inference. Implicit measures of intentional understanding have been used in infancy research for some time, yet they become conspicuously absent once these infants learn to talk. This is puzzling, given the wealth of information that implicit measures have generated in other domains of research with adults. This proposal therefore serves to highlight the importance of such a dual-process model, and the potential of a host of implicit behavioral measures that can be used in populations of all ages.

Adopting the framework proposed in this paper holds promise in furthering scientific knowledge of folk psychology. Assuming that an understanding of intentions is a universal human ability, it will be difficult to find individual differences in this ability after the second year of life. However, there is a much higher potential to find variability in peoples’ understandings of accidental behavior. Because data on individual differences is so crucial in constructing theories and gaining more nuanced understandings of behavioral phenomena (Underwood, 1975), the NICED framework will likely prove quite useful in this regard.

In sum, the change of framework proposed in this paper has implications for future research, not only within developmental psychology, but also in fields of inquiry as diverse as social psychology, judgment and decision-making,
and clinical psychology. Furthermore, an explicitly formulated framework outlining how and why our explanations of behavior change across the lifespan has considerable practical implications. Recognizing that adults have an automatic tendency to over-interpret intentions may yield insight into the most mundane aspects of life (e.g., seemingly innocent lane changes that trigger road rage), as well as to the most profound (e.g., laws distinguishing between murder and manslaughter, religious beliefs in supernatural agents). In light of this, acknowledging that accidents are not as readily understood as intentions is not a trivial matter.

Acknowledgements

The authors would like to thank Deb Kelemen for her helpful feedback on the manuscript. This writing was supported in part through a National Science Foundation graduate research fellowship awarded to the second author.

References


