

INTENTION, INTERPRETATION, AND MORAL RESPONSIBILITY: CHILDREN'S  
CHANGING BELIEFS ABOUT TRUTH AND RIGHTNESS

by

BRYAN W. SOKOL

B.A., St. Louis University, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Psychology

We accept this thesis as conforming  
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1998

© Bryan W. Sokol, 1998

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology

The University of British Columbia  
Vancouver, Canada

Date 13 Nov. 1998

## Abstract

Available research meant to bring out what children of various ages judge to be right and what they take to be true has evolved along separate pathways. That is, with the exception of a now defunct literature prompted by Piaget's (1932/1965) early concerns with the study of the role of intentions and consequences in morally relevant actions, recent research into children's beliefs about belief, and counterpart studies of their changing conceptions of moral responsibility have been largely non-overlapping enterprises. The research reported here represents an attempt to bridge this gap by examining the ways in which more or less mature conceptions of the knowing process influence the particular manner in which both adults and children judge moral culpability. Study 1 focused on the moral deliberations of a sample of 56 university students and was meant to provide a benchmark of the ways in which adults weigh the misdeeds of a story character who acts out of either simple ignorance or misguided interpretation. Study 2 similarly examined the moral judgments of 54 5- to 7-year-olds who differed in their grasp of the possibility of false belief and of the interpretive character of the knowing process. Results showed that children who subscribe to an interpretive or constructivistic "theory of mind" responded as did the adults of Study 1, by excusing persons for misdeeds that followed from mistakes of simple ignorance, while assigning greater responsibility to those whose wrongdoings arose out of having misinterpreted available evidence. This finding was in contrast to the responses of typically younger children characterized by a non-constructivistic understanding, or simple "copy theory", of mental life who failed to judge mistakes of ignorance and misinterpretation differently. It is argued that children's conceptions of morality, and particularly their judgments of moral responsibility, change as a consequence of developments in their beliefs about simple matters of fact and more complex matters of interpretation.

## Table of Contents

Abstract .....	ii
List of Figures .....	v
Acknowledgements .....	vi
Introduction .....	1
Characterizing the Moral Point of View: The Importance of Mental States .....	2
Understanding Desires, Beliefs, and Intentions: Alternative Manifestations of Human Agency .....	6
Building an Adult Conception of Human Agency: From "Copy" Theories to an "Interpretive" Understanding of Mental Life .....	10
Turning to the Law for Clarification: Mistakes of Fact and Mistakes of Law .....	13
Distinguishing Simple Ignorance from Genuine Interpretation: What Do Children Think? .....	15
Choosing a Procedural Approach: A Revival of "That Highly Moral Drama"—Punch and Judy .....	18
Study 1: Characterizing the Endpoint—Adults' Responses to Punch and Judy .....	21
Method .....	21
Participants .....	21
Stimulus Materials .....	22
Procedure .....	23
Story Questions .....	24
Scoring a Constructivistic Understanding .....	26
Results and Discussion .....	27
False Belief Conditions .....	29
Interpretive Conditions .....	30
More or Less "Consolidated" Varieties of Constructivism .....	31
Study 2: The Child's Point of View .....	33
Method .....	35
Participants .....	35
Procedure .....	35
Memory Probes .....	36
Control Questions for the Interpretive Scenes .....	36
Comprehension Questions—"Badness" Ratings .....	37
Criteria for Inclusion in the Study .....	37
Criteria for Categorizing Children as Subscribing to a "Copy" or Constructivistic Theory of Mind .....	38
Response Measure .....	40

Results and Discussion .....	40
Between-Groups Comparisons—Differences in the Use of Intention Information .....	41
Disentangling Age from Epistemic Stance .....	44
Within-Groups Comparisons—Differentiating Mistakes of Fact and Mistakes of Law .....	45
General Discussion .....	47
Bibliography .....	49
Appendix 1	
Adult Interview Schedule .....	56
Appendix 2	
Child Interview Schedule .....	59

## List of Figures

Figure 1: Adults' "Badness" Ratings.....	29
Figure 2: Children's "Badness" Ratings.....	42
Figure 3: Manslaughter Ratings for Children Coded as either "Transitional" or "Fully" Constructivistic.....	46

## Acknowledgements

Financial support for this research was provided in part by an operating grant from the Natural Sciences and Engineering Council of Canada to my supervisor, Michael Chandler. I would like to acknowledge Craig Hennessey, Willy Percy, Larry Hennessey, and Fran Dowie for their invaluable assistance in creating the Punch and Judy stimulus materials—the high quality video production would not have been possible without their contributions of technical expertise and creative energy. In addition, I would like to thank Michael Chandler for his expert mentorship and guidance throughout the process leading to the production of this thesis document; Jeremy Carpendale for his involvement in the early phases of conceptualizing and planning this project; Geoff Hall and Larry Walker for their willingness to serve as committee members and offer constructive criticism of my work; Preena Sandhu for her help in the data collection process; my family and friends for their constant encouragement; and Melissa Apprill for both her tireless support throughout every phase of this endeavor and her invaluable insights regarding the parallels between this work and the domain of jurisprudence.

Finally, my deepest thanks must be extended to the children and staff at Spare Time Fun Centre, the University of British Columbia's Child Care Services program, and S7UMUX Kindergarten for their enthusiastic participation in this research.

## Introduction

Seemingly reduced to silence by a fear of somehow violating Hume's classic prescription (1739/1978) against mixing claims about what "is" the case with value statements about what "ought" to be, developmental psychology has paid relatively little to attention to the often complex ways in which so-called "facts" and "values" ordinarily commingle in the moral deliberations of children and adults. In particular, the contemporary theories-of-mind enterprise, which is all about the manner in which children come to understand the process of belief entitlement, and the substantial literature on the development of moral reasoning are almost completely non-overlapping (for a rare exception see Wainryb & Ford, in press). The impetus for the present research undertaking grew, in part, out of the vacuum created by this isolationist policy, and so can be seen as an attempt to bridge the artificial divide that now separates the research literatures on children's changing views of mental life from their developing moral reasoning competence. More specifically, the broader research program, of which this particular study is a specific part, represents an attempt to explore how turning points in children's conception of their own and others' mental lives, and their changing views about the active, interpretive character of the knowing process, might influence their deliberation about various morally relevant actions.

The brace of studies to be reported builds on the work of philosophers such as Bratman (1984), Harman (1976), Searle (1983), and Wren (1974), who argue that a mature understanding of mental life necessarily involves more than just a simple appreciation of the existence of basic desires and beliefs. Rather, according to these theorists, the essential leavening ingredient that is otherwise missing from any such limited belief-desire psychology is a notion of agency that, at



its core, includes an appreciation of minds as being fundamentally active and interpretive, as opposed to merely passive and reactive. Notwithstanding the fact that even infants appear to have at least a tacit appreciation of human agency in highly structured action-contexts (see, for instance, Baldwin, Markman, Bill, Desjardins, & Irwin, 1996; Frye, 1991; Meltzoff, 1995; Premack, 1990; Spelke, Phillips, Woodward, 1995), it is the guiding assumption of this present research that young children do not come to a working understanding that the process of knowledge acquisition is itself an active, interpretive achievement until they are roughly 7 or 8 (e.g., Carpendale & Chandler, 1996; Chandler & Lalonde, 1996; Chandler & Sokol, in press). That is, it is not until they have come to an initial "constructivistic theory of mind" that young school-aged children begin to view themselves and others as "epistemic agents," whose active influence over the knowing process shapes the course of their own moral deliberations.

#### Characterizing the Moral Point of View: The Importance of Mental States

Outside the world of academic psychology, the prospect that an understanding of others' mental lives is deeply implicated in the process of adjudicating moral culpability is a notion of long standing. A long list of both classic (e.g., Aristotle, 1985; Kant, 1785/1959) and more contemporary philosophers (e.g., Duff, 1990; Harré, 1982; Hart, 1958; Ross, 1973; Strawson, 1982; Wren, 1974) have all insisted that the proper determination of an individuals' actions as being either right or wrong necessarily involves an understanding and consideration of those mental states that presumably precede and guide such actions. Some have even argued that a psychological, or "inward," focus concerning such things as beliefs and intentions is the only proper aspect of concern in moral evaluation. This is evident, for instance, in Kant's (1785/1959) famous remark that "when the question is of moral worth, it is not with the actions which we see

that we are concerned, but with those inward principles of them which we do not see" (p. 23). Without the weighing of such inward, psychological factors, moral evaluation, it has long been argued, amounts to no more than simple talk about the negative consequences that follow from people's outward behavior. Although classical stimulus-response psychology has shown an affinity for such limited "behavioristic" accounts, there are good reasons to suppose that any such outcome-oriented view is especially problematic when brought to bear on moral issues as they are ordinarily construed. As Kohlberg (1969), in particular, has argued, behavioristic accounts of morality fail to define many, if not the large bulk of, morally relevant actions. To take only one celebrated example, Henry David Thoreau and his cell mates all "did time" for the same manifest failure to pay taxes, although their crimes were presumably attempts to only weasel out from beneath their tax burdens, while his was allegedly a purposive "matter of conscience," inspired by his pacifist leanings. On a strictly behavioral level, such counterpart violations of the law are indistinguishable, while, as Kohlberg suggests, it is more properly "what the people involved think they are doing which sets [their] behavior apart" (cited in Turiel, 1990, p. 38, emphasis added).

Although there are certainly many important elements going into the process of moral evaluation, the clear contribution of an understanding of others' mental states should have served as an obvious cue for those involved in the study of children's theories of mind to become actively interested in this area of study. Unfortunately, this has not been the case (Chandler & Carpendale, 1998; Chandler & Sokol, in press), and the theories-of-mind enterprise has, consequently, remained essentially cut off from the study of children's moral reasoning development. The responsibility for this evident tunnel vision cannot, however, be laid entirely

at the door of theorists of mind. As it is, those working in the current mainstream of moral development research are often equally guilty of having ignored the role of beliefs in the process of moral deliberation.

Interestingly, the current fragmented state of affairs concerning the study of truth and rightness was not always the norm among developmentalists. Rather, as is widely known, there was in the early 1970's a brief flurry of research aimed at testing Piaget's (1932/1965) claim that children's moral deliberations undergo a transformative process involving a shift of focus from an early preoccupation with overt "objective" factors, to a later occurring emphasis on less obvious "subjective" cues, all relevant to any comprehensive assignment of moral responsibility (for reviews see Karniol, 1978; Keasey, 1977; Lapsley, 1996, esp. Chap. 2). Unfortunately, this potentially rich area of study went fallow, in large part, because of the overly narrow fashion in which those involved operationalized Piaget's distinction between objective and subjective modes of moral evaluation, by taking them to be equivalent to the ability to take both "consequences" and "intentions" into account in forming moral judgments (Dean & Youniss, 1991; Keasey, 1977; Youniss & Damon, 1992). Despite these limitations, some interesting findings did emerge. One of these was that preschool children—children young as 3 years old (e.g., Nelson, 1980)—appear to use information concerning intentions in the making of moral judgments, particularly if such information is made especially salient (Chandler, Greenspan, & Barenboim, 1973). The limited conclusion drawn from these facts has been that, while young children tend to initially overlook more elusive sorts of information provided by others' intentional states, they do gradually come to be more responsive to such obscured "subjective" variables and give them increasing weight in their moral evaluations as they grow older (Buchanan & Thompson, 1973;

Costanzo, Coie, Grumet, & Farnill, 1973; Farnill, 1974; Grueneich, 1982; Gutkin, 1972; Hebble, 1971; Helwig, Hildebrandt, & Turiel, 1995; Leon, 1980; Piaget, 1932/1965; Surber, 1977).

Although a makeshift developmental story has been constructed from these early findings, it is, nevertheless, apparent that, at least for most investigators working in this area, the ability of children to compute both the intentions behind, and the consequences that follow from, any morally relevant action is not only assumed to be in place by a very early age (e.g., Shultz, Wright, & Schleifer, 1986), but, also, is typically understood to change only in the quantitative, or "algebraic," ways that such factors are combined (e.g., Anderson, 1980; Lane & Anderson, 1976; Leon, 1980; Surber, 1977; Zelazo, Helwig, & Lau, 1996). On this popular computational view, Piaget's original insight regarding the role of "subjectivity" in children's moral reasoning development has been effectively submerged by the current tendency to see the mind in strict "objectivist" terms. As it stands, the received view on children's moral reasoning competence is little more than an attempt to characterize the sort of "mental calculus" presumed to underlie the earlier reaches of the belief entitlement process (Chandler & Carpendale, 1998; Chandler & Sokol, in press). This approach, in turn, has resulted in an essentially anti-developmental treatment of how an understanding of intentions, and its place as a cornerstone in the process of moral deliberation, works as an accomplished fact presumed to exist in equal measures among persons of almost every age. For a more complete developmental story to be told about children's moral reasoning competence, however, an attempt must be made, as I will argue, to produce a clearer picture of just how it is that children's notions of intention are necessarily situated in the broader context of their conceptions of mental life more generally.

### Understanding Desires, Beliefs, and Intentions: Alternative Manifestations of Human Agency

In opposition to the claims of more behavior-oriented investigators of psychology's mainstream, ordinary adults, as well as psychologists (e.g., Shultz, 1991; White, 1995) and philosophers (e.g., Harré, 1982; Wren, 1974) of a certain stripe, have all been quick to appreciate that any talk of intentions, and more particularly, moral responsibility, must necessarily begin with a consideration of the more fundamental matter of human agency, and how, in particular, it figures into children's emerging conceptualization of basic desires and beliefs. That is, at least within our Euro-American version of common sense, morality and intentionality are ordinarily understood to be the province of "agents" as opposed to "patients." On this common view, persons are understood to bear moral responsibility only for those of their intention-guided behaviors that they have had some active hand in bringing about. In other words, people whose trigger finger is controlled by wires, or persons who, through mental incompetence, are not free to actively form intentions, are ordinarily excused from responsibility for their actions (see Fincham & Roberts, 1985; or, for a more philosophical discussion, Strawson, 1962/1982). On this commonsense view, behavioral outcomes that merely befall, or are in some way visited upon us, and so lack any real intentional component, are simply coded as events without personal moral relevance. Within any such standard issue "folk psychology"—an account which also standardly views all intentional actions as the cross product of one's desires and beliefs—there are two obvious points of entry where human agency might be seen to gain a potential toehold.

The first and most accessible of these entry points is in the realm of "desire," where the wanting of one thing as opposed to another seems self-evidently the doings of an active chooser. As it happens, there is an ample supply of available research evidence addressing this issue.

Specifically, even if one were reluctant to accept research findings suggesting that even infants in arms are capable of recognizing particular kinds of goal-directed behavior (Frye, 1991; Meltzoff, 1995), then certainly available evidence from observations of slightly older children's interactions with others (Wolf, 1982; Wolf, Rygh, & Altshuler, 1984), as well as from their use and understanding of speech (Astington, 1986; Dunn, 1991; Montgomery, 1998; Wellman & Hickling, 1994), would be sufficient to convince most that preschool-aged children already have a fairly solid grasp of personal agency, at least in the limited sense of recognizing the operative role of others' explicit wants and desires. All such accomplishments notwithstanding, most advocates of this common sense view would go on to argue that while understanding what others want and desire is a worthwhile beginning, any mature view of how others' actions come about must also include some place for the role of belief (Harman, 1976; Moses, 1993). This has led some researchers (e.g., Astington, 1991; Meltzoff, 1995; Wellman, 1990) to begin trading on the work of philosophers such as Bratman (1984) and Searle (1983), who distinguish between an early understanding of presented-directed, intentions-in-action and more complex future-directed, prior intentions. On this view, these two different faces of human agency can be characterized in terms of either an individual's overt goal-directed acts that take place in the immediate present, in the first case, or as representing a person's abstract, future-oriented plans, in the second.

As Astington's (1991) research, in particular, suggests, young preschool children seem to have what is, at best, only a highly limited understanding of how human agency manifests itself. In particular, such young persons appear to easily comprehend the goal-directed character of particular concrete intentions-in-action, while still seeming to have only the vaguest inkling about the abstract nature of prior intentions. In fact, as Astington reports, it is not until the age of 4,

when children typically demonstrate an explicit understanding of false beliefs, that they also show signs of grasping personal agency in a more abstract, mentalistic fashion. That is, before coming to grasp the apparent role that future-directed intentions have in guiding others' actions, young children's understanding of human agency appears to be limited to those present-directed, intentions-in-action that can be readily matched with another's clearly advertised desires. By this account, then, a mature view of agency—at least as such an understanding bears upon the domain of “desire”—is seen to start with children's primitive intentions-in-action and reach its high point when a solid metarepresentational conception of belief (see Perner, 1988, p. 142) has been added to their repertoire of social-cognitive skills. In other words, a mature conception of agency, by this line of reasoning, is simply the result of combining a more abstract understandings of beliefs to an already worked-out conception of basic desires.

The problem with this account, as I mean to bring out, has to do with its failure to recognize that an understanding of belief is yet another potential place where an early understanding of human agentiveness might initially manifest itself. That is, depending upon the sort of tacit epistemology to which one subscribes, the beliefs that inform and instrument human action can be understood in either fundamentally passive or active terms. According to the first of these prospects, beliefs are simply those epistemic scars that people passively suffer as the reactive consequence of informational assaults from the “outside” world, and so, are not in any way an active creation of their own doing (Chandler, 1988). By contrast, any more constructive, or interpretive, view of the knowing process regards beliefs as those active achievements that automatically result from a process of co-construction brought about by agents who both select and interpret available inputs. On this view, the agentic powers of individuals can be seen to be

working both at the basic level of desire (an achievement which is well within the competence range of preschoolers) and at the more cognitive level of belief formation—an accomplishment that proves to be a later arriving understanding which appears only when children come to more fully appreciate the interpretive character of the knowing process.

Taken all together, then, the proposed picture regarding the various expressions of human agency that emerges is one that arrives in essentially three parts. First, there are those presumably very young children who have managed to link up whatever basic level conception of “agents” they may have harbored since infancy with a working understanding of desire. This leads such children to an early and limited grasp of intentions, at least insofar as they are expressed in certain limited, action-oriented contexts. Second, there are those children whose emerging conceptions of mental life, at least as they are demonstrated through common measures of false belief understanding, allows them to join up their limited notions of desires with a newly arriving view of beliefs. Consequently, such children are, consequently, able to talk more abstractly about their own and others’ so-called prior intentions, and to express how people’s actions follow from the particular beliefs that they are understood to hold at any given moment. Still, even such young “belief-desire” psychologists are limited—or, at least, so I mean to argue—to a passive conception of the belief formation process, and so, still have only a truncated view of human agency and intentions that remains locked to expressions of easily understood desire-outcomes. Finally, as the research here is meant to show, there are children with a more sophisticated, interpretive view of the knowing process who, in managing to get their heads into the third part of the developmental picture being previewed here, begin to look very much more like adults who recognize both themselves and others as being “epistemic agents” involved in



actively "authoring", and so, in some sense having "ownership" rights to (Blasi, 1991, 1995; Ross, 1973; R. Taylor, 1963, 1966), not only what they outwardly do, but also what they inwardly think and know.

In terms of how such changes prepare children to differently understand moral responsibility, the focus of the follow-up picture here should be clear enough. Children of the first two varieties (that is, children with an early understanding of desires and those who understand the possibility of false belief) are quite capable at judging the good or bad consequences of others' outward wants and desires. Still, these children's inherently passive conceptions of others' mental lives keeps their moral sensibilities on what they see as the only relevant plane, that is, on what they and others objectively accomplish in seeking to fulfill their desires. On the other hand, children of the third and more constructivistic sort, as I will argue, are well on their way towards seeing others as "moral agents," who by virtue of having minds with powerful interpretive abilities, are responsible not only for things that they objectively act out and physically accomplish, but also for what they subjectively come to know and think. Before further exploring the moral implications of such a more elaborated view of epistemic life, however, it is best to first try and get clearer about exactly what is meant by a having a constructivistic theory of mind.

#### Building an Adult Conception of Human Agency: From "Copy" Theories to an "Interpretive" Understanding of Mental Life

While research concerning children's changing views about the domain of desires is relatively thick on the ground, just the opposite is the case concerning research into the place of human agency in the development of young persons' conceptions of belief. That is, not

withstanding the fact that common sense ordinarily allows for the expectation that people have an active hand in the construction of their own beliefs (e.g., every thought one has ever had about the world of biases or prejudices is a thought about how people actively shape and put interpretive twists upon their experiences), precious little has been said within the contemporary psychological literature about how young persons come to share in such interpretive views about mental life.

Despite the shortage of evident signposts about how to best proceed, there is at least one standout way that suggests itself as a means of beginning to explore the role of human agency in the construction of knowledge. In particular, it would appear that young children are committed, at least initially, to a sort of "causal," or passive (Pillow, 1988, 1995), conception of beliefs—an understanding that amounts to what Chandler and Boyes (1982) as well as others (e.g., Wellman, 1990) have called a "copy theory" of mind. By this account, young persons first begin the epistemic enterprise of understanding their own and others' beliefs by effectively treating minds as "passive recorders" that simply "bear the scars of information which has been imbossed upon them" by the external world (Chandler & Boyes, 1982, p. 391). Such a copy theory naturally leads children to reduce all potential "knowers" to mere patients whose thinking is understood to lack just those constructive, or agentic, components that will later become the focal point for their more mature, adult-like theories of mind. In fact, it is just this evident absence of any such early understanding of epistemic agency that provides the basis for the current suggestion that young children possess a limited understanding of moral responsibility, at least until they come to a fuller appreciation of how people's beliefs might be characterized in the language of active, interpreting agents.

The general prospect that young persons come only gradually and fitfully to the view that mental life necessarily involves an active or constructive, as opposed to only a reactive or passive, aspect has been a topic of some real, but limited, standing in the recent theory of mind literature (e.g., Perner & Davies, 1991; Pillow, 1988, 1995; Ruffman, Olson, & Astington, 1991; Sodian & Wimmer, 1987; Wellman & Hickling, 1994), and, more particularly, the central focus of a line of research explicitly dealing with children's understanding of interpretation (Carpendale, 1995; Carpendale & Chandler, 1996; Chandler & Carpendale, 1998; Chandler & Helm, 1984; Chandler & Lalonde, 1996; Lalonde, 1996; Sokol & Chandler, 1998). In contrast to the singular focus on children's false belief understanding that has characterized the large bulk of contemporary work on children's theories of mind, this new program of research explicitly takes up the question of when, in the usual developmental course, young persons first come to view the knowing process as involving a bi-directional relationship between the "world" and the "mind". That is, even very young children are quick enough to appreciate that every act of knowing necessarily involves some reactive accommodation of the mind to match inputs supplied by the external world. What they are missing, however, is any real appreciation of the fact that "knowing" also ineluctably involves a second, "world-to-mind direction of fit" (Searle, 1983) through which the mind actively transforms the world by assimilating it—that is, making it "fit"—the mind's own already present knowledge structures. It is because, as adults, we ordinarily subscribe to such an "interpretive" theory of mind that we are prepared to allow for the possibility that two or more persons with precisely the same information at their disposal can, and regularly do, still succeed in coming to sharply different views, or readings, of a particular shared experience. On the basis of this new research (see particularly Carpendale &

Chandler, 1996, or Chandler & Lalonde, 1996) it has become increasingly clear that the initial signs of this richer understanding of others' active epistemic lives appear relatively late in children's socio-cognitive growth. Specifically, it is not until young children are typically 7 or 8, and so, well beyond the tender ages of when false belief understanding is already well in place, that they first begin to negotiate such an interpretive turn and show indications of a genuinely constructivistic theory of mind.

#### Turning to the Law for Clarification: Mistakes of Fact and Mistakes of Law

Following from this line of reasoning, the broad thesis, which this research undertaking is meant to test, is that a constructivistic theory of mind serves to provide the necessary underpinnings for that sense of epistemic agency that, as I have been arguing, is also the basis for a more adult-like conception of moral accountability and the "ownership" of one's actions (see Blasi, 1991, p. 23; Blasi, 1995, p. 239). Still, it is one thing to have broad theoretical reasons for exploring such a proposed relationship, and it is quite another to work out the specific ways in which such a relation might manifest itself in the actual reasoning of young children, or, for that matter, in the thoughts of adults. To this end, my research colleagues and I have been pushed rather far afield in search of some potential "source model" for envisioning the possible relations between matters of agency, intention, and interpretation, and their role in the framing of judgements concerning moral responsibility. In the end, this search has led us to the domain of jurisprudence, where all of these issues, as it turns out, manage to converge in the legal canon under the general rubric of mens rea, or, more simply, culpable mental states (Kadish & Schulhofer, 1995).

Among the broad range of legalistic concepts that are drawn together under this area of the

law, two notions have proven to be of particular relevance. Specifically, we have begun trading on what those in the legal profession refer to as the distinction between “mistakes of fact” and “mistakes of law” (Black, 1990, p. 1001). What, for purposes here, is the real nubbin of this otherwise complex legal distinction is that “mistakes of fact” refer to ways of getting things wrong that can be wholly laid at the door of ignorance-based false beliefs (e.g., one might well know all about the existing seatbelt law in some adjacent state, but simply be ignorant of the fact of having inadvertently wandered across the state line, leaving one duty bound to buckle up). “Mistakes of law,” by contrast, turn on errors of interpretation regarding the legality, or (in the case being presented here) morality, of a given action (e.g., one might know all about the existence of a seatbelt law, without having imagined—not even in a million years—that being parked in lovers’ lane actually qualifies as “conducting” an automobile). As this distinction would imply, ordinary or blissful ignorance, based on not having clear access to all the relevant and critical facts of some matter (as opposed to “willfully” choosing to ignore such facts), tends to be seen by the law as a bonafide mitigating excuse. By contrast, misinterpreting the law is standardly seen as an all together worse and, often times, unacceptable excuse. This is understood to follow for the reason that in venturing some homespun interpretation of events, it behooves one to do a good job at it. Failing to do so by seizing on some interpretation that is incoherent or poorly warranted by available evidence is, consequently, seen to bring righteous approbation down on one’s own head, and to be all together more blameworthy than simply being blindsided by an unexpected fact.

Such a legalistic way of doing business, I would suggest, highlights a commonly held and morally relevant assumption about the relationship between persons and the mind. That is, it

points to an often tacit conception of mental life according to which the knowing process is taken to be the work of active epistemic agents who are obliged to necessarily interpret the world around them in ways for which they are expected to take some moral responsibility. By this shared standard, then, mistakes of fact, or at least all those forms of counterfactual belief that arise simply because one was inadvertently in the wrong place at the wrong time, can be laid off to the doings of an environment perpetrated on people as patients, and who can be reasonably excused from responsibility for things about which they were uninformed. By contrast, badly misreading one's circumstances and, as a consequence, behaving in some way that runs afoul of legal and often attendant moral prohibitions (that is, by making a so-called mistake of law) naturally invites recrimination.

#### Distinguishing Simple Ignorance from Genuine Interpretation: What Do Children Think?

For those at all acquainted with the current burgeoning body of research literature detailing children's changing theories of mind, the rehearsal of this legal distinction should touch a familiar chord. That is, the parallel between the research findings concerning children's changing conceptions of mental life, on one hand, and these two adult rules of thumb for establishing relative degrees of culpability, on the other, is, at least as my research colleagues and I see it, clear enough. The recognition of a so-called mistake of fact, in particular, is just that parallel instance for an adult of what a card-carrying theorist of mind would label as a demonstration of false belief understanding—or knowing that, as a result of simple ignorance, people can come to take as subjectively true things that are, in fact, objectively false. Clearly, any young person capable of passing one of the now standard “unexpected transfer” or “unexpected contents” tasks (i.e., tasks that have become the accepted litmus test (Wellman, 1990) of early false belief understanding)

ought to be capable of recognizing what the law would deem to be a mistake of fact, and use this insight as a tool in tempering their judgments of legal or moral responsibility. Put more concretely, if there were a law against searching in the wrong place for one's missing chocolate bar, then Wimmer and Perner's (1983) much imitated protagonist "Maxi" should be seen by most 3- and 4-year-olds as not only having violated such a law, but also as deserving some relief from his legal responsibilities for doing so because, through no fault of his own, he had simply been in the wrong place at the wrong time when the chocolate was moved. Still, it is important to note, as I have worked to make clear elsewhere (Chandler & Sokol, in press), that nothing in this description touches on those more genuinely interpretive matters that, in the case being presented here, relate to so-called mistakes of law.

The particular turn of the epistemic wheel that brings children closer to those interpretive issues associated with such mistakes of law is, I contend, their achievement of having come to what I and my colleagues have been calling a constructivistic theory of mind (Carpendale & Chandler, 1996; Chandler & Lalonde, 1996). That is, a proper understanding of a mistake of fact, or, as I have suggested, any standard false belief scenario, requires no more than a simple understanding of the old cliché that "seeing is believing." Children old enough to have come to the simple realization that one has to at least "see" an event occur in order to know something about it, and who are equally capable of reversing this logic by also appreciating that "not seeing" has the consequence of leaving one in the dark about particular matters, have all the cognitive prerequisites necessary for understanding what it means to make a "mistake of fact." By contrast, coming to appreciate that people can reach different and sometimes badly mistaken views, even when they have seen everything there is to see about precisely one and the same

thing, requires (in addition to simple false belief understanding) some basic comprehension of interpretation, and how, in particular, the knowing process is informed by an active, mind-to-world contribution on the part of human agents.

Given the foregoing line of reasoning, it follows that it is not until children have arrived at an interpretive, or constructivistic, theory of mind that the way is also opened for them to understand that, when faced with some morally hazardous judgment call, circumstances could very well turn in such a fashion as to lead an individual into some ill-conceived interpretation, and consequently into some mistake of law. That is, if misinterpretation is possible, then there is nothing to prevent someone from running afoul of the law by somehow putting an inappropriate spin on the events of their experience. One could expect, then, that young persons mature enough to appreciate that people can fall into error, either because they are ignorant of relevant facts, or because they manage to badly interpret their circumstances, might come to side with adults in holding others more responsible for their active misinterpretations than for their passively acquired ignorance by judging mistakes of law more blameworthy than mistakes of fact.

Taken together, these various considerations amount to a set of testable expectations concerning how those with and without a constructivistic conception of mental life are likely to respond to stimulus materials involving a story protagonist's misguided actions that result from either ignorance-based mistakes of fact, or more interpretive-based mistakes of law. First, there is the broad expectation that the moral reasoning of children with a constructivistic theory of mind will more closely resemble that of adults than will their non-constructivistic counterparts, and that this difference will reflect their newfound ability to more fully appreciate the role of personal agency in shaping their own and others' actions and choices. This follows from the



general argument that children who lack an interpretive understanding of belief formation, and so still harbor in some fundamental sense a passive conception of persons' own epistemic lives, will tend to view themselves and others as lacking any moral accountability for what amounts in their eyes to mere, epiphenomenal thoughts about the world. Second, and in a closely related fashion, it is expected that only children who subscribe to such a constructivistic theory of mind will be able to appreciate the difference between cases involving mistakes of fact as opposed to mistakes of law. As a consequence, only these constructivistic reasoners will adhere to the adult-like view that a story protagonist who falls into error by actively misinterpreting his circumstances should bear more responsibility for his "epistemic" negligence than should a counterpart story character whose misdeeds arise merely as a result of guileless ignorance. In contrast, children without such a constructivistic understanding of mental life will fail to see any real light between any two such morally relevant mistakes, and so will collapse interpretively driven mistakes of law into the more passive and easily understood category of mistakes solely attributable to simple ignorance. They, it is predicted, will consequently assign equal responsibility to those whose wrongdoings follow from what, to more mature eyes, are seen as sharply different excusable mistakes of fact on the one hand, and more reprehensible mistakes of law on the other.

#### Choosing a Procedural Approach: A Revival of "That Highly Moral Drama" — Punch and Judy

Two things were clearly needed in order to test these various expectations: (1) a set of child friendly testing materials in which story characters could be presented as navigating various legally and morally hazardous circumstances that arise from either ignorance-based mistakes of fact, or through interpretive driven mistakes of law; and (2) an easily understood response measure by means of which 5- to 7-year-olds (i.e., children of an age that previous research has

suggested are on the cusp of coming to a constructivistic theory of mind) could potentially assign different degrees of moral responsibility to these characters.

Following a procedural course laid down a quarter of a century earlier by investigators concerned with related questions about the place of intentions and consequences in children's moral judgments (e.g., Anderson, 1980; Buchanan & Thompson, 1973; Costanzo, Coie, Grumet, & Farnill, 1973; Farnill, 1974; Leon, 1980; Surber, 1977), it was easy enough to arrive at a whole cafeteria of well practiced, Likert-type response measures, all designed for use in rating children's judgments about relative degrees of "badness", "naughtiness", moral responsibility, etc. The particular response measure selected from this arsenal consisted of a purpose-built 5-point "badness meter" by means of which children could express their views about someone's lack of moral rectitude by moving an arrow up (or down) a continuous scale marked off by lengthening rows of one to five faces of "culprits" signifying various levels of "badness." After a brief training session, this device proved to be an easily managed (and, perhaps more importantly, an engaging) way for young children to rate how "bad," or wrong, they judged various actions to be.

In contrast to the available supply of easily modifiable response measures left lying about by a previous generation of moral judgment investigators, no appropriate stimulus materials intended to dramatize the differences between mistakes of facts and mistakes of law were ready to hand, and it became necessary to fabricate these from the ground up. Drawing upon a suggestion by the philosopher Daniel Dennett (1978, 1997), the prototype case that my research colleagues and I eventually came to was afforded by the almost 500-year-old case of Punch and Judy (Baird, 1973), and their so-called "highly moral drama" (Damon, 1957, p. 5). As Dennett (1978) has pointed out, one particular vignette in the standard Punch and Judy repertoire has all

of the makings of what he termed a “minimally complex” measure of false belief understanding. In this episode, upon which Wimmer and Perner’s (1983) unexpected transfer measure is loosely patterned, Punch intends to do away with Judy by tying shut a box into which she has fallen, and then pushing it off of a cliff. While Punch is off stage searching for rope, he (but not the audience) fails to see that Judy manages to slip out of the box just before he returns to push, what others know to be an empty container, off the stage—an act that, by any account, qualifies as a failed attempt to commit murder. Using this prototype story as a springboard to other similar narratives involving cases of both simple false belief and misguided interpretation, a total of four Punch and Judy puppet shows were produced in a video format, all of which involved Judy somehow falling into a box and Punch being mistaken as to her true whereabouts. More specifically, two of these episodes consisted of cases of “involuntary manslaughter” and turned on the fact that, without malice of forethought (i.e., either out of ignorance or through some interpretive error), Punch threw over a cliff a box into which, unbeknownst to him, Judy had accidentally fallen. In the two remaining episodes, both of which qualify as a failed attempt to commit murder, Punch tried to take his revenge by capitalizing on the fact that Judy had fallen into a box that he then planned to push over a cliff. Here, paired with the classic episode in which Punch's plot to do away with Judy fails out of simple ignorance, was a second vignette in which Punch once again plotted to throw his wife off a cliff, but again failed, this time as a consequence of having wrongly interpreted a message concerning her true whereabouts. By presenting various combinations of these puppet shows to participants whose job it was to rate Punch's "badness" for acting as he did, it became possible: (a) to determine which respondents did and which did not act out of a constructivistic theory of mind; and (b) to evaluate the

hypothesis that only those with an evident appreciation of how misinterpretation is possible are able to distinguish and appropriately rate so-called mistakes of fact and mistakes of law. This was done twice, once with a sample of adults, and once with a group of 5- to 7-year-old children.

### **Study 1: Characterizing the Endpoint—Adults' Responses to Punch and Judy**

Before putting these stimulus materials and response measures to work in evaluating the abilities of young children, it seemed prudent to first road-test them with a sample of adults who could be expected to generally subscribe to a constructivistic theory of mind. It was anticipated that, if given the proper opportunity, such adults would tend toward leniency in judging those whose wrongful actions were the result of simple ignorance, while judging more harshly those whose mistakes were the product of their own ill-conceived interpretations. Because, as with all such normative expectations, it was anticipated that any group of adults would naturally include persons who subscribed more or less wholeheartedly to a fully constructivistic theory of mind, the appropriateness of the particular methods and procedures so far developed was understood to depend on their ability to bring out different patterns of response that, in turn, would serve as rough benchmarks against which the responses of children could be evaluated. In order, then, to "proof" these procedures, and to fine tune them for use with a target population of 5- to 7-year-olds, they were first pilot tested with a convenient sample of adults.

### **Method**

**Participants.** A total of 56 adults (mean age=22.3 years, SD=5.1 years, range=17.9 to 40.7 years) participated in this first study. Of these, 26 were men and 30 were women. All participants were university students (none from the school of law) of diverse racial and ethnic backgrounds.

Stimulus Materials. The Punch and Judy vignettes that made up the four stimulus conditions turned primarily on whether Punch was either ignorant of Judy's whereabouts because of his absence from the stage during a critical moment of the plot ("False Belief" Scenes 1 and 2), or confused about Judy's location because of interpreting her remarks in a way that proved incorrect ("Interpretive" Scenes 3 and 4). Both sets of scenes included parallel instances of Punch either trying, but failing, to harm Judy (i.e., scenarios that qualified as instances of "attempted murder") or inadvertently managing to "do her in" by unwittingly throwing off the stage a box in which she coincidentally had become trapped (i.e., cases of "involuntary manslaughter.") The narratives associated with these four conditions unfolded as follows:

*False Belief Scenes*

1. Attempted Murder. Punch maliciously plans to push a box into which Judy has stumbled out into the trash, but while he is off stage searching for string to secure the box, she escapes and hides. When the unwitting Punch returns, he ties shut what everyone but he understands to be an empty box, and pushes it directly into the trash, mistakenly believing that he has gotten rid of Judy once and for all.

2. Manslaughter. Judy accidentally falls into a box while Punch is off stage. When Punch enters the scene with the good intentions of cleaning house, he unwittingly throws the box containing Judy into the trash. Although he presumably kills Judy in the process, he is ignorant of having inadvertently done so.

### *Interpretive Scenes*

3. Attempted Murder. Punch and Judy are on stage, accompanied by two boxes: a box painted orange and a box into which they are together loading oranges. With Punch off stage, Judy stumbles while on-loading more oranges and shouts for help, announcing that she has fallen into “the orange box.” When Punch returns primed to take advantage of the situation by throwing Judy into the trash, he fails in his plan, not out of simple ignorance, but by wrongly interpreting Judy’s statement of “I am in the orange box” to mean she is in the box that is painted orange. Like Condition 1, then, Punch pushes an empty box off the stage, wrongly believing that Judy is trapped inside.

4. Manslaughter. Judy is by herself on stage with two boxes: one painted green and another that is white, but marked with a large green number “1.” Punch is off stage when Judy falls into the green colored box. When she calls for assistance, a well-intentioned Punch enters inquiring about how to help; She reports: “Check the green one.” Punch interprets this to mean that she is in the box marked by the large green number “1,” whereupon he innocently moves the green colored box, into which Judy has actually fallen, out of the way by pushing it off the stage. Similar to Condition 2, then, Punch innocently pushes Judy into the trash, this time, however, as a result of having misinterpreting her cry for help.

Procedure. The adult participants were tested in groups of five or more using a paper-and-pencil version of the same interview questions intended for later use with the 5- to 7-year-old children of Study 2. Because the Punch and Judy story materials were created for such use

with young children, these adult participants were advised to overlook the child-centered character of these videos, and to, instead, focus their attention on the potential seriousness of Punch's actions. In particular, they were instructed to recognize that Punch's attempts to throw Judy in the trash were, in fact, life threatening.

In a similar attempt to produce a response measure better suited for adults, the "badness meter" developed for use with young children was replaced with a more standard 5-point rating scale. The participants were asked to determine, in a manner likened to that of a judge deciding a criminal's guilt, just "how bad," or wrong, they held Punch's various actions to be by circling a number on the scale between 1 and 5. Before beginning the story procedure, the participants rated several prototypical social and moral violations of varying severity levels (Nucci, 1981; Nucci & Turiel, 1978; Smetana, 1981; Tisak & Turiel, 1988), both as a way of prompting them to use the full range of the scale and, more importantly, as a means of obtaining a baseline rating with reference to which it was possible to compare their judgments regarding the seriousness of intentionally harming others in life-threatening ways with the cases of Punch's behavior (see Appendix 1).

Following each of the four videotaped Punch and Judy shows, these adult participants were asked to answer a set of questions and to make their rating. To control for potential order effects, the presentation of the story conditions was counterbalanced such that half of the adult participants saw the set of false belief scenes first and the interpretive scenes second; the other half viewed the vignettes in the reverse order.

Story Questions. Written responses, including appropriate justificatory remarks, were obtained following each scene. Because the questions were open-ended, participants were

instructed to focus their responses by answering only the specific questions being asked, and to draw on only those details provided in the particular scene in question (e.g., they were asked not to respond to questions about Scene 3 on the basis of information provided in Scene 1).

In the two False Belief sequences (i.e., Scenes 1 and 2), participants were asked, not only to rate how wrong or bad Punch's actions were, but to provide an explanation of why they made the rating they did. In the scene involving Judy's escape from the box that Punch was preparing to throw away (Scene 1), participants were directly asked about Punch's beliefs concerning Judy's whereabouts as he pushed the box off stage. Participants' responses to this line of questioning were taken as a measure of their false belief understanding.

Questions concerning the two interpretive scenes (i.e., Scenes 3 and 4) in which Judy made an ambiguous statement about her whereabouts (i.e., "I'm in the orange box" and "Check the green one," respectively) were used to code whether participants did or did not subscribe to a constructivistic theory of mind. In particular, three additional questions were put to the participants prior to their evaluating Punch's behavior (see Appendix 1 for details). These included requests for them to explain the reasons which guided Punch's box selection, to evaluate Judy's remarks, and to determine who was to blame for the resulting misunderstanding concerning Judy's true location, as follows:

1. Explanation for Punch's box selection.—Following each of the Interpretive Scenes, participants were first asked to explain how Punch managed to choose the wrong box in each particular case. Respondents were specifically directed to organize their answers to this question around the verbal instructions that Judy had provided.



2. Evaluation of Judy's statement.—Participants were then asked whether Judy's ambiguous directions concerning her whereabouts were "really wrong." Here Judy's exact statements of either "Check the green one" or "I'm in the orange box" were repeated while still pictures of the boxes for the appropriate scene were held up for participants to see.

3. Determination of blame for the misunderstanding.—Finally, just before asking them to rate Punch's action on the 5-point scale, participants were asked about who should be blamed for the "mix up," or misunderstanding, about which box Judy had fallen into. Four choices were provided from which participants were to choose one: Punch, Judy, Both, or Neither. Participants were asked to give their reasons for their selection.

Scoring a Constructivistic Understanding. Although it was expected that all the adult participants would appreciate the ambiguity in both Interpretive Scenes 3 and 4, responses indicating an adequate understanding of the interpretive problem in only one of these two conditions was taken as sufficient for coding them as subscribing to a constructivistic theory of mind. This perhaps liberal scoring criterion was adopted for the reason that it coincided with scoring standards employed in Study 2 with children, who, because of their more fragile and often situationally-dependent understanding of such interpretive matters, might have difficulty meeting more stringent standards.

More specifically, to be scored as holding to a fully constructivistic conception of the mind, participants needed to make clear three critical points in responding to the set of questions just described. First, they had to somehow recognize that Punch had understandable reasons for

his mistaken belief concerning Judy's whereabouts as a result of having simply taken up what qualifies as one possible interpretation of her ambiguous directions ("Explanation" question). Second, they needed to avoid labeling Judy's statement as simply wrong, by indicating instead that her remarks, while technically correct, were unclear or confusing in the context that they were offered ("Evaluation" question). Third, and finally, participants had to give some indication of the fact that getting clear about such interpretive problems imposed a two-way obligation, both on the part of Judy, who was under an obligation to eliminate ambiguity in her remarks, and on the part of Punch, who should have asked for clarification when matters proved to be confusing ("Determination" question). Taken together, the clearest responses to these three probes should reflect, then, not only a full understanding of the potential double meaning that was present in Judy's utterance, but also an appreciation of the fact that Punch and Judy's different understandings of what was said (based on their differing interpretations of how the boxes should be referenced) both represent defensible, or at least, not unwarrantable, points of view. Moreover, such a constructivistic way of viewing things should include an understanding of the shared responsibility for the interpretive mix-ups that resulted. That is, because interpretive misunderstandings are ultimately the result of two (or sometimes more) persons' differing interpretations of the same matters at hand, both parties have an obligation to be clear about what is said and meant.

### Results and Discussion

Step one in the overall data analysis involved first coding participants as clearly expressing or not expressing a constructivistic theory of mind. As will be more fully described later, there was no a priori reason to suspect that all of these adult participants would

automatically succeed in equal measure at giving voice to the broadly constructivistic theory of mind to which ordinary adults are assumed to subscribe. In order to allow for this prospect, a scoring distinction was introduced that allowed not only for the possibility that some participants might entirely fail (although nobody did), but that they might also evidence consolidated and less than fully consolidated versions of a constructivistic view. Even though this coding process was largely mechanical, a second rater independently scored half of these protocols. There was 96% inter-rater agreement in assigning participants to one of these three categories, Cohen's  $\kappa = .94$ .

The remaining steps involved the application of various statistical procedures. For these, it should be noted, the responses of three participants were omitted: two because they did not appropriately justify their responses by restricting their judgments to the particular scene they had observed, and one for not offering any justificatory remarks at all. In total, then, the responses of 53 adults were analyzed (23 male; 30 female). Because preliminary analyses showed no statistically significant differences between the response patterns of men and women, or between those who viewed the video tapes in different orders, all the participants' responses were combined in a repeated-measures ANOVA examining their average ratings of Punch's actions across each of the conditions (including their baseline rating of intentionally inflicting personal injury on another). This analysis revealed a significant effect for condition,  $F(4, 208) = 134.84$ ,  $p < .0001$ , which was further explored in subsequent post hoc analyses of all ten of the possible pairwise comparisons between the five conditions.<sup>1</sup> Each of the resulting findings is discussed in relation to participants' responses to the questions meant to examine their false

---

<sup>1</sup> In accordance with Bonferonni-Dunn's procedure, the error rate was appropriately adjusted to .005 in order to reflect a familywise alpha of .05 for these various comparisons; all tests were two-tailed.

belief and interpretive understanding, as well as their explanations for the ratings of Punch's actions.

**False Belief Conditions.** Without exception, all the adult participants easily recognized that, because he had been in the wrong place at the right time in both Scenes 1 and 2, Punch was laboring under a legitimate false belief about where in fact Judy was. That is, it was uniformly understood in each case that, because of his untimely absence, Punch simply could not know that Judy fell into the box in Scene 2, or escaped from it in Scene 1. In turn, participants' recognition of these facts appeared to have a direct influence on their ratings of Punch's actions in each case (see Figure 1).

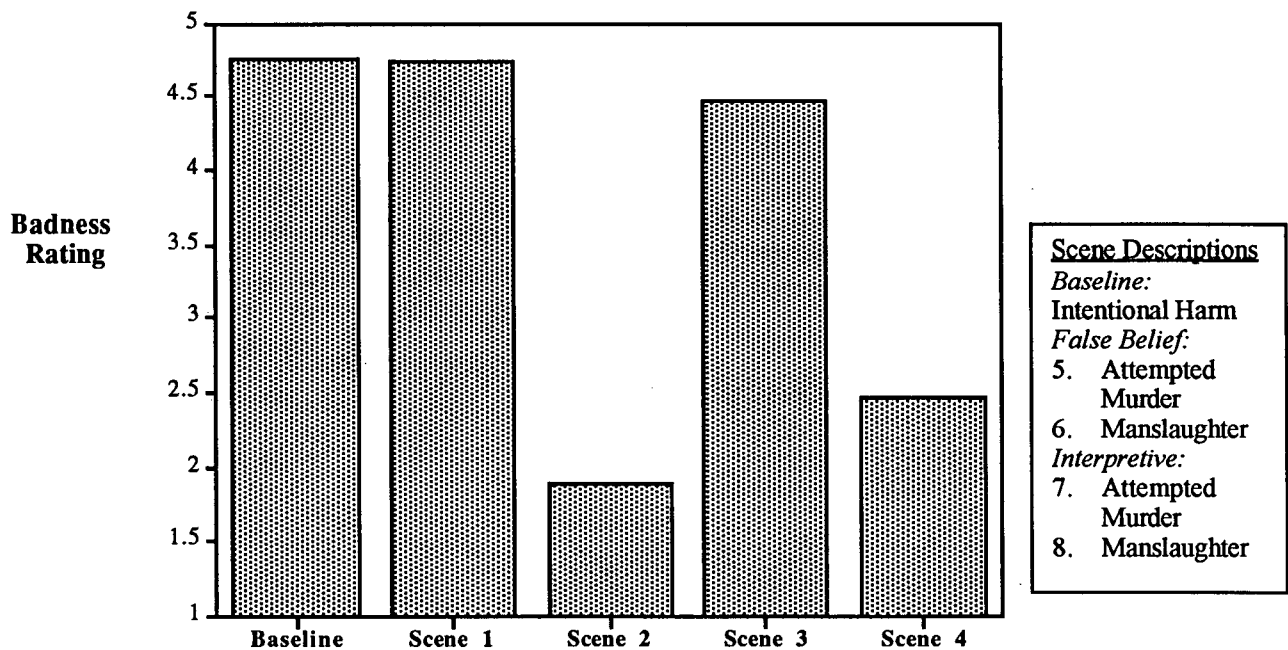


Figure 1: Adults' "Badness" Ratings

In the first scene, involving Punch's failed attempt at murder, the average rating of the act on the 5-point "badness" scale was nearly at ceiling, 4.74 ( $SD=.68$ ). When compared to the more "generic" baseline rating of 4.76 ( $SD=.68$ ) for acts of willful injury to others that participants produced during the preliminary phase of the procedure, it was clear that, insofar as he was rated near the extreme end of the scale, Punch might as well have succeeded in carrying out his murder plot. That is, as these two closely matched ratings suggest, both the actions of trying to kill somebody and actually doing so were viewed by the adult participants as being virtually identical, and so, the trivial difference that was observed here failed to approach statistical significance,  $t(52)=.16$ ,  $p=.88$ . The case was just the opposite, however, with Scene 2 involving Punch's ignorance-driven act of involuntary manslaughter. That is, despite the fact that in this story condition Punch actually succeeded in pushing Judy into the trash, the adult participants took his ignorance, and so, his evident lack of malicious intent, as an acceptable mitigating factor. Besides frequently mentioning this fact in their justificatory remarks, the diminished responsibility attributed to his act was clearly reflected in the participants' ratings, which on the whole averaged only 1.89 ( $SD=1.17$ ) on the 5-point rating scale. When compared to both the baseline condition and their responses to Scene 1, this average score for Scene 2 differed by nearly 3 full points,  $t_s(52)=16.62$  and  $16.24$  respectively,  $p_s<.0001$ .

Interpretive Conditions. Although adult participants' responses to Interpretive Scenes 3 and 4 proved to be somewhat more varied than expected, two points, at least, were clear enough when it came to their ratings of Punch's actions. First, in much the same way that participants rated Punch's actions in the False Belief Scenes according to whether he was guilty of either attempted murder (Scene 1) or involuntary manslaughter (Scene 2), they again recognized the

parallel circumstances at work in Interpretive Scenes 3 and 4, and so, again evidently took this difference into account as they evaluated his actions, 4.47 ( $SD=1.04$ ) and 2.47 ( $SD=1.09$ ), respectively,  $t(52)=10.5$ ,  $p<.0001$ . Second, when compared to ignorance-based Scenes 1 and 2, participants' responses to these more interpretive scenes revealed a clear appreciation of the distinction between mistakes of fact and mistakes of law that was hypothesized to underpin adults' judgments of moral responsibility. That is, in comparing the "badness" ratings produced by participants in response to the ignorance-based act of manslaughter in Scene 2 with the ratings of Scene 4, where Punch behaved in an almost identical fashion but this time as a result of misinterpretation, it was evident that, as hypothesized, interpretive errors were held to be more blameworthy than mistakes due to simple ignorance. Specifically, the average "badness" ratings for these two scenes were 1.89 ( $SD=1.17$ ) and 2.47 ( $SD=1.09$ ), respectively,  $t(52)=3.94$ ,  $p=.0002$ .

More or Less "Consolidated" Varieties of Constructivism. The unexpected variation in adults' responses came by way of the range of answers offered to the "Determination" question, having to do with who should be blamed for each of the misunderstandings that arose in Interpretive Scenes 3 and 4. The general expectation had been that, when presented with an opportunity to apportion blame, ordinary adults, all of whom are imagined to be capable of subscribing to a thoroughgoing constructivistic understanding of the knowing process, would find both Judy and Punch negligent for having failed to either communicate clearly or for not seeking clarification in what amounted to a life-threatening situation. That is, if Punch and Judy were in some sense each equally right to read the ambiguous conditions in the opposing ways that they did—a point with which all of these adult participants agreed—then both puppet characters

should be seen as contributing in some measure to the confusion that resulted, and so ought to be jointly seen as blameworthy. Put concretely, then, essentially all of the adults in this sample were originally expected to choose the “both” option when asked to indicate who was responsible for Judy ending up in the trash. What was apparent from the data, however, is that not all adults actually exhibited such a “consolidated” view, or at least failed to do so in this assessment context. That is, 51% of the respondents, although clearly indicating in the “Explanation” and “Evaluation” questions that both Punch and Judy held warrantable views of the ambiguous conditions, did not appear to follow through on the logical implications of these insights when deciding which puppet characters should be held responsible for the interpretive mix-up that resulted. As it was, then, just 49% of the adult participants demonstrated a thoroughgoing, “consolidated” understanding by properly answering the “Determination” question.

To determine just how much of an impact this difference in performance might have had on the relative ratings of Punch's actions, a subsequent 2 x 5 (Epistemic Stance x Condition) repeated measures ANOVA was conducted in which those who demonstrated a “consolidated” understanding and those who “did not” were treated as two distinct groups. The resulting interaction was examined [ $F(4, 204)=3.46, p=.009$ ], again using Bonferroni-Dunn's post hoc comparison procedure. Although for the most part these two groups consistently rated Punch's actions in the same way across the first three conditions, an important difference was revealed in Scene 4—the one opportunity where participants could express the view that, because interpretative errors involve the active contribution of more than one person, some sort of “distributed liability” must then exist that wouldn't allow Punch off the hook so easily. As it

was, while adults who were "consolidated" in their reasoning about these interpretive problems saw every reason to differently evaluate Punch in the two manslaughter conditions, the half of those participants who evidenced a "non-consolidated" epistemic stance in response to these problems simply overlooked any reasons why these otherwise similar cases should differ. That is, only those with a consolidated epistemic stance distinguished between False Belief Scene 2 and Interpretive Scene 4 by rating them differently, 2.08 ( $SD=1.23$ ) and 3.04 ( $SD=.87$ ), respectively,  $t(25)=4.91$ ,  $p<.0001$ . The implication of this difference is that while all the adults in the sample deserved some credit for subscribing to a constructivistic understanding of mental life, it is still the case that only those individuals with the fullest, or most thoroughgoing, constructivistic views actually make a strong distinction between mistakes of fact and mistakes of law by treating interpretive-based errors as being more blameworthy than those involving simple ignorance.

### **Study 2: The Child's Point of View**

Although the questionnaire-style format employed with adults in Study 1 is obviously different from the one-on-one interview procedure used with children in Study 2, the two studies are, in all other respects, procedurally identical. That is, following the intended road-testing efforts of Study 1, only slight modifications were made to Study 2's interview procedure that, for the most part, simply involved adding appropriate sorts of memory probes and control questions that were seen as necessary to confirm that children actually understood the videotaped stories and the questions asked about them. The only other alterations to the interview schedule involved providing a more highly structured response format for the key questions used to determine participants' interpretive competence. In particular, this involved changes to the



response options of the "Determination" question. Specifically, there was some concern that the four available response options (i.e., Punch, Judy, Both, or Neither) used in Study 1 permitted a certain ambiguity concerning participants' intended meaning in answering the question of who was responsible for the interpretive misunderstanding in Scenes 3 and 4. As a result of a careful review of adult participants' justificatory remarks to this question, as well as a following-up interview with nearly half of the adult respondents about their understanding of the various response options, it became clear that had they been questioned as to whether the misunderstanding between Punch and Judy was either entirely the fault of one or the other puppet character, or was, instead, somehow due to a failure on both of their parts, then nearly all of these adult respondents would have selected the "both" option. As it was, however, the "either," "neither", or "both" response options actually given to adults apparently pulled for participants to answer by indicating who they thought was "primarily" to blame (e.g., although some believed that Punch contributed somewhat to the resulting mix-up concerning Judy's true whereabouts by not seeking needed clarification, they still thought Judy was primarily to blame for evidently misspeaking, and so, saw this as a better response option than "both"). The procedural modification that was introduced in Study 2 in an effort to better elicit young participants' clearest thoughts about the key "Determination" question consisted, then, of narrowing the response options for apportioning blame to either "All Judy's Fault," "All Punch's Fault," or "A Little of Both." Only children who chose this latter option were coded as demonstrating a "fully" constructivistic view. Those children, on the other hand, who successfully passed all but this "Determination" question were coded as "transitional" in their interpretive competence. That is, children coded as transitional made it very clear in their

responses to the "Explanation" and "Evaluation" questions that they at least understood how the ambiguous context of Judy's message could leave both Punch and Judy with different, though warrantable, interpretations, but otherwise failed to follow through on the implications of this insight when it came to appreciating that each, by virtue of their active contribution to the interpretive misunderstanding, deserved a share of the blame for what happened.

### Method

Participants. A total of 54 children (27 boys and 27 girls) were recruited from local schools and after-school care programs for participation in this study. Four boys and two girls were dropped from the study because they either failed to demonstrate an adequate understanding of how to use the badness meter, or were unable to pass the various memory probes and control questions. Of the remaining 48 children, there were 15 kindergartners (mean age=5 yrs, 5 months, SD=4 months, range=59-70 months), 16 first-graders (mean age=6 yrs, 7 months, SD=3 months, range=74-83 months), and 17 second-graders (mean age=7 yrs, 8 months, SD=6 months, range=85-103 months). These child participants were from predominantly middle class backgrounds, and were ethnically diverse.

Procedure. All four of the filmed Punch and Judy vignettes described in Study 1 were presented to individual children in a one-on-one setting. As in Study 1, half the children saw the false belief scenes first and the interpretive scenes second; the other half observed the scenes in the reverse order. Following each vignette, children were queried using a semi-structured interview procedure meant to determine their understanding of information being presented in the conditions, and, particularly, to gauge their interpretive abilities. Just as with the adults, the child participants were asked to rate just how wrong they thought Punch's behavior was in each of the

four scenes, using the "badness meter" previously described. In order to insure that these young participants understood how to use this rating device, and as a means of securing a baseline measure of how they judged instances of intentional harm, children were asked to rate each of six prototypical social and moral violations similar to those in Study 1. Before continuing with the procedure, children were required to demonstrate their understanding of this procedure by differently rating at least three of these six test cases (i.e., they needed, for example, to assign different "badness" ratings to cases of talking out of turn and of stealing another's favorite toy). This practice phase not only served to prompt children to use the full range of the scale, but, similar to Study 1, also made it possible to obtain a baseline rating to the question of "how bad is it to throw somebody in the garbage on purpose"—an actual event that they would later be asked to rate (see Appendix 2).

Memory Probes. To insure that children were adequately attending to the critical details in each scene, memory probes, which involved asking children about Judy's true whereabouts, began each round of questions. For the interpretive scenes, in addition, children were asked to repeat Judy's exact instructions to Punch (e.g., "I'm in the orange box" or "Check the green one") in order to confirm that they hadn't managed to somehow mishear her ambiguous statement. Still pictures of the different colored boxes involved in these scenes were used as visual aids and as a way for all participants to easily designate which box Judy was in.

Control Questions for the Interpretive Scenes. In like fashion to the "belief" question posed in Scene 1, which is detailed more below, participants were also asked following Interpretive Scenes 3 and 4 to identify which box Punch "thought" Judy to be referencing with her remarks (e.g., "I'm in the orange box" or "Check the green one"). Similarly, children were

asked to comment on Judy's belief by indicating which box she was in fact intending by her comments. Taken together with the memory probes, then, children had to be able to indicate what Judy actually said, what her intended meaning was, and what Punch thought her statement meant. Obviously, without understanding these various puzzle pieces there would be no opportunity for participants to appreciate the interpretive problem posed in Scenes 3 and 4.

Comprehension Questions—"Badness" Ratings. Before children rated Punch's behavior for each scene, a series of three comprehension questions were asked in order to confirm that their evaluations included some basic understanding of all the particular scene's critical informational elements. Specifically, this included for all the conditions a question about the extent of personal injury Judy either suffered or not as a result of Punch's doings. This "harm" question was always, however, countered with questions about Punch's evident "knowledge" and "intent." That is, in the cases of involuntary manslaughter (i.e., Scenes 2 and 4), participants were asked if Punch knew Judy's true whereabouts and whether or not the outcome of the act was intentional or accidental. In contrast, in the scenes involving attempted murder (i.e., 1 and 3), the "harm" question was offset with a question about whether Punch was indeed "trying" to hurt Judy by pushing the box off the stage. On the basis of these last sorts of "mental state" questions, a determination could be made if children had at least some basic level understanding of the role that Punch's intentions played in each of his actions.

Criteria for Inclusion in the Study. In order to be retained in the study, child participants were required to appropriately respond to the various memory, control, and comprehension questions just described for each of the scenes. Children who did not accurately respond to each of these questions were again shown the scene that they failed and directed to focus on the

relevant details. As it was, six young participants were consistently unable to pass these questions after a second viewing; consequently, their responses were omitted from the data set.

Criteria for Categorizing Children as Subscribing to a “Copy” or Constructivistic Theory of Mind. On the basis of available reviews of the theories of mind literature (e.g., Lewis & Mitchell, 1994; Moses & Chandler, 1992) all of the 5- to 7-year-old participants were expected to understand the possibility of false beliefs, and, as described earlier, to subscribe to at least a “copy” theory of mind. That is, they were all expected to understand the basic notion that persons who are left in ignorance (e.g., who didn’t see that Judy had escaped from the box while Punch was out of the room) are prone to holding false beliefs about particular matters. As a way of verifying this general expectation, all participants in Study 2 were explicitly asked following Scene 1 where Punch “thought” Judy was at the moment he tied the box shut and pushed it off the stage. Explicitly stating that Punch mistakenly believed Judy to be trapped inside the box (a response given by all 48 participants) was taken, then, as a measure of their false belief understanding.<sup>2</sup>

Like Study 1, the children in Study 2 were coded as to whether they did or did not subscribe to a constructivistic theory of mind on the basis of their responses to three critical “Explanation,” “Evaluation,” and “Determination” questions that were posed immediately after participants viewed each interpretive scene (i.e., 3 and 4). Unlike the adults who had an open-ended response format, however, the child participants were provided, when necessary, with a more highly structured set of response options for each question (see Appendix 2 for details).

---

<sup>2</sup> Because the comparability of this task with other similarly structured, standard “unexpected transfer” tasks had not been previously established, a fifth videotaped vignette was produced involving the same Punch and Judy characters that replicated in all essential ways the details of Wimmer and Perner’s (1983) classic “Maxi” task. A pilot sample of 15 5-year-old children were shown both this newly filmed sequence and Scene 1, and, without exception, the young participants either succeeded or failed on both (i.e., 11 passed and 4 failed).

Specifically, in the "Explanation" question, regarding Punch's guiding reasons for choosing the box he did, children were asked if his reasons were "good" ones or not (i.e., "silly"), and then prompted to explain why. In the "Evaluation" question, involving whether Judy's ambiguous directions were indeed wrong, participants were presented with a specified set of response choices including, in the case of being wrong, potential reasons such as simply being mistaken, possible foolishness, or intentional deception. If she was, instead, seen to be right, then participants were asked to elaborate on why and to generate an alternative statement for Judy that might be better suited for the context. Finally, and as mentioned earlier, in the "Determination" question, children's response options were constrained to being either entirely one or the other puppet's fault, or alternatively, the fault of both. The "all or none" character of these response alternatives was supplemented by a series of careful probes meant to confirm that children, in fact, were not simply attempting to sound evenhanded in their judgments when they chose the "both" option just as a way to avoid being labeled a harsh critic. In other words, an appropriate justification, in which children spelled out how they thought Punch and Judy were each responsible for the mix-up, was required before this response selection was accepted. A "fully" constructivistic view was credited to those children who, similar to the adults, adequately answered each of these three criteria questions; a "transitional" view was assigned to children who responded appropriately to only the first two questions; and, a "non-constructivistic" status was held by the remaining children who, although seeming to understand matters of simple false belief well enough, could not in any way see why the inherently ambiguous content in the interpretive scenes might be confusing.

Response Measure. There was a separate 5-point “badness meter” for each vignette. A still picture of the scene’s critical elements (e.g., Punch standing beside a box containing Judy; Punch preparing to throw away an orange colored box; etc.) was attached to each response meter as a visual aid. Instead of asking child participants to independently evaluate Punch’s behavior with no reference to their prior ratings, and thus, risk having children “reset” their evaluative stance toward each separate scene, they were presented with their previous evaluation, which was still available on the badness meter that they had just used, and asked if the current case of Punch’s morally hazardous actions should be rated as being “more, less, or the same” on the scale. The badness meters from earlier trials, then, were left in view as a visual record of children's previous responses. Consequently, children always had access to each of their previous ratings, which, it was anticipated, would better enable them to make relative judgments.

### Results and Discussion

As with the adults in Study 1, the child participants were first scored as having a constructivistic or non-constructivistic understanding of mental life. Overall, 24 children were categorized as having a non-constructivistic theory of mind, while the remaining 24 were characterized as subscribing to at least some basic level constructivistic understanding. Of those participants who held some sort of constructivistic view, 14 were scored as being “transitional,” and 10 were coded as having a “fully” achieved constructivistic understanding. Again, as in Study 1, a second rater independently coded half of the protocols. This time there was 92% inter-rater agreement on the categorization of participants, Cohen’s kappa=.88. Preliminary analyses indicated no significant order effects or gender differences, so boys’ and girls’ responses, as well

as all those who had viewed different presentational orders, were combined in the remaining statistical analyses.

Following the line of analysis previously employed in Study 1, a 2 x 5 (Epistemic Stance x Condition) ANOVA with repeated measures on the last factor was conducted to examine the various ratings offered by children who were categorized as either constructivistic or non-constructivistic. As expected, there was a significant interaction between children's differing epistemic stances and their evaluations of Punch's behavior,  $F(4, 184)=21.39, p<.0001$ . A series of one-tail t-tests,<sup>3</sup> again using Bonferroni-Dunn's adjustment procedure for setting a familywise error rate, were used to explore the statistical relationship between the various comparisons of specific interest for both within and between groups of constructivistic and non-constructivistic reasoners. These findings are broken into three parts and reported in conjunction with the set of predictions that were held regarding how children's various commitments to the general issue of human agency and intentions might manifest themselves under these testing conditions, and, more specifically, how young children viewed the difference between errant acts based on simple ignorance (i.e., mistakes of fact) versus active misinterpretation (i.e., mistakes of law).

Between-Groups Comparisons—Differences in the Use of Intention Information. The fact that all the children in Study 2 had to get past certain basic comprehension questions about Punch's intentions and the outcomes of his actions in order to be included in the present analysis should be taken as up-front evidence that an all or none intentions-consequences split, which is sometimes used to explain how children frame their judgments of moral responsibility, proves to

---

<sup>3</sup> Based on what was learned from Study 1, comparisons that were of direct interest in this study could be narrowed to just eight and the direction of their differences could be accurately predicted. Consequently, the familywise error rate was adjusted to .00625 to reflect an alpha level of .05, and, instead of two-tail tests, one-tail tests were performed.



be only part of the developmental story unfolding in these data. That is, the relevant division here is seen as one between having a more or less full conception of human agency that, by virtue of the presence or absence of a constructivistic understanding of mental life, comes to manifest itself in children's use and appreciation of intention information.

The contrasting "badness ratings" for these child participants judged to have non-constructivistic versus constructivistic theories of mind provided support for this contention that moral judgments vary as a function of epistemic stance (see Figure 2). Specifically, the ratings for constructivistic children in the attempted murder scenes (1 and 3), not unlike those of adults, were both relatively high [4.00 ( $SD=1.25$ ) and 3.83 ( $SD=1.27$ ) respectively] on the 5-point scale, and did not differ statistically from each other [ $t(23)=.60$ ,  $p=.28$ ]. As it was, these children were, then, clearly of the opinion that, despite having failed to harm Judy as planned, Punch's bad

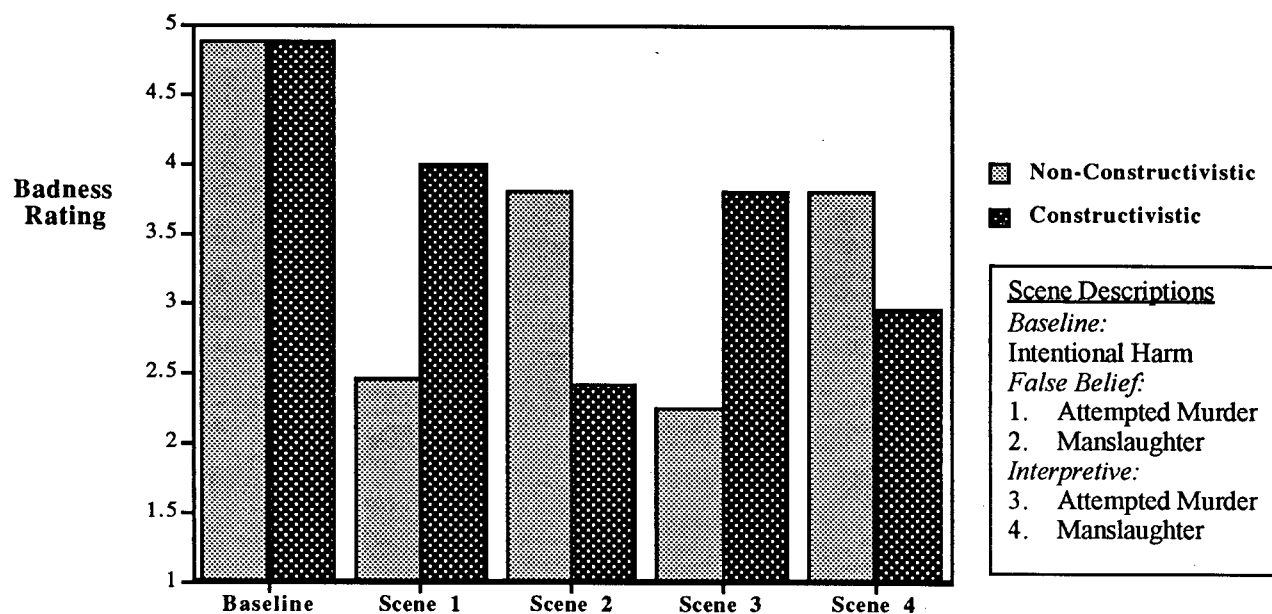


Figure 2: Children's "Badness" Ratings

intentions were nevertheless sufficient to make him guilty as sin. In direct contrast to these results, non-constructivistic children regularly focused their attention on the outcome of Punch's actions (i.e., whether Judy was harmed or not), and so, judged him relatively leniently in Scenes 1 and 3 [2.46 ( $SD=1.47$ ) and 2.25 ( $SD=1.23$ ), respectively]. The between-groups differences in these cases of attempted murder reached statistical significance. This difference, which was approximately 1.5 points on the badness scale [ $t(46)=3.91$  and  $4.39$ ,  $p \leq .0002$ ] in each case, gave some indication, then, of the sorts of diverging views of agency hypothesized to be guiding children's assignments of moral responsibility.

This hypothesis was further corroborated by the pattern of results for the two scenes involving involuntary manslaughter (i.e., Scenes 2 and 4), which proved to be just the opposite of that observed for the cases involving attempted murder. That is, this time it was the children who subscribed to a non-constructivistic theory of mind who rated Punch relatively high on the scale, 3.83 in both cases ( $SD=1.09$  and  $1.01$ , respectively), for the apparent outcome-oriented reasons that, no matter what Punch's intentions, Judy ended up in the trash. By contrast, the constructivistic participants, though also aware that Judy had been dealt with badly, were quick to excuse Punch on the basis that he was ignorant of the possible implications of his actions. That is, these children rated Punch rather more leniently [2.41 ( $SD=1.18$ ) and 2.96 ( $SD=1.12$ ), respectively]. Again, the between-group differences for each of these two cases of manslaughter (i.e., 1.42 and .88) proved to be statistically significant [ $t(46)=4.33$  and  $2.84$ ,  $p \leq .003$ , respectively], indicating that, depending on children's epistemic outlook on such matters, different parts of the problems could be seen to come into focus, and so, cause their moral judgments to diverge. Before saying anything more about these evident differences in what

children took to be the crucial elements upon which to base their judgements, however, it is important to rule out the possibility that these between-group differences are not simply a function of the participants' age.

Disentangling Age from Epistemic Stance. Because being coded as having a constructivistic or non-constructivistic theory of mind was itself significantly related to participants' age ( $r = +.39$ ), there is room to speculate that the evident connection between children's moral judgments and their epistemic stance is really an artifact of simply growing older. Consequently, it became important to try to rule out the possibility that information regarding children's particular epistemic stance simply amounted to a commentary about their date of birth. As it is, when calculated separately for the manslaughter and attempted murder scenarios, as is required by the contrastive pattern of results for these two story sets, the first order correlations between children's moral judgments and their theory of mind were  $-.46$  and  $+.52$ , respectively. These same measures of relationship are reduced in only modest ways to  $-.40$  and  $+.41$  ( $F(1,93) = 17.4$  and  $18.2$ ,  $ps < .001$ , respectively) when age is partialled out. This result, then, strongly suggests that the role of epistemic status in explaining these child participants' moral judgments is largely independent of age.

Taken together with the between-group differences previously reported, it seems clear enough from these correlational findings that the transition from an early "copy" theory of belief entitlement (in which knowledge is seen as the passive byproduct of simple exposure to of all the facts) to a more active, constructivistic stance (which obliges persons to take some responsibility for their own active interpretations of experience) has a significant impact on the way in which children frame their moral judgments. In particular, it would appear that those children

subscribing to a non-constructivistic notion of belief entitlement are restricted to a "desire-side" view of human agency that leaves them with something less than a full appreciation of what intentions (i.e., Punch's) fundamentally represent. These children, as a consequence, view only those "objective" matters of the situation (i.e., Judy's harm) as the variable of any real import in their adjudication of moral responsibility. In contrast, children who have already come to an understanding of the interpretive nature of belief, seem to also recognize the similar "subjective" core of morally relevant actions, and how particularly others' thoughts and intentions make a critical difference in determining moral responsibility.

Within-Groups Comparisons— Differentiating Mistakes of Fact and Mistakes of Law.

The analysis of differences occurring within groups of non-constructivistic and constructivistic children again closely followed the steps taken in the results section of Study 1. That is, first, the general way in which children of each epistemic category rated Punch in the two manslaughter conditions (False Belief Scene 2 and Interpretive Scene 4) was broadly examined for potential statistical differences, before going on to further sub-divide the constructivistic category into "transitional" and "fully" constructivistic reasoners. The key matter of interest here was the question of whether constructivistic children, like adults, would also differently judge these two, otherwise similar, cases on the basis of how Punch's mistaken knowledge about Judy's whereabouts actually arose. The specific purpose, then, of this within-group set of comparisons was to explore the mistake of fact and mistake of law distinction presumed to be guiding the moral deliberations of such young constructivistic thinkers.

As it was, the pattern of results for children subscribing to a constructivistic view closely resembled that of adults. That is, these children also rated Punch more harshly in the

interpretive-based instance of manslaughter (Scene 4), than in the case where his actions stemmed from simple ignorance (Scene 2). The average difference (i.e., .54) between these two scenes again proved to be statistically significant,  $t(23)=2.85$ ,  $p=.0045$ . Non-constructivistic children, on the other hand, completely failed to distinguish these cases, and, as predicted, rated Punch identically in Scenes 2 and 4.

The step of further sub-dividing the constructivistic category into “transitional” and “fully” constructivistic views shed additional light on this difference between ignorance- and interpretive-based errors (see Figure 3). Specifically, only the 10 children expressing a “fully” constructivistic view strongly differentiated between Punch's ignorance- and interpretive-based errors. Their ratings were 2.2 (SD=.79) and 3.1 (SD=.88) for Scenes 2 and 4, respectively, and

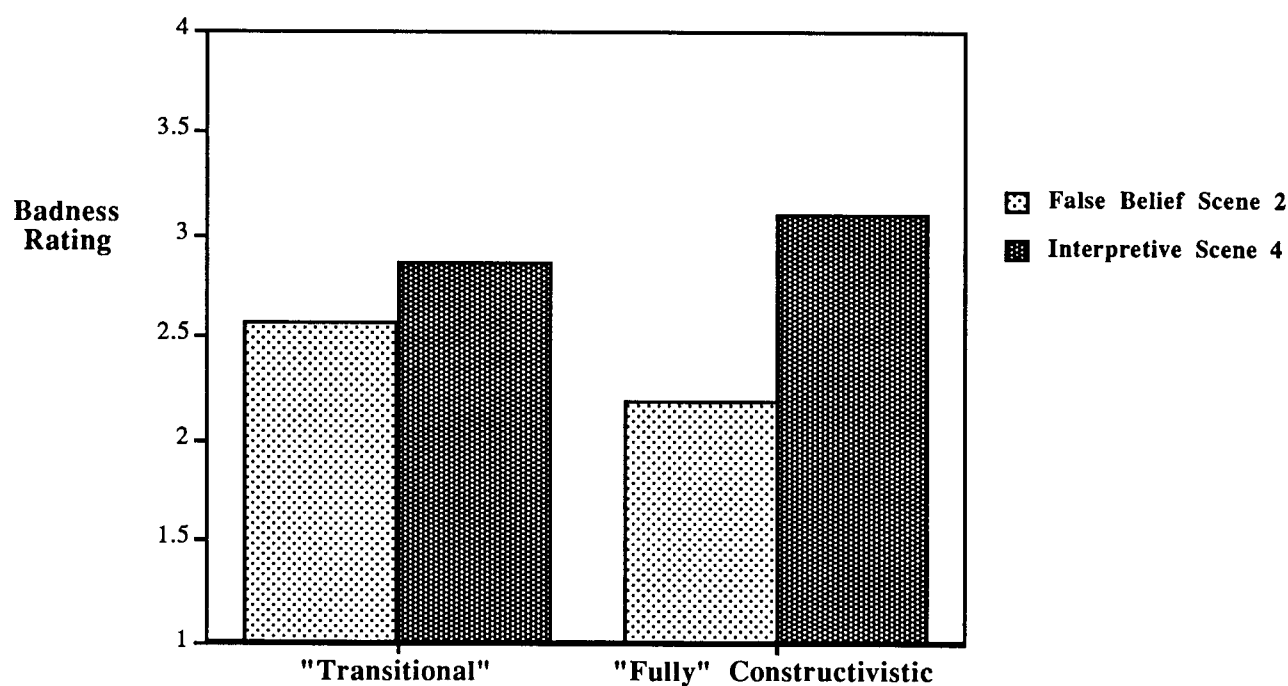


Figure 3: Manslaughter Ratings for Children coded as either "Transitional" or "Fully" Constructivistic

thus showed a difference [ $t(9)=5.01, p<.0004$ ] of almost one full point on the scale. Transitional children, although otherwise indistinguishable from their more fully constructivistic counterparts, saw only a trivial difference (i.e., .29) which did not reach levels of significance,  $t(13)=1.0, p=.17$ . This difference between these two groups of children was also reflected by their views about Punch's specific contribution to the interpretive mix-up in Scene 4. That is, fully constructivistic participants raised the issue of Punch's negligence about 80% of the time, while this issue came out in discussion just 17% of the time among transitional children (Fisher's Exact Test,  $p=.0009$ ).

### General Discussion

The take home message from these various results comes in two parts. First, it is apparent that school-aged children who already have come to an interpretive, or constructivistic, view of the knowing process are alert to a dimension of the moral order that is lost on their less mature counterparts, who, in not having rounded this "interpretive turn" continue to view simple ignorance as the only intellectual crime. In particular, children subscribing to a constructivistic theory of mental life, in contrast to their non-constructivistic peers, evidence a more thoroughgoing understanding of human agency that makes place not only for desires, but also includes an understanding of how people make an active, or interpretive, contribution to what they think and know. The practical upshot of this for children's moral deliberations, as I've worked to demonstrate, is that such young "constructivistic," as opposed to plain "copy," theorists come to see the possibility of a new kind of epistemic negligence. That is, in light of this new posture toward mental life, they come to hold themselves and others as agents who are not only morally accountable for the objective consequences of their desires, but also for what

they subjectively interpret to be the case. With this insight in place, such young persons are in the epistemic position to join the adult world in distinguishing between so-called mistakes of fact and mistakes of law.

The final point, which has served as the guiding theme throughout this study, concerns how the fact-value division prescribed by some is not inviolable, particularly at the level of practical moral judgments. Specifically, the empirical efforts here demonstrate in fairly clear ways how beliefs about belief and beliefs about morality are actually deeply interpenetrating matters. Given that this is the case, then perhaps such issues of truth and rightness shouldn't be allowed to stand on separate pedestals, but instead, should be brought back down to their humble beginnings in the naïve thoughts of children.

## Bibliography

- Anderson, N. H. (1980). Information integration theory in developmental psychology. In F. Wilkening, J. Becker, & T. Trabasso (Eds.), Information integration by children (pp. 1-45). Hillsdale, NJ: Lawrence Erlbaum.
- Aristotle. (1985). Nicomachean ethics. (T. Irwin, Trans.). Indianapolis, IN: Hackett Publishing.
- Astington, J. W. (1986). Children's comprehension of expressions of intention. British Journal of Developmental Psychology, 4, 43-49.
- Astington, J. W. (1991). Intention in the child's theory of mind. In D. Frye & C. Moore (Eds.), Children's theories of mind: Mental states and social understanding (pp. 157-172). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Baird, B. (1973). The art of the puppet. New York: Bonanza Books.
- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., & Irwin, J. M. (1996). Infants' reliance on social criterion for establishing word-object relations. Child Development, 67, 3135-3153.
- Black, H. C. (1990). Black's law dictionary (6<sup>th</sup> ed.). St. Paul, MN: West Publishing Co.
- Blasi, A. (1991). The self as subject in the study of personality. In D. Ozer, J. M. Healy, Jr., A. J. Stewart (Vol. Eds.) and R. Hogan (Series Ed.), Perspectives in personality: Self and emotion (Vol. 3, Part A, pp. 19-37). London: Jessica Kingsley Publishers.
- Blasi, A. (1995). Moral understanding and the moral personality: The process of moral integration. In W. M. Kurtines & J. L. Gewirtz (Eds.), Moral development: An introduction (pp. 229-253). Boston, MA: Allyn and Bacon.
- Bratman, M. (1981). Intention and means-end reasoning. The Philosophical Review, 90, 252-265.
- Bratman, M. (1984) Two faces of intention. The Philosophical Review, 93, 375-405.
- Buchanan, J. P., & Thompson, S. K. (1973). A quantitative methodology to examine the development of moral judgment. Child Development, 44, 186-189.
- Carpendale, J. I. M. (1996). On the distinction between false belief understanding and the acquisition of an interpretive theory of mind. Unpublished doctoral dissertation, The University of British Columbia, Vancouver, Canada.



- Carpendale, J. I. & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. Child Development, 67, 1686-1706.
- Chandler, M. J. (1988). Doubt and developing theories of mind. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.) Developing theories of mind (pp. 387-413). New York: Cambridge University Press.
- Chandler, M. J., & Boyes, M. (1982). Social-cognitive development. In B. B. Wolman (Ed.), Handbook of developmental psychology (pp. 387-402). Englewood Cliffs, NJ: Prentice-Hall.
- Chandler, M. J., & Carpendale, J. I. M. (1998). Inching toward a mature theory of mind. In M. Ferrari & R. Sternberg (Eds.), Self-awareness: Its nature and development (pp. 148-190). New York: Guilford Publications.
- Chandler, M. J., & Helm, D. (1984). Developmental changes in the contribution of shared experience to social role taking competence. International Journal of Behavioral Development, 1, 145-156.
- Chandler, M. J., & Lalonde, C. (1996). Shifting to an interpretive theory of mind: 5- to 7-year olds' changing conceptions of mental life. In A. J. Sameroff & M. M. Haith (Eds.), The five to seven year shift: The age of reason and responsibility (pp. 111-139). Chicago: The University of Chicago Press.
- Chandler, M. J., & Sokol, B. W. (in press). Representation once removed: Children's developing conceptions of representational life. In I. E. Sigel (Ed.), Theoretical perspectives in the development of representational (symbolic) thought. Mahway, NJ: Lawrence Erlbaum Associates.
- Chandler, M. J., Greenspan, S., & Barenboim, C. (1973). Judgments of intentionality in response to videotaped and verbally presented moral dilemmas: The medium is the message. Child Development, 44, 315-320.
- Costanzo, P. R., Coie, J. D., Grumet, J. F., & Farnill, D. (1973). A reexamination of the effects of intent and consequence on children's moral judgments. Child Development, 44, 154-171.
- Damon, S. F. (1957). Punch and Judy. Barre, MA: Barre Publishing Company, Inc.
- Dean, A. L., & Youniss, J. (1991). The transformation of Piagetian theory by American psychology: The early competence issue. In M. Chandler & M. Chapman (Eds.), Criteria

- for competence: Controversies in the conceptualization and assessment of children's abilities (pp. 93-109). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dennett, D. C. (1978). Beliefs about beliefs. The Behavioral and Brain Sciences, 4, 568-570.
- Dennett, D. C. (February 1997). Making tools for thinking. Paper presented at the Tenth Annual Vancouver Cognitive Science Conference, Vancouver, British Columbia, Canada.
- Duff, R. A. (1990). Intention, agency, and criminal liability: Philosophy of action and the criminal law. Oxford: Basil Blackwell.
- Dunn, J. (1991). Young children's understanding of other people: Evidence from observations within the family. Children's theories of mind: Mental states and social understanding (pp. 97-114). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Farnill, D. (1974). The effects of social-judgment set on children's use of intent information. Journal of Personality, 42, 276-289.
- Fincham, F. D., & Roberts, C. (1985). Intervening causation and the mitigation of responsibility for harm doing. Journal of Experimental Social Psychology, 21, 178-194.
- Frye, D. (1991). The origins of intention in infancy. In D. Frye & C. Moore (Eds.), Children's theories of mind: Mental states and social understanding (pp. 15-38). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Grueneich, R. (1982). Issues in the developmental study of how children use intention and consequence information to make moral evaluations. Child Development, 53, 29-43.
- Gutkin, D. C. (1972). The effect of systematic story changes on intentionality in children's moral judgments. Child Development, 43, 187-195.
- Harman, G. (1976). Practical reasoning. Review of Metaphysics, 79, 431-463.
- Harré, R. (1982). Personal being: A theory of individual psychology. Oxford: Basil Blackwell.
- Hart, H. L. A. (1958). Legal responsibility and excuses. In S. Hook (Ed.), Determinism and freedom in the age of modern science (pp. 81-104). New York: New York University Press.
- Hebble, P. W. (1971). The development of elementary school children's judgment of intent. Child Development, 42, 1203-1215.
- Heider, F. (1958). The psychology of interpersonal relations. New York: John Wiley & Sons.

- Helwig, C. C., Hildebrandt, C., & Turiel, E. (1995). Children's judgments about psychological harm in a social context. Child Development, 66, 1680-1693.
- Hume, D. (1978). A treatise of human nature. (2<sup>nd</sup> ed.) (L. A. Selby-Bigge & P. H. Nidditch, Eds.). Oxford: Oxford University Press. (Original work published in 1739)
- Kadish, S. H., & Schulhofer, S. J. (1995). Criminal law and its processes: Cases and materials (6<sup>th</sup> ed.). New York: Little, Brown and Company.
- Kant, I. (1959). Foundations of the metaphysics of morals (L. W. Beck, Trans.). New York: Macmillan. (Original work published 1785)
- Karniol, R. (1978). Children's use of intention cues in evaluating behavior. Psychological Bulletin, 85, 76-85.
- Keasey, C. B. (1977). Children's developing awareness and usage of intentionality and motives. In C. B. Keasey (Ed.), Nebraska Symposium on Motivation (Vol. 25, pp. 219-260). Lincoln, NE: University of Nebraska Press.
- Kohlberg, L. (1969, March). The relations between moral judgment and action: A developmental view. Paper presented at the Institute of Human Development, University of California, Berkeley, CA.
- Lalonde, C. E. (1996). Children's understanding of the interpretive nature of the mind. Unpublished doctoral dissertation, The University of British Columbia, Vancouver, Canada.
- Lane, J., & Anderson, N.H. (1976). Integration of intention and outcome in moral judgment. Memory & Cognition, 4, 1-5.
- Lapsley, D. K. (1996). Moral psychology. Boulder, CO: Westview Press.
- Leon, M. (1980). Integration of intent and consequence information in children's moral judgments. In F. Wilkening, J. Becker, & T. Trabasso (Eds.), Information integration by children (pp. 71-97). Hillsdale, NJ: Lawrence Erlbaum.
- Lewis, C., & Mitchell, P. (1994). Children's early understanding of mind: Origins and development. Hove, UK: Lawrence Erlbaum Associates.
- Meltzoff, A. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. Developmental Psychology, 31, 838-850.

- Montgomery, D. E. (June 1998). A Wittgensteinian perspective on children's acquisition of mental terms: The case of intentions. In D. E. Montgomery (Chair), Language and theory of mind. Paper symposium conducted at the 28<sup>th</sup> Annual Symposium of the Jean Piaget Society, Chicago, IL.
- Moses, L. J. (1993). Young children's understanding of belief constraints on intention. Cognitive Development, 8, 1-25.
- Moses, L. J., & Chandler, M. J. (1992). Traveler's guide to children's theories of mind. Psychological Inquiry, 1, 285-301.
- Nelson, S. A. (1980). Factors influencing young children's use of motives and outcomes as moral criteria. Child Development, 51, 823-829.
- Nucci, L. (1981). Conceptions of personal issues: A domain distinct from moral and societal concepts. Child Development, 52, 114-121.
- Nucci, L. P., & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. Child Development, 49, 400-407.
- Perner, J. (1988). On representing that: The asymmetry between belief and desire in children's theory of mind. In D. Frye & C. Moore (Eds.), Children's theories of mind: Mental states and social understanding (pp. 139-155). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Perner, J., & Davies, G. (1991). Understanding the mind as an active information processor: Do young children have a "copy theory of mind"? Cognition, 39, 51-69.
- Piaget, J. (1965). The moral judgment of the child (M. Gabain, Trans.). New York: The Free Press. (Original work published 1932)
- Pillow, B. H. (1988). The development of children's beliefs about the mental world. Merrill-Palmer Quarterly, 34, 1-32.
- Pillow, B. H. (1995). Two trends in the development of conceptual perspective taking: An elaboration of the passive-active hypothesis. International Journal of Behavioral Development, 18, 649-676.
- Premack, D. (1990). The infant's theory of self-propelled objects. Cognition, 36, 1-16.
- Ross, S. D. (1973). The nature of moral responsibility. Detroit, MI: Wayne State University.
- Ruffman, T. K., Olson, D. R., & Astington, J. W. (1991). Children's understanding of visual ambiguity. British Journal of Development Psychology, 9, 89-102.

- Searle, J. R. (1983). Intentionality: An essay in the philosophy of mind. Cambridge: Cambridge University Press.
- Shultz, T. R. (1991). From agency to intention: A rule-based, computational approach. In A. Whiten (Ed.), Natural theories of mind: Evolution, development, and simulation in everyday mindreading (pp. 79-83). Oxford: Blackwell.
- Shultz, T. R., Wright, K., & Schleifer, M. (1986). Assignment of moral responsibility and punishment. Child Development, 57, 177-184.
- Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. Child Development, 52, 1333-1336.
- Sodian, B., & Wimmer, H. (1987). Children's understanding of inferences as a source of knowledge. Child Development, 58, 424-433.
- Sokol, B. W., & Chandler, M. J. (June 1998). Beliefs about truth and rightness: Implications of developing a theory of mind for children's evaluations of right and wrong. Paper presented at the 28th annual Symposium of the Jean Piaget Society, Chicago, IL.
- Spelke, E. S., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), Causal cognition: A multidisciplinary debate (pp. 44-78). Oxford: Clarendon Press.
- Strawson, P. (1982). Freedom and resentment. In G. Watson (Ed.), Free will (pp. 59-80). Oxford: Oxford University Press. (Reprinted from Proceedings of the British Academy, 48, pp. 1-25, 1962)
- Surber, C. F. (1977). Developmental processes in social inference: Averaging of intentions and consequences in moral judgment. Developmental Psychology, 13, 654-665.
- Taylor, R. (1963). Metaphysics. Englewood Cliffs, NJ: Prentice-Hall.
- Taylor, R. (1966). Action and purpose. Englewood Cliffs, NJ: Prentice-Hall.
- Tisak, M. S., & Turiel, E. (1988). Variation in seriousness of transgressions and children's moral and social conventional concepts. Developmental Psychology, 24, 352-357.
- Turiel, E. (1990). Moral judgment, action, and development. In W. Damon (Series Ed.) & D. Schrader (Vol. Ed.), New directions for child development: Vol. 47. The legacy of Lawrence Kohlberg (pp. 31-49). San Francisco: Jossey-Bass.

- Wainryb, C., & Ford, S. (in press). Young children's evaluations of acts based on beliefs different from their own. Merrill-Palmer Quarterly.
- Wellman, H. M. (1990). The child's theory of mind. Cambridge, MA: MIT Press.
- Wellman, H. M., & Hickling, A. K. (1994). The mind's "I": Children's conceptions of the mind as an active agent. Child Development, 65, 1564-1580.
- White, P. A. (1995). The understanding of causation and the production of action: From infancy to adulthood. Hove, UK: Lawrence Erlbaum Associates.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition, 13, 103-128.
- Wolf, D. (1982). Understanding others: A longitudinal case study of the concept of independent agency. In G. E. Forman (Ed.), Action and thought: From sensorimotor schemes to symbolic operations (pp. 297-328). New York, NY: Academic Press.
- Wolf, D. P., Rygh, J., & Altshuler, J. (1984). Agency and experience: Actions and states in play narratives. In I. Bretherton (Ed.), Symbolic play: The development of social understanding (pp. 195-217). Orlando, FL: Academic Press.
- Wren, T. E. (1974). Agency and urgency: The origin of moral obligation. New York: Precedent Publishing.
- Youniss, J., & Damon, W. (1992). Social construction in Piaget's theory. In H. Beilin & P.B. Pufall (Eds.), Piaget's theory: Prospects and possibilities (pp. 267-286). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. Child Development, 67, 2478-2492.

## Appendix 1

**Adult Interview Schedule****Preliminary Questions.** *(Baseline Rating)*

On a scale from 1 to 5, where 1 is only a "little bit bad" and 5 is "really, really bad," please rate the degree of badness for the following actions: (Please circle your answer)

Hitting another person

1      2      3      4      5  
only a little bit bad      really, really bad

Putting library books back in the wrong place

1      2      3      4      5  
only a little bit bad      really, really bad

Stealing a valued belonging from another person

1      2      3      4      5  
only a little bit bad      really, really bad

Talking when it's not your turn

1      2      3      4      5  
only a little bit bad      really, really bad

Closing somebody up in a box and throwing them in the garbage  
(i.e., intentionally bringing serious personal injury to another)

1      2      3      4      5  
only a little bit bad      really, really bad

**Scene 1: Judy Escapes.** *(Attempted Murder)*

Where does Punch think Judy is when he throws the box out into the garbage? (Please circle one)  
In the Box or Someplace else *(False Belief Question)*

Using as a guide your rating of throwing somebody into the garbage (see above), how bad do you think it is now for Punch to throw the box with Judy in it out into the garbage?

1      2      3      4      5  
only a little bit bad      really, really bad

Briefly explain your reasons for rating Punch's action as you did.

**Scene 2: Throwing out Boxes.**      *(Involuntary Manslaughter)*

Using as a guide your rating from Scene 1, how bad do you think it is now for Punch to throw out the box?

1      2      3      4      5  
only a little bit bad      really, really bad

**Briefly explain your reasons for rating Punch's action as you did.**

**Scene 3: "I'm in the Orange Box".**     *(Attempted Murder)*

Why does Punch end up choosing the wrong box after Judy has told him which box she is really in? (*Explanation Question*)

Is Judy's reply to Punch about which box she has fallen into wrong? Why or Why not.  
(Evaluation Question)

Who is at fault for the mix up here about which box Judy has fallen into? (Please circle one)  
Punch , Judy , Both , Neither

**Briefly explain your reasons for selecting this answer.** *(Determination Question)*

Using as a guide your first rating of throwing somebody into the garbage (see above), how bad do you think it is now for Punch to throw out the orange-colored box?

1 2 3 4 5  
only a little bit bad really, really bad

**Briefly explain your reasons for rating Punch's action as you did.**

**Scene 4: “Check the Green One”.**      *(Involuntary Manslaughter)*

Why does Punch end up choosing the wrong box after Judy has told him which box she is really in? (*Explanation Question*)

Is Judy's reply to Punch about which box she has fallen into wrong? Why or Why not.  
(Evaluation Question)

Who is at fault for the mix up here about which box Judy has fallen into? (Please circle one)  
Punch , Judy , Both , Neither

**Briefly explain your reasons for selecting this answer.** *(Determination Question)*



Using as a guide your rating from Scene 3, how bad do you think it is now for Punch to throw the green box with Judy in it out into the garbage?

1      2      3      4      5  
only a little bit bad      really, really bad

Briefly explain your reasons for rating Punch's action as you did.

## Appendix 2

### **Child Interview Schedule**

#### **Preliminary Questions.** *(Baseline Rating)*

Show me how bad it is to hit another person. (1 2 3 4 5)

Is putting your toys away in the wrong place more bad or less bad ?

Show me how bad it is. (1 2 3 4 5)

Show me how bad it is to steal another person's favorite toy. (1 2 3 4 5)

Is talking when it isn't your turn more bad or less bad ?

Show me how bad it is. (1 2 3 4 5)

Show me how bad it is to go out of turn when you're playing a game. (1 2 3 4 5)

Is lying to your teacher more bad or less bad ?

Show me how bad it is. (1 2 3 4 5)

Finally, show me how bad it is to put another person in a box and throw them in the garbage.  
(1 2 3 4 5)

#### **Scene 1: Judy Escapes.** *(Attempted Murder)*

##### **Memory Probes**

Before Punch left to get the rope, where was Judy? (Probe: Before Punch left to get the rope, was Judy in the box or someplace else?)

When Punch comes back in with the rope, where is Judy? (Probe: When Punch comes back with the rope, is Judy in the box or someplace else where Punch doesn't see her?)

##### **False Belief Question**

Where does Punch think Judy is when he throws the box out into the garbage? (Probe: Does Punch think Judy is in the box or someplace else?)

##### **Comprehension Questions**

Was Judy in the box that Punch threw out? (Yes / No)

Does that mean that Judy could have been hurt or not hurt when Punch threw out the box?

Was Punch trying to hurt Judy? (Yes / No)

##### **Badness Rating**

Remember how bad you first said it was to throw someone out into the garbage. (adjust meter to appropriate level -- see the badness meter practice phase) In the story now, is it the same, more bad, or less bad for Punch to throw out the box?

Using the badness meter, show me where you would put the pointer for Punch? (1 2 3 4 5)  
Why is it (the same, more, less bad)?

**Scene 2: Throwing out Boxes.** (*Involuntary Manslaughter*)

**Memory Probe**

Where is Judy when Punch throws the box out into the garbage? (Probe: When Punch throws out the box, is Judy in the box or is she someplace else?)

**Comprehension Questions**

Since Judy was in the box, does that mean that she could have been hurt or not hurt when Punch threw it into the garbage?

Did Punch know that Judy was in the box that he threw out? (Yes / No)

Does that make what Punch did an accident or did he throw her out on purpose?

**Badness Rating**

Remember last time when Punch tried to throw the box with Judy in it out into the garbage? You said it was this bad. (adjust meter to appropriate level) In the story now, is it the same, more bad, or less bad for Punch to throw Judy into the garbage?

Using the badness meter, show me where you would put the pointer for Punch? (1 2 3 4 5)  
Why is it (the same, more bad, less bad)?

**Scene 3: "I'm in the Orange Box".** (*Attempted Murder*)

**Memory Probes**

What box did Judy fall into? (Probe: Did Judy fall into the orange-colored box or the brown box with oranges in it?)

What did Judy say when Punch asked what box she fell into? (Judy said: "I'm in the orange box.")

**Control Questions**

When Judy says "I'm in the orange box," what box is she thinking of? (Probe: When Judy says "I'm in the orange box," is she talking about the orange-colored box or the brown box with oranges in it?)

After hearing Judy say "I'm in the orange box," what box does Punch think she fell into? (Probe: Does Punch think Judy is in the orange-colored box or does he think she is in the brown box with oranges in it?)

### **Explanation Questions**

Why is Punch thinking that Judy is in this box (point to orange-colored box) after Judy has told him what box she is really in?

What are Punch's reasons for choosing this other box (point to orange-colored box)? (Probe: Does Punch have a good reason for choosing this box or is he being silly?) (if "silly" go to Determination Question)

### **Evaluation Questions**

When Judy says "I'm in the orange box" she is talking about this box (point to brown box), but Punch thinks she is in this other one. Is saying "I'm in the orange box" really wrong or does it sort of make sense?

If "really wrong"—Why does Judy say it? (Probe: Is it wrong for Judy to say "I'm in the orange box" because she is making a mistake, because she is being silly, or because she is trying to trick Punch?) (Go to Determination Question)

If "makes sense"—Is saying "I'm in the orange box" the only way for Judy to say where she is at or could she have said something different to make sure that Punch would pick the box she fell into? What could she have said?

### **Determination Question**

So Punch picked this other box (point to orange-colored box) even though Judy really fell into the brown box with oranges in it. Was this mix-up about the boxes all Judy's fault, all Punch's fault, or a little bit of both?

### **Comprehension Questions**

Was Judy in the box that Punch threw out? (Yes / No)

Does that mean that Judy could have been hurt or not hurt when Punch threw out the box?

Was Punch trying to hurt Judy? (Yes / No)

### **Badness Rating**

Remember how bad you first said it was to throw someone out into the garbage. (adjust meter to appropriate level -- see the badness meter practice phase) In the story now, is it the same, more bad, or less bad for Punch to throw the orange-colored box out into the garbage?

Using the badness meter, show me where you would put the pointer for Punch? ( 1 2 3 4 5 )  
Why is it (the same, more, less bad)?

**Scene 4: "Check the Green One".** (*Involuntary Manslaughter*)

**Memory Probes**

What box did Judy fall into? (Probe: Did Judy fall into the green box or the white box with the green "1" on it?)

What did Judy say when Punch asked what box she fell into? (Judy said: "Check the green one.")

**Control Questions**

When Judy says "check the green one," what box is she thinking of? (Probe: When Judy says "check the green one," is she talking about the green box or the white box with the green "1" on it?)

After hearing Judy say "Check the green one," what box does Punch think she fell into? (Probe: Does Punch think Judy is in the green box or does Punch think she is in the white box with the green "1" on it?)

**Explanation Questions**

Why is Punch thinking that Judy is in this box (point to white box) after Judy has told him what box she is really in?

What are Punch's reasons for choosing this other box (point to white box)? (Probe: Does Punch have a good reason for choosing this box or is he being silly?) (if "silly" go to Determination Question)

**Evaluation Questions**

When Judy says "check the green one" she is talking about this box (point to green box), but Punch thinks she is in this other one. Is saying "check the green one" really wrong or does it sort of make sense?

If "really wrong"—Why does Judy say it? (Probe: Is it wrong for Judy to say "check the green one" because she is making a mistake, because she is being silly, or because she is trying to trick Punch?) (Go to Determination Question)

If "makes sense"—Is saying "check the green one" the only way for Judy to say where she is at or could she have said something different to make sure that Punch would pick the box she fell into? What could she have said?

**Determination Question**

So Punch picked this other box (point to white box) even though Judy really fell into the green box. Was this mix-up about the boxes all Judy's fault, all Punch's fault, or a little bit of both?

**Comprehension Questions**

Since Judy was in the box, does that mean that she could have been hurt or not hurt when Punch threw it into the garbage?

Did Punch know that Judy was in the box that he threw out? (Yes / No)

Does that make what Punch did an accident or did he throw her out on purpose?

**Badness Rating**

Remember last time when Punch tried to throw the box with Judy in it out into the garbage? You said it was this bad. (adjust meter to appropriate level) In the story now, is it the same, more bad, or less bad for Punch to throw the green box with Judy in it out into the garbage?

Using the badness meter, show me where you would put the pointer for Punch? (1 2 3 4 5)  
Why is it (the same, more, less bad)?