

CHILDREN'S CONCEPTIONS OF AGENCY AND MORALITY: MAKING SENSE OF THE
HAPPY VICTIMIZER PHENOMENON

by

BRYAN W. SOKOL

B.A. (Honors), Saint Louis University, 1991
M.A., The University of British Columbia, 1998

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

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Psychology

THE UNIVERSITY OF BRITISH COLUMBIA

December 2004

© Bryan W. Sokol, 2004

Abstract

This thesis explores a puzzling aspect of young children's reasoning about the emotional consequences associated with harming others – or what, in the moral development literature, has become known as the “happy victimizer phenomenon.” Past research on this so-called “phenomenon” has consistently found a strong tendency among most 5-, but not 7-, year-olds to imagine that those who get what they want by victimizing others will experience only positive, or “happy,” emotions. Despite an otherwise good showing on alternative measures of moral understanding, such young children regularly appear to leave sadness, guilt, or remorse out of the picture. The three studies reported here each provide converging evidence that this apparent “moral transition” in the early school years is, in fact, one facet of a broader re-structuring in children's evolving conceptions of human agency. Building on recent initiatives in the theories-of-mind literature, Study One explored the relations between children's understanding of the interpretive nature of the knowing process (i.e., an “interpretive theory of mind”) and performance on both traditional measures of the happy victimizer phenomenon and alternative procedures designed to draw out the agentic dimensions of the standard assessment conditions. While a strong association between children's theory-of-mind scores and their level of emotion understanding was found in the traditional testing conditions, the alternative procedures demonstrated that children's emotion attributions were heavily influenced by how closely they attended to victimizers' actions. Study Two further examined these findings by more directly testing the relation between an interpretive understanding and children's reasoning about the motivational states that typically underlie others' actions. Using procedures designed to assess children's understanding of “deviant causal chains,” Study Two found that an interpretive theory of mind placed important constraints on children's views about the differences between desires

and intentions. Finally, Study Three worked to rule out potential reductionistic explanations of the present “agency account” by showing that children’s success at reasoning simultaneously about two aspects of a stimulus event – a design feature shared by all the measures used in these studies – was a distinct, and far simpler ability, than the later-arriving agentive notions assessed in Studies One and Two.

Table of Contents

Abstract.....	ii
List of Figures and Tables	vi
Acknowledgements	vii
CHAPTER 1: Introduction	
An Overview of Three Studies	1
CHAPTER 2: Conceptual Foundations	
The Happy Victimizer Phenomenon	5
Critical Literature Review	8
Studies on Mixed Emotions: Another Mixed Bag	18
Summary and Assessment: Where Do We Go From Here?.....	23
Harris and Bandura: Agency or Apparition?	26
Piaget and “The Will”	30
Interpretation and Human Agency	35
CHAPTER 3: Study One	
Exploring the Cognitive Constraints Surrounding the Happy Victimizer Phenomenon	40
Design Modifications	41
Method	43
Results	52
Conclusions	63

CHAPTER 4: Study Two

Exploring the Building Blocks of Human Agency:

Children's Conceptions of Intention and Desire	70
The Inherent Shortcomings of Belief-Desire Psychology	72
The New Frontier of Intentions	76
Method	79
Results	81
Conclusions	85

CHAPTER 5: Study Three

Agency or Otherwise? Exploring Alternative Explanations of an

Interpretive Theory of Mind	87
Method	90
Results	91
Conclusions.....	93

CHAPTER 6: General Discussion

Limitations and Future Directions	94
REFERENCES	99
APPENDIX A	116
APPENDIX B.....	119
APPENDIX C.....	122
APPENDIX D	125

List of Figures and Tables

Figure 1: A ship arriving too late to save a drowning witch.....	45
Figure 2: Mixed emotion attributions grouped by theory-of-mind score	56
Figure 3: Primary emotion attributions in the victimizer protocol.....	61
Figure 4: Primary emotion attributions in the non-victimzier protocol.....	62
Figure 5: Participants' success in the separate conditions of the <i>IDM</i> protocol.....	84
Table 1: Children's seriation performance	92
Table 2: Children's theory-of-mind performance	93

Acknowledgements

Portions of this research were supported by the Natural Sciences and Engineering Research Council of Canada, the Michael Smith Foundation for Health Research, and the Human Early Learning Partnership.

The amount of thanks that I owe to my colleagues, friends, and family who have both helped sustain me throughout the production of this thesis and contributed in various ways to its completion is simply immeasurable. In particular, none of this work would have been possible without the support and input from the following people: Melissa and Noah, Michael Chandler, the children and staff at the Delta Christian School, George Alder, Jed Allen, Tracey Burns, Jeremy Carpendale, Laura Failows, Jessica Flores, Erica Gehrke, Sheena Grant, Darcy Hallett, Gabriela Ionita, Chris Jones, Leigh Koopman, Tobias Krettenauer, Chris Lalonde, Stuart McKenzie, Michelle Morgan, David Paul, Colleen Rostek, Ulrich Teucher, Cynthia VanGinkel, and Leah Yin. Finally, I wish to thank the members of my examination committee for their thoughtful feedback and constructive criticism of my research: Anita DeLongis, Geoff Hall, Ara Norenzayan, Kiran Soma, Larry Walker, Kim Schonert-Reichl, and Charles Helwig. My deepest thanks to all of you!

I. INTRODUCTION

An Overview of Three Studies

The research to be detailed in this thesis is driven by two broad considerations. One of these – a perennial problem for developmentalists – concerns the difficulties routinely encountered in making both theoretical and empirical sense of the observation that, for just about everyone, but especially for young children, our best thoughts about morally hazardous matters frequently fail to spill over into morally appropriate behaviors and feelings. There are, of course, many sorts of “slips twixt cup and lip,” but few as puzzling as the fact that young children’s morally relevant thoughts, emotions, and actions so often appear to live in separate watertight compartments. What I will go on to describe under the rubric of the “happy victimizer phenomenon” is a live instance of an important corner of this age-old problem.

The second broad issue upon which this thesis is meant to touch concerns the place of “agency” (earlier writers would have said “will”) in the formation of young people’s developing moral lives. For generations, and throughout most of the 20th century, psychologists have taken a special (perhaps even perverse) pride in developing accounts of human affairs from which all talk of agency, or will, was carefully expunged. That was then. Of late, there has, at least in developmental circles, been a considerable resurgence of interest (e.g., Martin, Sugarman, & Thompson, 2003) in writing the notion of agency back into the equation for moral maturation. For all of the reasons that originally led to the banishment of will and other such potential “ghosts-in-the-machine” (Koestler, 1967) from the psychological scene, re-introducing the notion of agency into serious developmental work is not to be undertaken lightly, but soberly and discreetly. What follows, then – at least discreetly – is an account that leads from the two broad considerations just outlined to a particular set of more narrowly focused studies designed to

explore these overlapping matters of agency and morality in the developing mental lives of young children.

By way of a quick overview, the first study (see Chapter III) in this set of three is aimed at providing a fuller explanatory account of the so-called “happy victimizer phenomenon.” In general, this phenomenon or finding highlights an unusual disjuncture in young people’s moral conceptions – one in which early school-aged children, who have otherwise been shown to understand that acts of victimization are wrong, nevertheless attribute positive, or “happy,” emotional responses to those who intentionally bring harm upon others. Until quite recently, the traditional experimental procedures used in this area of research have turned largely on the straightforwardly descriptive task of documenting the age-graded differences between older and younger children’s reasoning about harmful actions and the subsequent experience of moral or immoral emotions. The observation that typically characterizes much of this work is that younger children, in contrast to their older counterparts, respond to questions about a victimizer’s misdeeds by focusing almost exclusively on the positive sentiments associated with the immediate, material rewards such wrongdoings produce, while generally overlooking the broader moral ramifications of such actions. Although there are occasional hints about the kinds of cognitive factors that might underlie such “outcome oriented” responses, very few studies have actively sought to identify the causal mechanisms, or reasoning processes, that potentially shape children’s inferences about a victimizer’s transgressions. Accordingly, one of the primary goals of this thesis has been to address this shortcoming in the happy victimizer research literature. In brief, the tentative conclusions that I draw from Study One are that, accompanying important changes in children’s theories of mind, are other attendant advances in their growing

understanding of human agency – changes that ultimately lead to the emergence of new ways of organizing or structuring key aspects of others’ morally relevant actions.

Study Two (see Chapter IV) was designed to provide further support for this conclusion by investigating, in a more head-on fashion, two defining aspects of action: an agent’s “desires” and his or her “intentions.” Although sometimes used interchangeably, philosophers (e.g., Bratman, 1987; Chisholm, 1976; Searle, 1983) have long argued that both of these motivational states hold very different criteria of satisfaction. In particular, a desire can be fulfilled or satisfied in any number of ways, so long as the individual in question eventually gets what he or she wants – if the “end” is achieved, the “means” by which it is accomplished are of secondary importance. Intentions, by contrast, have more narrow satisfaction criteria, and can only be fulfilled if the means the agent had in mind are precisely met. That is, one must not only get what one wants, but also achieve these ends in the way they were planned or intended. Following this line of reasoning, individuals can be seen to satisfy their desires without also fulfilling their intentions, just as they may follow through on their intentions without actually meeting their desires. Using a series of so-called “deviant causal chains” to illustrate these alternative possibilities, Study Two examines the abilities of young children to drive a conceptual wedge between intentions and desires in relation to their evolving theories of mind, as well as other important cognitive markers.

Finally, and as a way to begin ruling out potential reductionistic readings of this first pair of studies, Study Three (see Chapter V) introduces an important control procedure into the measurement strategies used to gauge children’s theories of mind. In particular, it could be argued that the underlying competence tapped by the theory-of-mind measures used in the first two studies has nothing to do with understanding agency per se, but is instead reflective of

children's more generic ability to "hold two things in mind at once." I investigate this alternative explanation in the last study by drawing on Inhelder and Piaget's (1969) classic research dealing with children's seriation abilities.

II. CONCEPTUAL FOUNDATIONS

The Happy Victimizer Phenomenon

The cognitive-developmental approach that currently dominates the literature on children's moral development has long been plagued by questions about how to deal with matters involving affective or emotional content (e.g., Eisenberg, 1986; Eisenberg, Shea, Carlo, & Knight, 1991; Hoffman, 1984, 2000). One of the more recent eruptions of this latent problem has surfaced in empirical work on the so-called "happy victimizer phenomenon." A variety of studies (e.g., Arsenio & Kramer, 1992; Barden, Zelko, Duncan, & Masters, 1980; Keller, Lourenço, Malti, & Saalbach, 2003; Lourenço, 1997; Murgatroyd & Robinson, 1993; Nunner-Winkler & Sodian, 1988; Yuill, Perner, Pearson, Peerbhoy, & van den Ende, 1996) exploring this phenomenon all provide converging evidence that, before the ages of 6 or 7, children who otherwise appear to have a rich understanding of right and wrong nevertheless consistently fail to show an appreciation of the broader emotional consequences of various immoral activities. More specifically, they uniformly attribute positive (or "happy") emotional reactions to individuals who achieve their goals, despite having victimized or harmed others in the process.

The prospect that young children (or anyone, for that matter) might make such cold-blooded attributions is not, of course, wholly unanticipated. Viewed by the lights of certain classical "associationist" theories, for example, the prospects of "getting what one wants" and of "being happy that one did" can be seen to so regularly go hand-in-hand that proceeding as though something similar always occurs, regardless of possible way-side costs or collateral damage, would seem a likely default option, not only for young children, but for the rest of us as well. Similar expectations might also be seen to flow from the more contemporary theories-of-mind literature, whose contributors never seem to tire of demonstrating that children initially

structure their earliest understandings of others' mental and emotional lives around the basic notion of "desire" (e.g., Wellman, 1990; Wellman & Banerjee, 1991). On this account, children as young as 2 or 3 are judged to be capable of grasping that a satiated desire is also a reliable cause for happiness. Despite this confluence of views, all of which serve to make young children's evident callousness less of a surprise, the happy victimizer phenomenon continues to be sufficiently counter-intuitive that it has earned its place as a special problem to be solved by those working in the area of children's socio-moral development. The reason for this is because, as Arsenio and Lover (1995) have pointed out, the "happy victimizer conception seems to conflict with the core assumptions of several influential sociomoral theories...in which an awareness of victim harm, a sense of empathic distress, and a fear of external sanctions should all lead children to expect that victimizers will feel *negative* emotions, such as sadness, fear, and/or guilt" (p. 101). The trouble is, if the available research literature is to be believed, such young children do not. Instead, they expect victimizers to be unapologetically happy about their misdeeds.

Given these antagonistic intuitions, the scene is well set for debate. In one corner, there are all those from the associationist and theories-of-mind camps who, placing their bets on the likes of Wellman and Woolley's (1990) work, argue that the conceptual machinery of young children is driven by the most "basic aspect of desire-dependent emotions [or, simply that] ... wanting and getting leads to happiness" (cited in Wellman & Banerjee, 1991, p. 193). By such lights, children are seen as being bound – even in moral or value-laden situations where one might hope they would act otherwise – to follow what might be called a rudimentary "pleasure principle." In the opposite corner, are all those prepared to hold out for a more optimistic (some might say "guilt ridden") view, according to which even preschool children are imagined to be

capable of overriding such “baser” reactions, at least when real moral matters are at stake. This would seem to be the required stance of anyone, for instance, who has been impressed by the precocity of preschool children in the various studies conducted by Turiel and his colleagues (e.g., Turiel, 1983, 1998; Turiel, Hildebrandt, & Wainryb, 1991; Turiel, Killen, & Helwig, 1987; Turiel & Smetana, 1984) – studies showing that children as young as 3 years of age seem to have well-formed intuitions concerning the principles constituting the moral domain (i.e., matters involving human welfare or fairness). Assuming that all of this is true, the big question that naturally arises is, as Lourenço (1997) recently put it: “Why does the young child’s cognitive moral knowledge not lead to a corresponding emotional morality (i.e., to expect a wrongdoer to feel bad rather than good)?” (p. 426). Or, more generally, how ought we to account for the apparent disjunction between children’s seemingly rich knowledge about the defining characteristics of moral issues, on the one hand, and their relatively impoverished understanding of the emotional or affective consequences of these very same matters, on the other?

As it turns out, there are already a handful of fairly standard responses to this general question sprinkled throughout the various published studies on the happy victimizer phenomenon and the literature dealing with children’s early emotional development, all of which will be taken up in the following critical review. A central shortcoming in much of this work, as I alluded to in the overview of Chapter I, is its general failure to provide an adequate explanatory framework in which to investigate the causal mechanisms and potential cognitive constraints underlying children’s reasoning about acts of victimization. Accordingly, my own candidate solution to the tension between children’s so-called “*cognitive* moral knowledge” and their “*emotional* morality” differs from these more standard accounts by targeting the under-explored cognitive constraints surrounding the happy victimizer phenomenon. Moreover, it does this by taking

guidance from a frequently under-appreciated passage in Piaget's (1954/1981) work, *Intelligence and Affectivity* – a passage in which he addresses a similar problem concerning the relation between children's cognitive and affective growth. *There*, in this early account, Piaget argued that human agency – or “the will” as he called it – was “the affective analogue of intellectual decentration” (Piaget, 1954/1981, p. 64), and, as such, was at the center of understanding the overall relation between children's developing cognitive and emotional competence. *Here*, I mean to similarly argue that human agency, and particularly children's changing views of it, is central to making local sense of the smaller-scale tension surrounding the happy victimizer phenomenon. This argument, which forms the conceptual basis for Study One, flows largely from research dealing with children's early attempts to understand matters of interpretation, or what has been called the development of an “interpretive theory of mind” (Carpendale & Chandler, 1996, Chandler & Sokol, 1999, Lalonde & Chandler, 2002). By the end of this chapter, I plan to show, in particular, how this earlier work can be deployed in operationalizing human agency and the different forms or expressions it may take. Still, getting clear about why there are good reasons for moving from “interpretive this” to “happy victimizer that” – a conceptual maneuver that, at least at this early warm-up phase, might seem quite the stretch – requires a more substantive look at the research literature on children's understanding of moral emotions, as well as some of the other hypotheses already on offer for why young people are especially prone to making happy-victimizer emotion attributions.

Critical Literature Review

Early Research. Although Nunner-Winkler and Sodian (1988) are frequently credited with the first series of studies to directly explore the so-called happy victimizer phenomenon, previous work by Barden, Zelko, Duncan, and Masters (1980) actually provides the earliest

demonstration that young children expect the doers of misdeeds to experience only positive feelings about their actions (though, see also, Thompson & Hoffman, 1980). Specifically, Barden et al. (1980) found that kindergartners differed markedly from older school-age children (i.e., third and sixth graders) in their own reported feelings about engaging in undetected acts of dishonesty, such as, for example, stealing a toy without being caught. When given a choice between "happy," "sad," "mad," "scared," or "just okay" (neutral) the majority of 4- and 5-year-olds in their sample predicted that they would experience happiness despite their dishonesty, whereas, by contrast, the older children (9- to 10-year-olds and 12- to 13-year-olds) indicated that they would most likely experience a negative reaction, such as being scared or sad. The rationale that Barden et al. (1980) offered for this difference was twofold: 1) as a result of being enrolled in school longer, older children were assumed to have simply undergone more "moral training" than the younger children, and so, were thought to have easier access to the proper "canned" answer about how one should feel about their wrongdoings; and 2) by virtue of their being more alert to the seeming omnipresence of moral authorities, older children were perceived as less likely than their younger counterparts to enthusiastically dwell "on the [immediate] success of a dishonest act" (p. 975), as opposed to anxiously anticipating the future likelihood of their apprehension.

As reasonable as this pair of explanations might seem, the now seminal work of Nunner-Winkler and Sodian (1988), which, at least in part, was aimed at extending Barden et al.'s (1980) findings, provides very little in the way of empirical support for either. That is, while Nunner-Winkler and Sodian (1988) did in fact confirm Barden et al.'s (1980) initial findings that younger and older children hold different expectations about the likely feelings of victimizers, they nevertheless found little evidence (even across a wider range of transgressions: stealing, lying,

and overt aggression) that this was either due to differences in children's moral understanding, or because children grew more wary of external sanctions with age. Instead, Nunner-Winkler and Sodian (1988) report, just as most other subsequent researchers have (e.g., Keller et al., 2003; Lourenço, 1997; Nunner-Winkler, 1999), that an overwhelming majority of the children in their study readily recognized when moral norms were violated – in fact, only one child, out of the 100 plus 4- to 8-year-olds they individually interviewed, responded otherwise. They also found, again contrary to Barden et al.'s (1980) speculations, that older children's justifications for their emotion attributions rarely focused on the fear of future punishment, but tended instead to simply reiterate the fact that moral rules had been broken. That is, children did not appear to be responding on the basis of extrinsic considerations, as Barden et al. (1980) had presumed, but rather, seemed to base their judgments on an intrinsic understanding of morality that emphasized victims' harm (see also Smetana, 1995; Turiel, 1998). Little about such findings, it would seem, gives Barden et al. (1980) much to be happy about.

Still, not all of what Nunner-Winkler and Sodian (1988) learned about children's happy victimizer conceptions proves entirely inconsistent with Barden et al.'s (1980) views. In particular, Nunner-Winkler and Sodian report that young children do indeed seem to "dwell" on successful outcomes to action, even when such actions involve victimization. The real upshot of their 1988 study, however, has been to show just how entrenched children's focus on success can be. More specifically, Nunner-Winkler and Sodian report on a series of three studies, the first of which was directed at essentially replicating Barden et al.'s (1980) earlier work, and the latter two aimed at better gauging just how deep the happy victimizer attribution pattern ran in younger children. To accomplish this, they constructed a variety of story conditions in which the salience of the victimizer's gain and the victim's losses was systematically varied. They found that even

in cases stripped of the “tangibility of the profit gained” (Nunner-Winkler & Sodian, 1988, p. 1331), as they called the condition where victimizers received no clear material rewards for their actions, little could be done to deter 4- and 5-year-olds from their predominantly “outcome-oriented” responses. What is perhaps even more striking is that such responses continued to hold even in cases where the victim was shown to be severely hurt (i.e., crying and bleeding) as a direct result of the victimizer’s misdeeds. Indeed, if there were any indication in the stories that the victimizer somehow achieved his/her goals or wants, then, no matter how cold-blooded the means, (s)he was construed as being happy. Interestingly, later work by Wiersma and Laupa (2000), studying 3- to 5-year-olds’ emotion attributions, has shown (pace Nunner-Winkler and Sodian) that when participants see only a person transgressing with no perceivable goals the happy victimizer effect washes out. Certainly, from all this, it would seem reasonable to conclude that “getting what one wants” is the most fundamental heuristic that young children apply in organizing their emotion inferences. Still, the explanation that Nunner-Winkler and Sodian (1988) give to their findings departs from any such straightforward rule-based account, and instead elaborates, if not always convincingly, on matters of moral motivation and identity development.

Specifically, Nunner-Winkler and Sodian (1988) interpret their results in terms of a *moral* attributional shift in early childhood, beginning with the basic outcome-oriented attributions common among 4- and 5-year-olds, and ending with the intrinsic moral orientation evidenced in most 8-year-olds. On the developmental account they propose, children first come to understand moral principles in, as they say, “a purely *informational* sense” (p. 1336) before they begin to treat them as personally binding. They ultimately conclude, much like Barden et al. (1980), that children’s actions, at least initially, appear to be guided less by moral or

normative considerations than by “instrumental cost-benefit calculations” (Nunner-Winkler, 1999, p. 254) of a solely personal nature. They go on, however, to argue that “young children’s ‘amoral’ emotion attributions ... signal[s] a morality that has not gained motivational force in guiding their behavior” (p. 1337). Although on this final point, Nunner-Winkler and Sodian (1988) do not go so far as some (e.g., Rich, 1986) who have argued that “an absence of appropriate constitutive emotion when committing a morally wrong act shows that the person does not fully have a concept of what is morally right” (Rich, 1986, p. 212; cited in Blasi, 1995), they nevertheless imply that young children’s moral knowledge remains vacuous unless, or until, it acquires, as Blasi (1995) has claimed, “*its own* motivational power, one, namely, that is intrinsic in the nature itself of morality” (p. 237).

Consistent with current theorizing about the “moral self” (e.g., Blasi, 1984; Colby & Damon, 1992, 1993; Damon, 1984; Hart & Fegley, 1995; Wren & Noam, 1993), Nunner-Winkler and Sodian’s (1988) findings are often taken as important evidence for the argument that cognitive and motivational aspects of an individual’s identity exist initially as two independent conceptual systems (see, e.g., Damon, 1984, p. 109) that only very gradually, and not until adolescence or early adulthood, grow into a unified and integrated moral identity. Whatever other merits this broader account might have (for reviews, see: Bergman, 2002; Walker & Hennig, 1997), attempts to bolster such views by pointing to Nunner-Winkler and Sodian’s (1988) work is not without its share of obstacles. Perhaps the biggest of these obstacles is that findings from other emotion attribution research suggest a different, and considerably more parsimonious, explanation for why the happy victimizer phenomenon occurs.

Alternative Appraisals of the Early Work. The first of these obstacles has its origins in attribution theory, and particularly in the work of Weiner and his colleagues (Graham & Weiner,

1986; Weiner, 1986; Weiner & Graham, 1984). They argue, in particular, that emotion inferences involve a two-part appraisal process. Part one in this two-step sequence consists simply of evaluating an affectively charged situation in terms of a successful or failed outcome, regardless of its cause. The immediate upshot of this "primary appraisal," as it is called, is the attribution of a global positive or negative emotional reaction, such as "happiness" for success or "sadness" for failure. Because these emotions remain largely removed from any causal analysis of the situation, they have been labeled "outcome-dependent" (see Weiner, 1986, p. 121) affective responses. It is only with the next step, or "secondary appraisal," in this process that a more refined search for the outcome's cause takes place, and the initial, or primary, emotional response is potentially adjusted. This second, and later-occurring, step opens the way for a more highly differentiated set of emotion attributions, typically referred to as "attribution-dependent" affects, that, as Weiner and his colleagues have shown, turn on an individual's perceptions of a variety of causal dimensions. Those studied so far include whether: a) an event's causal locus is perceived as internal or external to the agent, b) the cause is stable or not (e.g., based on an agent's ability or sheer luck), and c) the outcome is under the agent's own volitional control (see Graham & Weiner, 1986, pp. 153-154, for other potential dimensions). Unlike the "primitive" emotions of happy and sad, then, these more complex affective inferences such as pride, guilt, anger, or gratitude all require some more in depth analysis beyond a situation's superficial outcome. To illustrate just how this two-part analysis works, Weiner (1986) paraphrases one of Kant's few culinary claims concerning the feelings associated with eating a good meal. "Everyone at a meal," he remarks, "might enjoy the food, but only the cook could experience pride" (Weiner, 1986, p. 128).

Notwithstanding whatever other culinary insights might be drawn from such an example, the place where this recipe for dividing outcome- and attribution-dependent emotions begins to gain real purchase for understanding the happy victimizer phenomenon is located in the program of developmental research that these attribution theorists have cooked up to investigate children's emotion inferences (for reviews, see Graham & Weiner, 1986; Thompson, 1989). Although few of these studies have been aimed at directly exploring moral issues (though see Thompson, 1987), most provide evidence, not unlike that found in happy victimizer research, that children between the ages of 5 and 7 primarily base their emotion attributions on only the outcomes of situations or events. That is, in the language developed by Weiner and his colleagues, these young children's inferences regarding the emotional consequences of a situation rarely go beyond the very global level of "primary appraisal." Such undifferentiated, outcome-dependent emotion attributions have also been shown to decrease significantly in older children (see, e.g., Graham, 1988), whose responses more typically highlight a growing sophistication in causal, or so-called "secondary," appraisal skills (e.g., Stipek & DeCotis, 1988). In short, these findings from the attribution literature suggest that, identity and motivational factors aside, the happy victimizer phenomenon may well have less to do with supposed moral considerations than with the limited causal analysis that younger children are capable of bringing to bear on the process of emotional inference. That is, on this account, young children base their judgments, not on the morally questionable causal means leading up to an outcome – means that would certainly be factored into any secondary appraisal – but exclusively on their first, or primary, evaluative pass – an enterprise meant only to determine whether an individual's goals have been successfully achieved.

Weiner (1986) has also argued that, while cognitive limitations prevent such secondary appraisals in young children from occurring, adults will also generally skip this second inferential step unless they perceive an outcome as particularly “negative, unexpected, and/or important” (p. 121). That is, in most run-of-the-mill circumstances, primary appraisals generally work well enough, and so, no additional cognitive efforts are made to proceed further or to adjust one’s initial emotional inference. According to this line of reasoning, the happy victimizer phenomenon, contrary to Nunner-Winkler and Sodian’s (1988) argument that it disappears in middle childhood, could be expected to persist, under many common circumstances, even into adulthood. As it turns out, this is exactly what some follow-up studies to Nunner-Winkler and Sodian’s (1988) work have found.

The second general obstacle that Nunner-Winkler and Sodian’s (1988) account faces, then, turns on the findings of other investigators that, as Lourenço (1997) has remarked, “do not support an age-related reversal in victimizer conceptions (i.e., positive to negative emotions)” (p. 426), as Nunner-Winkler and Sodian (1988) reported. Rather, such evidence presents a more complicated developmental picture, sometimes showing only a “partial or ‘subtle’ shift” (Lourenço, 1997, p. 435; see also Keller et al., 2003, p. 13) away from positive emotion attributions in older children, and at other times showing no significant shift at all, even among adults (e.g., Murgatroyd & Robinson, 1993). For example, Murgatroyd and Robinson (1993) reported that 69% of the 5-year-olds and 48% of the 7-year-olds in their study continued to imagine that victimizers were relentlessly happy (see Arsenio, 1988, for similarly high incident rates), as were approximately one-third of the first-year undergraduates they surveyed. This was especially the case when the target persons in their story conditions were portrayed as being in an argument. Although there are grounds for questioning whether Murgatroyd and Robinson’s

(1993) assessment procedures are really comparable to Nunner-Winkler and Sodian's (1988) – particularly since argumentative situations, like those they employed, could be perceived as providing justification for a victimizer's turn to violence – the evidence they do present nevertheless casts doubt on Nunner-Winkler and Sodian's (1988) conclusion that the happy victimizer phenomenon is largely in eclipse by 8 years of age. Clearly, there are circumstances where this is not the case.

The Good, the Bad, and the Morally Ambivalent. A similar conclusion seems to follow from research by Arsenio and Kramer (1992), although here the issue is less about adding support to the notion that the happy victimizer phenomenon persists into adulthood, and more about challenging the precise nature of the so-called "moral shift" evidenced in Nunner-Winkler and Sodian's (1988) work. What Arsenio and Kramer (1992) found is not a dramatic age-related *reversal* from amoral to moral attributions, but a far more gradual, or subtler, shift in children's emotions from almost exclusively positive to multi-valenced, or mixed (i.e., simultaneous positive and negative emotions). In particular, Arsenio and Kramer's (1992) work addressed what, in their words, has been a "basic research limitation" (p. 916) of other studies that narrowly focus on just the *victimizer's* potential feelings, often excluding any further probing of children's broader perceptions of overall *acts of victimization*. That is, in the now essentially standardized procedures used by most investigators of the happy victimizer phenomenon, children are typically only asked "how does the victimizer feel at the end of the story," with little emphasis placed on the victim or on the possibility that other potential emotional responses might be experienced as well. Arsenio and Kramer (1992) hypothesized that if young people were encouraged to reflect more closely on *both* sets of feelings that typically arise in situations of victimization – i.e., the *victimizer's* feelings of gain *and* the *victim's* sense of loss – then more

instances of conflicting or mixed emotion attributions would surface in children's responses. They tested this idea in a group of 4-, 6-, and 8-year-olds, finding that in fact children of all three ages could attribute mixed emotions to the broader acts of victimization (i.e., sad victim, happy victimizer), but that only the two older groups were capable of attributing such multi-valenced responses to just the victimizer alone (i.e., a victimizer who simultaneously experiences both happiness and sadness for his wrongdoing). Moreover, even this apparent step forward was in evidence when their research participants were presented with an additional probe question that pointedly inquired about how else the victimizer might be feeling. Ultimately, under these limits-testing conditions, two-thirds of the 6- and 8-year-olds interviewed by Arsenio and Kramer (1992) responded to such follow-up questions by attributing emotions to victimizers that were opposite (i.e., sad as opposed to happy) to those initially offered (see Yuill et al., 1996, for similar results).

The particular interpretive spin that Arsenio and his colleagues (e.g., Arsenio & Kramer, 1992; Arsenio & Lover, 1995) have eventually come to give this transition in children's thinking is one in which young persons are said to go from understanding "victimization as involving two separate sets of [differently situated] emotional reactions (victimizer & victim) to viewing victimization as a single set of coordinated and interacting reactions" (Arsenio & Lover, 1995, p. 109) all housed in the same person. By such lights, the happy victimizer phenomenon is seen to result from an initial failure on the part of young children to properly coordinate the two points of view that one can take in such situations, leaving most unable to work out how a "victim's pain and loss will moderate one's own happiness regarding the gain produced by 'successful' acts of victimization" (Arsenio & Lover, 1995, p. 109).

Given what turns out to be a rather long history of developmental research concerning young people's difficulties in solving multi-variate, perspective-taking problems (see Chandler, 2001), it is widely thought to be a truism that at least preschoolers regularly struggle, and frequently fail, to coordinate different points of view, or, for that matter, just about any two things in need of coordinating. In the context of happy victimizer research, this is well illustrated by Yuill et al.'s (1996) recent work. Not unlike Arsenio and Kramer (1992), they proposed that children's happy victimizer attributions could be traced back to the fact that the typical victimization story condition affords the possibility of two different "stances": a *personal stance* rooted in "a judgment of satisfaction at a fulfilled desire" and a *moral stance* emphasizing that "morally responsible people are expected to express remorse or shame for their wrongdoings" (p. 459). They found, as a consequence of making various contextual cues more or less salient, that even 5-year-olds could be successfully influenced to respond according to one or the other stance, though only the oldest 10-year-old children were able to spontaneously combine, or coordinate, both. However suggestive Yuill et al.'s (1996) findings might be, what now seems most needed, particularly if Arsenio and his colleagues' argument is to be extended, is new evidence dealing more directly with the matter of when and how children come to understand mixed, or simultaneously occurring, emotions.

Studies on Mixed Emotions: Another Mixed Bag

Harter's Touchstone Research. As is the case with the literature on the happy victimizer phenomenon, only a small handful of studies exist exploring children's understanding of mixed emotions, and the bulk of these seem to relentlessly re-work one, very prominent, piece of research. In this case that research belongs to Harter and her colleagues (Harter, 1983; Harter & Buddin, 1987; Harter & Whitesell, 1989; see also, Reissland, 1985), and, what is continually

being re-worked, is the step-like developmental sequence these investigations have proposed for children's maturing understanding of conflicting emotions. More particularly, Harter et al.'s research program involved asking children between the ages of 3 and 13 to generate examples of situations that might evoke two different emotions. From these examples, they identified three principal levels of understanding: 1) among 3- to 6-year-olds, an insistence that only one emotion can be linked to any given situation at a time; and a denial that two emotions might co-exist; 2) among 6- to 9-year-olds, a recognition of the possibility that two emotions may arise from a situation, but only in a closely ordered temporal sequence; and, finally, 3) among 9- to 13-year-olds, an acceptance that two, or more, emotions may occur at the same time, even in one individual. Collectively, they also found that the valence and the number of targets associated with each emotion formed the two most central dimensions of this sequence. For instance, children's ability to attribute others with two of the *same* valenced emotions in a situation (e.g., angry and discouraged) preceded, and formed the basis for, their later-arriving ability to attribute emotions with *opposing* valences. Altogether, Harter and her colleagues found evidence for five distinct stages in children's evolving conceptions of emotion. Of these, the most developmentally advanced – and, also, the most relevant to Arsenio and Kramer's (1992) case concerning the precise nature of the happy victimizer "shift" – was an understanding (at their stage 5) that one person could potentially feel two conflicting emotions about the same situation at the same moment in time.

The work of Harter and her colleagues has not, of course, gone entirely unchallenged. Some (e.g., Carroll & Steward, 1984) have proposed that certain cognitive accomplishments, such as whether children have achieved concrete operational thought, should be added to Harter's account in order to explain children's progress through their different levels of emotion

understanding. Others (e.g., Wintre, Polivy, & Murray, 1990; Wintre & Vallance, 1994) have argued for the inclusion of additional dimensions, such as emotional intensity, to expand the narrow scope of this work, which focused mainly on emotional valence and multiplicity. In the end, however, most of these “challenges” (although, see, Stein & Trabasso, 1989) can be seen to enrich, rather than to deconstruct, Harter’s early studies, and so, tend to leave her original account largely intact. In addition, the internal logic of Harter’s model has been bolstered by collateral lines of research, including the longstanding literature on the development of role-taking and interpersonal negotiation strategies (Selman, 1971, 1976, 1980, 1981). Selman (1976) has, for example, claimed that children generally begin their understanding of the social and emotional lives of others by first “read[ing] off another’s emotions (such as sad, mad, happy) in those situations where the child knows his own response” (p. 303), before later recognizing that “each person has his own uniquely ordered values or purposes” (p. 304) which sometimes “can be in conflict” (p. 305). Only at this later stage, that Selman and his colleagues have variously called “self-reflective” or “reciprocal role-taking,” do children rely on a cognitive strategy that “not only differentiates between subjective perspectives, but also allows those perspectives to be considered simultaneously” (Yeates, Schultz, & Selman, 1991, p. 372). Selman’s description, then, of children’s socio-cognitive growth provides a close parallel to Harter’s account, and would certainly seem to lend credence to the claim that understanding simultaneous conflicting, or mixed, emotions is a relatively late-arriving developmental achievement. Just how late, however, proves to be a sticking point in much of the literature.

Late- Versus Early-Onset Views. Other research (with the exception of Harris, 1983) has found that the age-norms suggested by Harter and her colleagues are quite conservative. At least this is true when the procedural requirement of having children generate (or recall) their own

examples of emotionally mixed circumstances is lifted (for discussion, see: Denham, 1998, pp. 95-99; Saarni, 1999; pp. 91-94). That is, when commenting on "ready-made" examples, as is the case in happy victimizer studies, children as young as 7 years old have demonstrated the ability to attribute mixed emotions to a single event. Donaldson and Westerman (1986), for instance, reported that the majority of their 7- to 8-year-olds recognized that a story protagonist, who had recently lost a pet cat, would likely feel both happy and sad about receiving a new kitten as a birthday present. Similarly, and though couched in a study of message-referent relations, Lightfoot and Bullock (1990) found that up to half of the second-graders they tested could simultaneously integrate two conflicting emotional reactions to a single situation. Finally, in a training study of conflicting emotions, Bennett and Hiscock (1993) presented evidence that nearly half of the 6-year-olds in their control group could already attribute such mixed emotions to a target person. These findings all stand in sharp contrast to Harter's claim that such abilities are typically evidenced only in children beyond the age of 11.

Notwithstanding whatever else is eventually made of this age debate, the fact remains that Arsenio and Kramer's (1992) report of 6-year-olds offering mixed-valenced emotion attributions can be made to fit with other research that has little to nothing invested in the broader moral questions being addressed here. Even more telling, however, is related work showing the positive effects of prompting children to consider other possibilities beyond their initial responses to such situations (that is, just as Arsenio and Kramer did). Gnepp and Klayman (1992), in particular, hypothesized that "younger children do not think *differently* about alternative emotional possibilities but rather *less* about them" (p. 146; italics added). Much like Weiner and his colleagues' claim that emotion inferences often stop short of a secondary appraisal of the situation, Gnepp and her collaborators (see, e.g., Gnepp, McKee, & Domanic,

1987) proposed that children prematurely cut off consideration of other plausible emotional reactions, even for situations that clearly warrant further reflection, once they have offered up their initial response (see also Harris, 1989, 1995; Peng, Johnson, Pollock, Glasspool, & Harris, 1992). For instance, in a study of children and adults' understanding of emotionally equivocal situations, Gnepp and Klayman (1992) found (as did Murgatroyd and Robinson, 1993) that not only did 6- to 12-year-olds typically offer only uni-valence emotion attributions, but so too did nearly one-third of the university-aged adults that they tested. These results dramatically changed, however, when an additional prompt [i.e., "Could (protagonist) have felt a different way, or not really" (p. 150)] was used at the very end of the interview procedure. In particular, 35% of 6-year-olds, 68% of 8-year-olds, and 90% of 11-year-olds were found, under these conditions, to offer mixed-valenced responses. These results mirrored similar improvements reported in an earlier study by Gnepp et al. (1987) in which 71% of a group of kindergartners and 90% of third-graders gave multi-valenced answers when prompted to do so. They also report in this earlier work that, despite an obvious age trend in children's understanding of mixed emotions, the general ability to entertain multiple possibilities in response to questions outside the domain of emotions proved robust for each of the four grade levels they tested (kindergarten, first-, second-, and third-grades). That is, children as young as 5 were found to offer dual responses to situations that clearly afforded them, suggesting that any earlier failures to do so were a function of the circumstances under which children were tested. A similar conclusion could also be drawn from Bennett and Galpert's (1993) finding that 5-year-olds, and perhaps even younger children, understand the possibility of multiple, conflicting desires.

Summary and Assessment: Where Do We Go from Here?

With these various threads of earlier research gathered together, it now becomes possible, not only to pinpoint some of the more critical holes in the happy victimizer literature, but also to begin piecing together a fuller explanation for why young children are especially vulnerable to making happy victimizer attributions. As a way, then, of narrowing in on the central problems of the literature, this is what can be said so far with a certain degree of confidence.

First, as several studies reviewed here show (Arsenio & Kramer, 1992; Lourenço, 1997; Murgatroyd & Robinson, 1993; Nunner-Winkler & Sodian, 1988; Yuill et al., 1996), the happy victimizer transition is less dramatic than initially described, and likely most pronounced only for those situations or contexts where respondents perceive that a clear moral violation has taken place. Other than work by Wiersma and Laupa (2000), however, the direct manipulation of such situational factors, such as transgressions that may be more or less prohibited (e.g., social-conventional or personal violations), remains a matter in need of further exploration. No research, for instance, has investigated whether children are capable of offering similar emotion attributions for more discretionary situations in which a target person is free to act pro-socially or not (e.g., being a "Good Samaritan"). Introducing this dimension of discretionary and obligatory morality (see Kahn, 1992), would likely yield more variability in the standard mix of happy victimizer findings. I will return to this issue in Study One.

Second, findings from Arsenio and Kramer (1992), as well as Yuill et al. (1996), highlight the fact that, by making the moral stakes in the typical victimizer scenario as salient as possible, and by encouraging children to take every opportunity to exhaust their thoughts about such situations, respondents are more likely to generate a greater number of mixed, or multi-valenced, emotion attributions than what was typically found in earlier research. That is, acts of

victimization have been found to generally afford two central dimensions along which children's responses may proceed – the victimizer's gain and the victim's pain. Children who succumb to the happy victimizer phenomenon may well do so because of a failure to properly balance, or coordinate, these two fundamental aspects of the situation. Closely related to this observation, Keller et al. (2003) have recently argued that the interview procedures standardly used in happy victimizer research may unduly bias children's responses by drawing their attention toward the more factual or descriptive, as opposed to deontic or prescriptive, aspects of the victimizer vignette. Their research demonstrates, in particular, that if child participants are asked what they call a "moral-prescriptive question" (p. 4), such as "how *should* the victimizer feel after the moral violation?", then more negative (or moral) emotion attributions are typically provided, even by the youngest of children. Future work in this area would do well to take steps to ensure that, at least for children, this balancing act between the descriptive and prescriptive is kept as straightforward as possible. Otherwise, whatever lop-sided, single-valenced responses happen to be observed may be more a function of the demand characteristics of the experiment than a true demonstration of children's lack of competence. Again, I re-visit these concerns when discussing the design of Study One in the next chapter.

Third, and closely related to the previous point, given the standard age range (i.e., 5- to 8-years-old) in which the happy victimizer transition is commonly found to occur, most research has indicated that the production of mixed, or multi-valenced, responses is best understood as a component part of the broader repertoire of skills that such young school-aged children typically possess. Still, this nascent ability may be frequently "masked" by a common tendency, sometimes seen even among adults, to prematurely settle on just one, only partially adequate, assessment of an event or situation. This evident default strategy can apparently be overridden,

however, through either additional prompting or by making conflicting aspects of a situation especially salient.

Fourth, and finally, although such a “masking” hypothesis may perhaps account for a sizeable portion of happy victimizer attributions, it is also very likely, as work among attribution theorists suggests (e.g., Graham & Weiner, 1986), that some children fall into the happy victimizer category due to the ineffective causal analyses they commonly apply to most situations. Although no research has directly shown precisely what sort of cognitive constraints might be behind such shortcomings in young children’s reasoning, Arsenio and Lover (1995) have recently alluded to the possibility that a child’s still fledgling “theory of mind” may be one of the more central factors contributing to the happy victimizer findings. In the end, then, such “underlying cognitive constraints,” as they say, “may mean that young children are more morally ‘obtuse’ than ‘resistant’” (p. 111).

With this last piece of the literature review in place, the “stretch” required to link the happy victimizer phenomenon and children’s acquisition of an “interpretive theory of mind” referred to in the introduction begins to seem like a more manageable distance to cover. That is, at least now, the beginnings of a conceptual case can be made for why the waxing and waning of the happy victimizer phenomenon may actually be best understood as one of the natural consequences that attend changes in children’s developing theory of mind – one that largely turns on the question of how they construe agency, and, as I hope to demonstrate, the inherently subjective nature of the knowing process. Still, before turning to a detailed account of the research program that is meant to bring these various matters of agency, interpretation, and children’s theories of mind altogether under a single roof, it is important to first clarify why a conceptual move in this direction seems especially promising. Whatever its merits, the idea is

certainly not a commonplace. At least, as the previous review makes clear, *agency* is not a term commonly evoked as a way of making sense of the standard findings, nor, more generally, is it a construct frequently taken up anywhere in most psychological theories. Two notable exceptions, found in the work of Paul Harris and Albert Bandura, stand against this claim, although neither, as I mean to show, is immune to serious criticism.

Harris and Bandura: Agency or Apparition?

Internalizing “External Audiences.” In turning to the work of both Harris and Bandura as a possible source model for explaining the happy victimizer phenomenon, Harris’ theory seems the most relevant. At least he discusses the happy victimizer findings directly, whereas Bandura does not. In particular, Harris (1989) has argued that the turning point in young children’s reasoning about emotional matters, and particularly acts of victimization, comes with the addition of a new recursive layer in their views about others’ agency, or, as he put it, a shift from “seeing people as simply agents” to “seeing them as observers of their own agency” (p. 92). Central to this account, which Harris acknowledges borrows heavily from the classic writings of Cooley and other early pragmatist philosophers, is the idea that children eventually come to internalize an “external audience” that, in the end, allows them to evaluate their own and others’ actions from a more distant, third-person point of view. Harris has done no empirical work to test this claim, however, and, consequently, it is more of a promising conjecture about, rather than a warranted explanation for, the happy victimizer phenomenon. Moreover, while some support for Harris’ claims might be garnered from the previously cited studies of Murgatroyd and Robinson (1993, 1997), who also seem to share in this more classic formulation of children’s reflective capacities, it is actually with their introduction of an additional “onlooker” into the

standard happy victimizer assessment procedure that the non-agentive features of this classic account become clear (for more general arguments, see Blasi, 1995; Davidson & Youniss, 1995).

More specifically, Murgatroyd and Robinson (1997) found that by using an altered victimizer story line – one where respondents heard that a third story character, the “onlooker,” was observing the victimizer’s actions, and then was seen to react with either approval or disapproval – both children and adults commonly attributed an emotion to the victimizer that matched the onlooker’s positive or negative response. In other words, Murgatroyd and Robinson (1997) provide evidence suggesting that emotional judgments are determined more by how others (i.e., the “audience” in Harris’ account) are imagined to think about a wrongdoing than as the result of applying a moral principle that would independently establish an action’s “wrongness,” no matter what others thought. Although, as Harris’ (1989) writings show, agency is certainly a word that gets paid lip-service in this kind of an explanation, nothing about these findings, nor the broader framework in which they have been presented, seems particularly agentive. If anything, the most that these results appear to demonstrate is the degree to which respondents can be seen to passively accommodate to the will of others, rather than exert one of their own.

Disengaging Moral Agency. A similar conclusion might be drawn about Bandura’s work on morality, although this is perhaps clearest only in his early research (1969; Bandura & McDonald, 1963) on modeling and imitation. Over the years, his particular brand of social-learning theory, or what he (1986) calls “social *cognitive* theory,” has worked to transform, as one critic remarks, “the conceptually lean idea of simply ‘observing a model’ into a semantically richer idea, and hence a more cognitive one” (Wren, 1991, p. 57). While there is some debate (see Martin, Sugarman, & Thompson, 2003; Rottschaefer, 1991, 1998) over just how deep the

agentive thread actually runs in Bandura's theory, his latest work (e.g., Bandura, 1989, 1991, 2001) is certainly one of the few places, outside the happy victimizer literature, where words like *agency*, *morality*, *emotion*, and *victimization* all converge. Specifically, his description of the self-regulatory processes that individuals may use to short-circuit, or disengage, the inhibitory mechanisms and negatively charged emotional reactions that typically attend (and often prevent) most acts of victimization would seem like a rich place to look for developmental insights into the happy victimizer phenomenon. As it turns out, however, development is at best a distant concern in this account.

The series of strategies and mechanisms for moral disengagement that Bandura has detailed in his various writings (e.g., Bandura, 1990, 1999, 2002) include: 1) moral justification, 2) euphemistic labeling, 3) advantageous comparison, 4) dehumanization of victims, 5) distortion or disregard of consequences, 6) displacement and diffusion of responsibility, and 7) attribution of blame to the situation. Although these mechanisms, on whole, are meant to address the broad question of how individuals exercise moral agency in general, only the more particular strategies of shifting blame and responsibility to others or the situation (i.e., the last two in the list) appear to involve the kind of personal agency that, at least I mean to argue, is so central to happy victimizer attributions. That is, accounts from the research literature suggest that children who see happiness as the primary emotional response for a wrongdoer in the happy victimizer story procedure often appear to be ignoring the active hand the victimizer has had in bringing harm to another. A positive emotion attribution, in this case, potentially indicates a failure to appreciate what some have called "responsible agency" (Bratman, 1997, Hart, 1968; Malle, Moses, & Baldwin, 2001), which presupposes intentionality in the form of planning, deliberation, and reasoning. Unfortunately, while Bandura (2001) has discussed each of these various aspects of

responsibility, he has never done so in a developmental context. In fact, the work he has done with young people (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996) seems to assume that there are no developmental differences in their notions of responsibility, nor, for that matter, in any of the other strategies they might use to disengage from morally hazardous actions.

Another related weakness in Bandura's account, as Wainryb (2000) points out, concerns his lack of clarity about "whether disengagement practices are deliberate or not" (p. 36). Certainly, at times, he presents them as the un-intended, or non-deliberate, consequences of broader societal structures operating on individuals who have little control over their actions. For instance, Bandura (1999) remarks that "some of the moral disengagement practices, such as diffusion and displacement of responsibility, are rooted in organizational and authority structures of societal systems" (p. 207). Nevertheless, he goes on to claim that "people are producers as well as products of social systems" (p. 207), suggesting that even at this broad societal level of analysis, individuals may be seen to play a conscious role in sustaining immoral or unjust practices, and should be held accountable. A similar ambiguity runs through his account of psychological functioning at the more individual level. Bandura (1990), for example, discounts the possibility of unconscious self-deception and argues, instead, that individuals ignore or deny the detrimental or immoral consequences of their actions by remaining willfully ignorant, or "keep[ing] themselves intentionally uninformed" (p. 189). People do not typically, as he says, "go looking for evidence of their culpability or the harmful effects of their actions" (Bandura, 1990, p. 189). On the other hand, he is equally quick to claim that "people may be misled by those they trust into believing that violent means are morally right" and that quite often "in these instances, people who have been persuaded are not lying to themselves" (Bandura, 1990, p. 190). Regardless of how this ambiguous strand in Bandura's work is eventually resolved, there can be

little doubt that in order for disengagement mechanisms to be purposefully and selectively activated, individuals must necessarily be aware of them. That is, their *use*, in order to count as deliberate *misuse*, requires a level of awareness that Bandura fails to make clear in his theorizing. Such awareness, however, is precisely what young children frequently appear to be missing in their thoughts and actions. It seems unlikely, for instance, that young children attribute positive emotions to victimizers because they somehow recognize that they have initiated a set of disengagement strategies that allow them to feel good about their wrongdoings. From a developmental perspective, it is more likely that children only gradually become aware of these self-regulatory practices. As it stands, however, these and other developmental concerns are largely ignored in Bandura's account.

The upshot of critically reviewing both Bandura's and Harris' work is that guidance about where to fit agency into a developmental account of the happy victimizer phenomenon is in short supply, even in the most promising corner of the current literature. A new and different source model is much needed. By way of an alternative, I mean to propose Piaget's (1954/1981) early account of the will as one such source, and with it, his suggestion that the solution to the happy victimizer puzzle has been under our noses (quite literally) all along.

Piaget and "The Will"

Although perhaps not a particularly well-known piece of his broader legacy, Piaget actually discusses the will at some length in a collection of lectures (1953 – 54) now published as *Intelligence and Affectivity* (1981). There, demonstrating his familiar fondness for analogies, he describes the will as "the affective analogue of intellectual decentration" (Piaget, 1954/1981, p. 64), and argues, more generally, that it reflects the broader "isomorphism" (p. 9) that can be seen to obtain between affective and intellectual structures. The overall intent of this work is to

present a series of “stages” of affective development that closely parallels his more commonly known account of cognitive or intellectual growth. As history has shown, however, this parallel track of affective stages never gained the same notoriety, perhaps because, as Piaget cautioned, “the comparison between affective states and acts of intelligence cannot be pushed too far” (Piaget, 1954/1981, p. 15). Without meaning to throw such cautions into the wind, there is, arguably, more pushing to be done, though, in the present case, it has more to do with drawing out some of the important ways that Piaget’s account can be shown to map onto the standard happy victimizer assessment procedures. To see this, a more descriptive summary of what Piaget actually said about the will and children’s affective development is necessary.

Intelligence and Affectivity. While common sense perhaps tells us that any account of the will, or human agency, is likely to focus on notions like self-determination and individual choice, Piaget’s discussion is, in fact, largely dominated by the more sober matters of moral duty and obligation, or what he called “normative affects” (Piaget, 1954/1981, p. 59). That is, for Piaget, the structure of the will, again perhaps counter to most intuitions, is said to have more in common with logical necessity than with personal freedom. This is because, for Piaget at least, the will is inherently rational, emerging only from a coordinated system of social and personal values. These values, in turn, may be seen as a “veritable logic of feelings” that, as Piaget (1954/1981, p. 13; see also p. 60) remarks, ultimately come to share the same “conservations and invariants” (p. 60) that arise in children’s intellectual growth. Just as children, then, come to appreciate the logical necessities involved in the conservation of matter, or number, or liquid, so to do they develop, according to Piaget (1954/1981), a “conservation of values” (p. 60) that is evident in how they conduct their daily social lives, and eventually commit themselves to the

personal ideals by means of which they plan their futures (compare Frankfurt, 1999, pp. 114-115).

As counter-intuitive as much of this might sound, what should not be surprising is Piaget's commitment to telling a fuller developmental story about how children's wills take shape. That is, obviously, children do not come into the world automatically equipped with a ready-made "scale of values" (Piaget, 1981, p. 9). Rather, on his account, values arrive as a bundle of largely arbitrary desires, or spontaneous impulses (i.e., "non-normative" feelings), that work to effectively *drive* the will. Initially, then, a child's will is not properly said to be his or her own, but is instead determined by considerations that are external to it. Insofar as this is the case, the child's conduct is sometimes said to be "volitionally adventitious" (Frankfurt, 1999, pp. 131-132), or as Piaget (1981) claims, "heteronomous" (p. 65; see also Piaget, 1932/1965).

For many, the prospect of a heteronomic will is likely to sound like no will at all, and with this Piaget would fundamentally agree. In fact, he argues that "to speak of will" properly a very specific set of conditions must apply in which "a conflict between two impulses or tendencies [is] present...[and] the impulse that is initially weaker must become the stronger of the two" (p. 61). An act of will, more specifically, requires that a subject act against the dominant impulse, as Piaget (1954/1981) says, by "subordinating [it] to a permanent scale of values" (p. 65). To help make this subordination process clear, Piaget (1954/1981) draws on the notion of decentration, suggesting that acts of will are essentially the intellectual equivalent to a "change of perspective" (p. 64). Similar to the perceptual manipulations occurring in Piaget's classic conservation problems, the subject masters, in this case, the immediate *affective* configuration of a situation by "connecting it with former situations and, if need be, by

anticipating future ones” (p. 63). That, according to Piaget, is how the mental operation of decentration works, not to mention, one should add, how *autonomy* in the child is first expressed.

A “Happy” Parallel. The point of rehearsing all of this should now be evident. That is, although more often cast as a problem requiring children to reason about various moral emotions, the circumstances, or story conditions, that participants in the standard happy victimizer study are asked to consider can also be seen as having all of the earmarks of Piaget’s classic account concerning the will. This becomes more evident in the following example. Imagine, that two children (of roughly the same age, and same sex) are playing together on a swing, when one of them – the victimizer – behaves badly by pushing the other to the ground so he can have the swing all to himself. What, you are led to wonder, must the victimizer in these circumstances be feeling? And how, you begin to ask, did he come to act this way?

One way in which to answer these two questions is to recognize that, in this case, the victimizer finds himself in a morally hazardous situation because he faces a conflict of two desires. The initially stronger desire – the one exerting an immediate pull on the victimizer’s deliberations – is the wishful desire to be on the swing. Still, in the background of this example, is the fact that these two children appear to have a history of sharing together, and may even be friends. The opposing desire, then – the one that Piaget would characterize as normative – is embedded in a broader set of publicly shared values that prohibit morally questionable actions. In short, this second, or background, desire concerns the maintenance of friendly relationships, and more generally, conforming to the rules of behaving in a morally upstanding way.

Following this line of reasoning, a case can be made that the victimizer, in failing to overcome his initial, affectively salient impulse, experiences a “collapse” of the will, and ends up pushing his friend from the swing. The parallel to be drawn from reviewing these story

elements, if not already obvious, turns directly on Piaget's general characterization of the will. That is, the general functioning of the will, and this particular instance of victimization, both appear to involve the presence of two opposing impulses, or desires, and the need to subordinate the more dominant one to the one that is less affectively charged.

In the end, the broader implications of this account are twofold. The first is that, while the methods used to assess the happy victimizer phenomenon are most often characterized as a set of procedural means for eliciting children's best thoughts about various moral emotions, it may be more accurate to describe them as a measurement strategy for exploring children's notions of the will. Second, and more central to the work undertaken in the first of the studies reported here (see Chapter III), is that, if the standard happy victimizer stimulus materials (i.e., the story conditions) are really all about matters of the will, then it follows that children's developing conceptions of agency – such as those to be detailed in the next section – should intersect in systematic ways with their understanding of the story, and particularly, the emotions they attribute to its characters. Put in terms of a testable hypothesis, it should be possible to predict children's responses to the happy victimizer procedure based on their scores from other measures meant to more directly gauge their understanding of agency. All that would be required to follow through on this idea would be to find such more direct measures.

As it is, my research colleagues and I are deeply embroiled in a program of study that closely fits this bill of particulars (Carpendale & Chandler, 1996; Chandler & Sokol, 1999; Lalonde & Chandler, 2002). In particular, the procedures we have developed to gauge children's evolving notions of interpretation can, with some minor conceptual adjustments, be reworked – or better, reinterpreted – as an assessment tool for identifying different developmental levels of understanding human agentiveness.

Interpretation and Human Agency

Without going into the long history behind these collaborative efforts (see Chandler & Sokol, 1999), it is nevertheless useful to briefly preface these remarks about interpretation by saying something about how the research program began. This work had its origins in a concern that the broad theories-of-mind enterprise was guilty of artificially truncating the course of children's developing conceptions of the knowing process by narrowly focusing attention on their abilities to understand the possibility of "false beliefs." One consequence of this singular preoccupation with false belief understanding has been the assumption that children's success on standard "unexpected change" or "unexpected contents" tasks – the so-called "litmus tests" (Wellman, 1990) for understanding representational life – is synonymous, as Perner (1991) has claimed, with a full understanding of "knowledge as representation with all its essential characteristics...[including] *interpretation*" (p. 275). Work from our lab, however, casts doubt on such strong claims. We argue, in opposition to Perner and other like-minded theory-theorists (e.g. Gopnik & Wellman, 1994), that the so-called "representational diversity" afforded by standard false belief tasks grows entirely out of the fact that the experimental conditions simply leave the relevant story protagonist ignorant of certain key pieces of information. Perhaps an appreciation of the difference between those playing with and without a full deck of information can be counted as a demonstration of representational diversity, after a fashion, but is clearly *not* what is ordinarily intended by talk of the interpretive nature of human understanding – talk that reasonably assumes that two individuals who have had full access to the same information can still take away different meanings from it. More specifically, the research we have conducted to date (Carpendale & Chandler, 1996; Chandler & Lalonde, 1996; Lalonde & Chandler, 2002; Sokol, Chandler, & Jones, 2004) shows that, while 3- or 4-year-olds may well appreciate that

those given access to different information will hold differing beliefs (the basis for claims about false belief understanding), it is typically not until they are 7 or 8 that children also recognize that those who have the same information can also make different interpretive sense of it. What makes these findings important here is not, however, that they draw attention to the shortcomings of standard, but profligate, claims about simple, ignorance-based measures of false belief understanding. Rather, it is that getting clear about what either does or does not constitute a genuine interpretive competence also allows us to better explicate the different forms of *agency* that run through young children's developing beliefs about belief. This is perhaps best seen in the distinction typically drawn between a *copy theory* and an *interpretive theory* of mind (e.g., Chandler & Boyes, 1982).

Copy-Theories Versus Interpretive-Theories. A frequent and widely accepted claim in theories-of-mind circles is that young children are first committed to a "causal," or passive (Pillow, 1988, 1995), conception of the knowing process. By this account, young persons first begin the epistemic enterprise by effectively treating their own and others' 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). An example of such a "copy theory" at work can be seen in children's early attempts at deception in which they rely on relatively straightforward efforts to keep others in the dark about certain critical details of a situation. The easiest of these to achieve – and so typically the developmentally earliest to be seen, even in children as young as 2-1/2 or 3 (Chandler, Fritz, & Hala, 1989; Hala & Chandler, 1996) – is that of simple secret keeping. These instances of basic deception turn on little more than efforts to withhold key information by keeping one's own counsel, or through some other rough equivalent of hiding one's light under a bushel. Still, because such "clamming up" activities have been

construed in certain functionalist quarters (see Mitchell, 1986; Sodian, Taylor, Harris, & Perner, 1991) as reflecting more of a pre-programmed behavioral response to adversity than a conscientious effort to deceive, deciding with some confidence whether young children do or do not understand the knowing process from a copy-theorist perspective generally requires obtaining evidence of a more active and outward sort – typically evidence demonstrating some discernable attempt to actively dis-inform others. What is important to realize even here, though, is that these more active attempts to dupe or mislead others – strategies also commonly present in young preschoolers – turn on the same basic principle of ignorance, and particularly, the notion that others only “believe what they see” or have otherwise gained informational access to. That is, while secret keeping strategies are effective because another’s state of ignorance is preserved by preventing new and *relevant* facts from coming to light, tactics of disinformation essentially work to achieve the same effect by introducing all *but* the relevant details of a situation. While perhaps in some ways this latter strategy reflects a more sophisticated approach to information management, it clearly remains well within the bounds of a copy-theory framework.

What remains more or less hidden in the background of this description of deception, and the copy theory of knowledge on which it rests, is exactly where such an early folk epistemology leaves children in terms of their understanding of the *source* of the mind’s activity and their consequent notions of *epistemic agency*. A copy theory, it needs to be made clear, naturally leads children to view epistemic agency as originating outside, or external to, the individual. This follows from the fact that, based on such a view, the mind is, as Rorty (1979) claims, no more than a “mirror of nature” that essentially reflects internally on the mind’s eye what can be seen externally in the world outside our skins. Although some (e.g., Perner & Davies, 1991) have claimed that the mind, even in this evident state of passive accommodation, is nevertheless

“doing” something, it is no more “active” – most will recognize – than is any other mirror or reflective surface when light strikes it. The mind’s activity, in this case, is a process that is neither initiated nor controlled by an active subject, and is akin instead to what philosophers of action (e.g., Frankfurt, 1988; Taylor, 1966; Velleman, 2000) have characterized as behavioral *re-actions*, or mere internal *events*, as opposed to more “full-blooded” (Velleman, 1993) or “meaningful” (Moya, 1990) *actions* that can be seen as belonging to, and initiated by, autonomous, self-moving agents. For the young copy-theorist, mental life amounts to “psychological and physiological *events*” that, as Velleman (1993) notes, may be said to “take place inside a person,” but then (and here’s the catch) “the person serves merely as the arena for these events: he takes no active part” (p. 189; italics added). In other words, for those children who hold a copy theory of mind, it is best to say that they recognize the mere *re-actions* of epistemic “patients” rather than the meaning-making *actions* of knowing “agents.”

Taking the Interpretive Turn. By contrast, when children come to hold what my research colleagues and I have called an “interpretive” view of the knowing process, they relinquish their earlier passive or reactive conceptions of the mind in favor of a more fully agentic understanding of mental life. Rather than viewing the mind as passively accommodating to the impact of the environing world, young people who subscribe to such an interpretive conception understand themselves and others as actively transforming the world by assimilating experience to the mind’s existing knowledge structures. By this interpretive account, then, the epistemic process is viewed as being made up of two complementary parts, including, not only of the “objective” uptake of information from the world, but also the “subjective” action of giving such input interpretive meaning. That is, children holding such an interpretive theory of mind no longer view the knowing process as just another simple instance of, what some have called (e.g.,

Bandura, 1986, p. 12), “mechanical agency” where all activity or movement originates from outside of persons. Rather, they attribute the mind’s activity to an autonomous and inherently active subject, who, by personally engaging the world, does more than simply react in some mechanistic fashion. Instead, such agents are seen to purposefully pursue and construct their own meaning from experience. Such young interpretive theorists, it should be clear, come to regard themselves and others, as what Scholastic philosophers once described as, “prime movers,” and having done so, they begin to lay the conceptual groundwork for appreciating the fullness and depth of real actions, as opposed to mere reactions or other behaviorally contingent moves that omit the participation of an active subject.

To quickly summarize, then, the boundary conditions for these two categories of epistemic agency are set by the following contrastive features. On the one hand, there are passive, externalistic, and mechanically oriented “copy theories” to which even preschool children, at the very beginning of their efforts to piece together some coherent understanding of mental life, appear to subscribe. On the other, there are more active, internalistic, and autonomous “interpretive” views of the knowing process to which older (typically school-aged) children eventually come, and which allow them to judge their own earlier, and now dated, copy theories of mind as inadequate for the task of capturing the fullness of human actions. With these definitional matters now out of the way, there is finally room to state more fully the hypotheses under consideration in Study One.

III. STUDY ONE

Exploring the Cognitive Constraints Surrounding the

Happy Victimizer Phenomenon

Given the range of issues covered in Chapter II's review of the literature, it is perhaps not hard to see why Keller et al. (2003) have recently referred to the happy victimizer phenomenon as a "multi-faceted" problem. Of all these facets or dimensions, however, there appears to be one that especially calls out for further investigation. This issue concerns the cognitive constraints that children's developing theories of mind would appear to place on the processes of emotion attribution. As the previous account on the relation between interpretation and agency worked to make clear, there are good reasons to expect that the transition to an interpretive view of mental life will also mark a change in the pattern of emotions that children can attribute to others, including the emotions of victimizers, happy or otherwise. More specifically, it follows, based on the argument outlined earlier, that young children who attribute positive (or "happy") emotions to story characters engaged in immoral activities ordinarily do so because they possess only a "copy theory of mind," and with it, a truncated view of human agency – a view in which *end-states* or outcomes (i.e., the external "operators" in a framework of "mechanical agency") are seen as the central organizers of others' behavior. Such entry-level views of mental life contrast with those of typically older children who standardly subscribe to a more interpretive conception of the knowing process – a view that, by definition, includes at least the beginnings of a working understanding of autonomous agency. For these young interpretive theorists, a fuller, person-centered, or *means-based*, orientation toward action is seen to be in place, which, in turn, allows them to make more complex, or mixed, emotion attributions. Such attributions would include, not only recognizing that others would likely feel happy about having satisfied a

desire, but also might experience sadness or guilt for having acted badly along the way. In fact, these two claims taken together form the central hypotheses to be tested in Study One's comparison between interpretive and non-interpretive children's reasoning about the happy victimizer problem.

Design Modifications

Beyond the hypotheses just outlined, several other issues relevant to the design of Study One need to be briefly mentioned. The first of these concerns the awkward fact that, in the story materials commonly used to study the happy victimizer phenomenon, it is possible for participants to suppose that the target protagonist (or victimizer), who is shown to behave badly in the single scenario portrayed, might in fact routinely bully others, and so, as a matter of course, be characteristically happy about the harm done to them. To guard against the prospect of such interpretations, the procedures adopted in Study One followed closely Arsenio and Kramer's (1992) suggestion that such vignettes depict the victimizer and victim as friends. This small addition to the standard storyline, they argue, helps to highlight that the victimizer's moral transgression is other than a routine event, and that he or she is not simply an uncaring "bully." Furthermore, such a modification also serves to draw out the more latent conflict of desires that, if Piaget's theory of the will is right, may be seen to arise in the typical happy victimizer scenario – that is, the conflict between the immediate impulse to get what one wants and the more tempered background concern to avoid harming others, especially one's friends.

Second, because children have been found to sometimes overlook or fail to register the harmful effects of the victimizer's actions, the stimulus materials in Study One were also designed to make especially explicit the fact that the victim was hurt and crying as a result of the victimizer's misdeeds. The intent behind such strong visual cuing was to ensure that participants

could not easily ignore the high moral stakes portrayed in the vignettes. Third, and for related reasons, changes to the semi-structured interview procedure that usually follows the presentation of each vignette were also introduced. These included the addition of a minor prompt (simply “*what else might the victimizer be feeling?*”) following the more standard test question (“*how does the victimizer feel?*”) – changes that, in line with attribution theory (e.g., Graham & Weiner, 1986), encourage participants to engage in a more thorough analysis of the victimizer’s behavior and emotions by prompting them to make both a primary and a secondary appraisal of the stimulus events. An alternative interview question was also developed, and presented to all participants as a separate condition. This more pointed question (“*how does the victimizer feel about acting like that?*”) was intended to especially direct respondents’ attention to the matter of moral agency in each story. The prediction here was that participants who first received this alternate question would be more focused on the moral dimensions of the situation, and so, more likely to provide moral (or negative) emotion attributions to the victimizer than would those who heard the less directive, standard form of the question first. Altogether then, these alterations to the traditional happy victimizer testing procedures were designed to work against the common tendency among young children to attribute exclusively positive feelings to the victimizer, and therefore constitute a conservative test for the phenomenon. In short, if the happy victimizer effect could still be elicited in the face of these efforts to facilitate more elaborate responses, then some of the easier criticisms leveled against earlier findings (and potentially those reported here) could be discounted.

Finally, because the bulk of happy victimizer studies deal almost solely with matters of prohibitive morality (i.e., those things one should not do), Study One also introduced an additional, contrasting stimulus condition meant to explore moral agency in a more discretionary,

or prosocial, context. In this contrasting case, the stimulus materials were structurally similar to the standard victimization vignette, with the one important exception that the target protagonist undertook no harmful actions. That is, bad things befell the “victims,” without the target protagonist being directly responsible. The design of this alternative, *non*-victimization condition was meant to place a less restrictive set of situational constraints on children’s reasoning about emotions – one that lifted the obligatory standards present in the traditional storyline (those that prohibit harming others) and left open, instead, a set of more elective, prosocial response options. One foreseeable consequence of this change was to make the potential conflict between competing desires and emotions, featured so clearly in the traditional vignette, far less pronounced, and thereby decrease the likelihood of detecting clear-cut differences between interpretive and non-interpretive children. Even with this liability, however, it was expected that a similar pattern of results would emerge, although it was anticipated that differences between those with and without an interpretive theory of mind would likely prove less robust in the *non*-victimization condition.

Method

Participants

A total of 77 children (36 boys and 41 girls), ranging from 5 to 7 years old, were recruited from a local parochial school and a university-based after-school care center. Of these, 24 were 5-year-olds ($M = 5$ years, 6.7 months, $SD = 3.1$ months, range = 60.0 – 70.6 months); 25 were 6-year-olds ($M = 6$ years, 5.2 months, $SD = 3.5$ months, range = 72.4 – 83.3 months); and, 28 were 7-year-olds ($M = 7$ years, 6.4 months, $SD = 4.1$ months, range = 84.0 – 95.9 months). The majority of the children were from white, middle-class families. All had agreed to participate in the study, and their parents or guardians had provided informed consent.

Stimulus Materials and Procedures

A within-subjects design with repeated measures was used. The measures, described below, constituted a battery of three separate procedures: a *Theory-of-Mind Protocol* and two parallel procedures for testing children's emotion understanding, the *Victimizer* and *Non-Victimizer Protocols*. The order in which participants received these three measures was counterbalanced, as were the two trials or conditions contained within each individual protocol. All the participants completed the procedures in two short, one-on-one interview sessions (lasting less than 15 minutes apiece) that took place at the children's school or daycare center, and were conducted no more than one week apart.

Theory-of-Mind Protocol. To gauge participants' understanding of the interpretive nature of knowledge construction, an interview procedure employing Price's (1953) classic "doodle" images as stimulus objects was adapted from measures used by Lalonde and Chandler (2002). The two particular doodles that were selected from this earlier study comprised the two different testing conditions of Study One's *Theory-of-Mind Protocol* (see Appendix A). The first condition, or *Standard Trial*, consisted of showing individual participants a line drawing of "a ship arriving too late to save a drowning witch" (see Figure 1). After discussing the various elements of the picture, the experimenter placed it in an envelope with a small rectangular cut-out that formed a viewing window. This window concealed all but a partial image of the ship's bow and the witch's pointed hat – a drawing, in other words, that amounted to no more than two triangles. This remaining, or restricted, view of the picture then became the ambiguous image, or "doodle," to which two different puppet characters, in this case Tweedle Dee and Tweedle Dum, were exposed. Both puppets, the experimenter indicated beforehand, were to be treated as "real people" who had never seen the picture before, nor heard any part of the discussion of what

the full drawing actually depicted. Following a quick reminder of these conditions, the experimenter then presented the first puppet to each participant, and while pointing to the restricted view of the picture, asked: “*What will Tweedle Dee say this is?*” The children’s responses were recorded, and after putting Tweedle Dee away, the procedures were repeated for the second puppet.

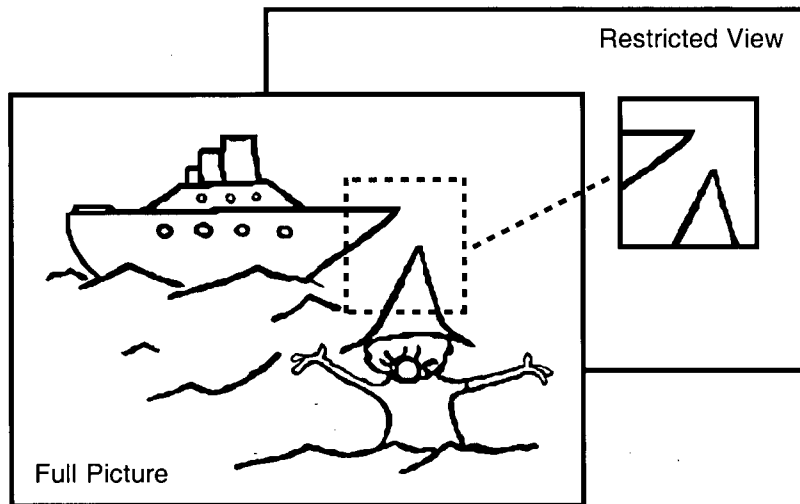


Figure 1: “A ship arriving too late to save a drowning witch”

The second testing condition proceeded much like the first, but began by introducing an opportunity for participants to guess what the doodle actually was before being shown the entire picture (in this case, “an elephant smelling a grapefruit”). In essence, the children in this second *Guess Trial*, who initially saw only the restricted “window” view (i.e., a rectangular portion of the elephant’s trunk and the circular outline of the grapefruit), were put in the same position as the puppet characters in the *Standard Trial* by being asked to imagine what the fuller, underlying drawing might be. Although a set of similar testing conditions was used by Lalonde and Chandler (2002; see Studies 5 & 6) to demonstrate, as they claim, “that only rarely do [participants’] guesses end up coinciding with the beliefs they attribute to the puppets” (p. 189), the possibility that children in this condition might use their own initial ignorance to get a “leg

up” on understanding the interpretive demands faced by the puppet characters seemed like an issue deserving of further exploration. Altogether, the minor procedural variations adopted in this second trial involved simply encouraging participants to make two guesses before being shown the full picture of the elephant and grapefruit, and, following the various questions concerning the possible interpretations available to the puppets (identical to those in the first condition), asking them to recall their own earlier guesses. Both of these trials were presented to the participants in counterbalanced order.

Victimizer Protocol. In addition to the *Theory-of-Mind Protocol* just described, a second interview procedure was administered to explore participants’ reasoning about the emotional consequences of victimizing, or harming, others. The stimulus materials for this procedure, or *Victimizer Protocol* (see Appendix B), were drawn from several sources in the literature, particularly Arsenio and Kramer (1992) and Lourenço (1997). Children were presented a storybook containing a four-paneled cartoon vignette of two young boys or girls (chosen to match the gender of the study participant) who were depicted as friends that regularly took turns playing together on a swing. As the short story unfolded, however, one of the characters (the study’s “target” protagonist) emerged as the victimizer. This character, in particular, was portrayed as refusing to wait his or her turn and, in the penultimate scene, was shown to maliciously push the friend out of the swing. The vignette concluded with the victimizer playing on the swing, his or her facial expression obscured (to avoid influencing children’s inferences about the character’s emotional state), while the friend lay visibly upset on the ground.

To control for potential comprehension and attention problems during the story’s presentation, participants were asked to repeat the vignette’s central elements (with the storybook closed), before then being asked to evaluate the victimizer’s actions and to generate an

alternative means by which he or she could have received a turn. These steps were meant to ensure that participants properly recognized the moral nature of the victimizer's misdeed and understood that non-violent strategies were available to take. Following these control questions (and with the book now re-opened to the last picture frame), participants' attentions were directed to the victim on the ground and asked: "*How does [the victim] feel here?*" After their responses were recorded, the experimenter then pointed to the victimizer, or target protagonist, now on the swing and asked what has come to be the standard test question and request for justification commonly employed in the literature: "*How does [the victimizer] feel at the end of the story?*" and "*Why does he/she feel like that?*" Finally, children who did not spontaneously offer more than one emotion attribution were provided the following simple prompt: "*What else might [the victimizer] be feeling?*"

With the exception of this secondary, follow-up prompt, the interview questions in this portion of the protocol deviated in only minor ways from the standard practices employed in other earlier studies of the happy victimizer phenomenon. For this reason, they were referred to as the *Traditional* portion, or condition, of the interview format. Where this protocol departed more dramatically from the traditional procedures was with the addition of a second, alternative phrasing of the standard test question. In this *Alternative* section of the protocol, children were asked to first recall the victimizer's behavior, and then given a more specific, or directive, form of the central question: "*How does [the victimizer] feel about acting like that?*" The secondary, follow-up prompts to this question remained the same as those in the *Traditional* section. The order of these individual sections within the protocol was alternated so that half the participants received the *Traditional* interview condition first and the other half received the *Alternative* format.

Non-Victimizer Protocol. Finally, a third, and parallel, set of stimulus materials was used to investigate children's understanding of circumstances that, like the *Victimizer* procedure, also involved characters that experienced physical harm, although, this time, not as the result of overt aggression or victimization. The cartoon vignette in this counterpart, *Non-Victimizer Protocol* (see Appendix C), was designed, in particular, to be of similar length, verbal complexity, and again to emphasize friendship relations, but with thematic content (in this case, an accident during a race) that diverged from the *Victimizer* story. The rationale behind this latter change grew out of a concern that participants' responses might otherwise be based on corresponding story elements, rather than on the purpose-built differences intended to separate the *Victimizer* and *Non-Victimizer* scenarios.

At the time these procedures were being developed, an unusual event in sporting news occurred that heavily influenced the choice of story for the protocol. This incident, at the Winter Olympics, involved a race participant who came from last place to win his first gold medal because all of the other competitors collided and fell just before reaching the finish line. Following the race, the winner acknowledged experiencing mixed feelings about his victory – one that clearly resulted from the others' dramatic misfortune (see "Australia Wins First Ever Gold," 2002).

The *Non-Victimizer* stimulus materials modeled on this incident similarly described a race participant who, after falling significantly behind the other participants, won because his or her friends unexpectedly slipped and fell. The last story frame, to which study participants' attentions were drawn during the interview, showed the winner (with his or her face obscured like that of the target protagonist in the *Victimizer* condition) crossing the finish line while the friends lay visibly hurt on the ground. As with the *Victimizer Protocol*, two different, alternating

interview conditions were employed as parts of the *Non-Victimizer* measure, which again, all participants received. The *Non-Victimizer* interview consisted of a parallel set of closely matched questions and follow-up probes, except this time the target protagonist was a boy or girl *racer*, not a *victimizer*, and participants were asked in the more directive, *Alternative* portion of the interview, not about the target's misdeeds, but about the special circumstances allowing him or her to win ("How does [the racer] feel about winning like that?").

Scoring

Theory-of-Mind Protocol. Participants' responses to the doodle materials making up the *Theory-of-Mind Protocol* were scored using procedures set out by Lalonde and Chandler (2002). Following this two-part scoring system, individual belief attributions for each puppet (again, in this case, Tweedle Dee and Tweedle Dum) were first coded according to one of three basic categories, before then being combined into scores of *Non-Interpretive*, *Transitional*, or *Interpretive*. As it turns out, just one of the basic scoring designations (the last one to be described) could be counted toward an *Interpretive* understanding of the knowing process. The other, *Non-Interpretive*, response categories generally reflected mistakes that, especially the youngest children, were prone to make while trying to keep their own privileged knowledge of what the pictures "really were" separate from what the puppets could feasibly know. In particular, a response indicating that Tweedle Dee, despite not having seen the complete drawing, would still think the doodle was "a ship and a witch" made obvious reference to fuller details of the underlying picture. Accordingly, such responses were coded as *Reality Errors*. Other similar mistakes tended to be less explicit, but still contained trace elements from the fuller picture (e.g., "A ship about to strike an iceberg"). These, in turn, were scored as *Contamination Errors*. Finally, responses that were entirely free of either error-type – that is, belief attributions

containing no references to the underlying pictures – were counted as important building blocks toward a more genuine understanding of “representational diversity,” or how, in particular, others can be seen to actively interpret ambiguous situations or events in any variety of divergent ways. Because only these latter response-types showed such an appreciation for the diversity of beliefs others can legitimately hold, only they received the scoring designations of *Divergent Beliefs*.

Clearly, recognizing that others may hold to beliefs different, or *Divergent*, from one’s own constitutes a fundamental part of what it means to subscribe to an interpretive theory of mind. Nevertheless, such belief designations, by themselves, provided only the first half of what Lalonde and Chandler (2002) operationalized as an *Interpretive* understanding. The second portion of the scoring procedure involved a closer evaluation of the pattern of combined belief attributions within each trial – that is, the attributions for both Tweedle Dee and Tweedle Dum for a particular doodle. This step in the scoring scheme followed directly from Lalonde and Chandler’s (2002) definition of interpretation. On this account, a defining factor of an *Interpretive* understanding is “that it is possible for two persons to be exposed to *precisely the same* information or stimulus event and yet to arrive at *different* opinions about what is still the self-same reality” (p. 192). According to this standard, participants in the present study needed to provide distinct *Divergent Belief* attributions for each puppet character (i.e., Tweedle Dee thinks it is this, while Tweedle Dum thinks it is that) in order for their response combinations to be scored as *Interpretive*. If, on the other hand, children gave responses for Tweedle Dee and Tweedle Dum that were identical, or, if either contained a *Reality* or *Contamination Error*, then the response pair was characterized as *Non-Interpretive*.

Together, these two scoring categories were expected to adequately capture the majority of participants' responses. Where this "either-or" strategy could not be made to fit, however, was with participants who successfully provided an *Interpretive* response pair in one trial, but then went on to fail to do the same in the other remaining trial. These "in-between" cases, involving both *Interpretive* and *Non-Interpretive* responses, required the inclusion of a *Transitional* category.

Several criteria needed to be met in order for children's responses to be designated as *Transitional*. Beyond the basic incongruence between the conditions, children also needed to show at least a partial understanding of representational diversity in their unsuccessful, *Non-Interpretive* condition. Specifically, for a case to receive an overall *Theory-of-Mind* score of *Transitional*, one of the two attributions in the unsuccessful trial had to be judged as meeting the *Divergent Belief* criteria listed above. A response pair in this condition that did not contain at least one *Divergent Belief* attribution – that is, a pair where both attributions were coded as *Reality* or *Contamination Errors* – failed then to meet the *Transitional* criteria. These cases, in turn, received an overall score of *Non-Interpretive*.

Altogether, then, participants' overall *Theory-of-Mind* score across the two trials could be fully *Interpretive*, *Transitional*, or *Non-Interpretive*. As described in the following Results section, each of these categories was converted to a numerical equivalent for purposes of statistical comparison.

Victimizer and Non-Victimizer Protocols. Lastly, the scoring procedures for the *Victimizer* and *Non-Victimizer Protocols* involved determining the overall valence, or "emotional charge," of the feelings that participants' attributed to the target protagonist (the "victimizer" or the race "winner") in each trial of the two vignettes. The most common emotion attributions,

happy and sad, were coded as expressing positive and negative valences, respectively. Because participants were prompted to provide more than one attribution – that is, both a primary and a secondary appraisal of the situation – a response pair in any given trial could be entirely positive, entirely negative, or a mixture of both. For this reason, trials were scored as *Single-Valence* or *Multi-Valence*. Where the prompt proved ineffective, and only one response was provided, the score, by default, was *Single-Valence*.

To quickly summarize, if the rationale guiding the hypotheses of Study One is valid, the distribution of these scoring categories should ideally conform to the following pattern of results. *Multi-Valence* attributions should strongly co-vary with *Interpretive* responses to the doodle materials, whereas *Single-Valence* attributions should more typically be associated with *Non-Interpretive* responses. No such strong relations were predicted to hold for the *Transitional* category, although given its “in-between” status (i.e., the balancing of *Interpretive* and *Non-Interpretive* responses), one serious possibility was that it too would comprise of a similar admixture of *Multi-* and *Single-Valence* attributions.

Results

Theory-of-Mind Protocol

Numerical equivalents for the *Theory-of-Mind* scoring categories were produced by using a system in which a single point was awarded for each *Divergent Belief* that participants attributed to the puppet characters. Given the protocol’s two trials, with two response opportunities within each, a maximum of 4 points was possible. Following the scoring criteria detailed in the Methods section, a total of 0, 1, or 2 fell into the *Non-Interpretive* range; a score of 3 was *Transitional*; and a perfect 4 was fully *Interpretive*. The mean *Theory-of-Mind* score for the entire sample ($N=77$) was 2.78 ($SD=1.33$), and the breakdown of the categories was fairly

even: 36.4% were *Non-Interpretive*, 22.1% were *Transitional*; and 41.6% were *Interpretive*.

Inter-rater agreement calculated on 25, or approximately one-third, of the protocols was 92%

[Cohen's kappa=.89, $p<.001$.]

Comparing the participants' scores using a 2 x 2 x 2 (Gender x Order x Trial) Between-Within ANCOVA with Age designated as a covariate revealed no statistically significant main effects, nor significant interactions, for any of the factors, except for Age [$F(1,72)=26.34$, $p<.001$]. In this case, the level of participants' interpretive understanding – that is, their *Theory-of-Mind* score – was found to be strongly age-graded [$t(76)=.52$, $p<.001$ (one-tailed)], as evidenced by a steady progression in mean scores for each age group: 1.67 ($SD=1.24$), 2.96 ($SD=1.24$), and 3.59 ($SD=.69$) for 5-, 6-, and 7-year-olds, respectively. Consistent with previous findings (Carpendale & Chandler, 1996; Lalonde & Chandler, 2002), these scores indicated that children do not generally achieve a fully *Interpretive* level of understanding until roughly 7 years of age.

For the purposes of exploring the happy victimizer phenomenon, replicating this basic pattern of age differences in the onset of an interpretive theory of mind was of critical importance for testing the study's main hypotheses. What was clearly less critical, but nevertheless of interest, was the potential impact of the "guessing" manipulation (i.e., the *Guess Trial*) in the adapted *Theory-of-Mind* interview procedure. Here the findings actually diverged somewhat from earlier results. That is, contrary to Lalonde and Chandler's (2002) findings, participants' use of their own guesses as attributions to the puppets was quite a common occurrence. For instance, of the 32 participants who were scored as fully *Interpretive*, 21, or nearly two-thirds, applied at least one of their own initial guesses when asked what the puppet characters would say the doodle was. A similar proportion of *Transitional* respondents, 11 of

17 (or 65%), did the same. Using the content of one's guesses, however, was not the only notable difference between the current and previous findings. It also appeared that the process of guessing had an impact on participants' responses and, in fact, was responsible for marginal improvements in their scores. Specifically, the test for an interaction between Trial and Order identified a clear statistical trend [$F(1,72)=3.78, p=.056$], indicating that participants who received the *Guess Trial* before the *Standard Trial* (Order #1) performed slightly better than those receiving the opposite order (Order #2). This difference was evident in the comparison of the *Standard Trial* mean scores – the trial expected to be the most difficult – for Orders #1 and #2, respectively: 1.36 versus 1.22 (out of a possible 2).

Although these latter results did not critically impact the central goals of the present study, some steps seemed in order to minimize whatever possible effects they might have on further statistical analyses. This was accomplished by computing the mean score for the two trials, which effectively washed out any order effect. An overall *Theory-of-Mind* score, then, was used in comparisons with the other measures.

Victimizer Protocol

Children's responses to the control questions indicated that, in all but one instance, participants regularly understood that the target protagonist's actions were wrong and that more suitable behavioral options were available – the simplest, and most common, suggestion being that he or she “could have asked for a turn.” Using a similar data reduction strategy to that employed with the *Theory-of-Mind* findings, response pairs in the *Victimizer Protocol* were given one point if they fit the *Multi-Valence* criteria, and no points if responses were *Single-Valenced* only. This numerical system allowed for a total *Emotion* score of 0, 1, or 2 across the protocol's two conditions. The overall average for the sample ($N=77$) was 0.82 ($SD=.82$).

Again, just as with the *Theory-of-Mind* scores, inter-rater agreement on one-third of the protocols was nearly perfect, or 96% [Cohen's kappa=.93, $p<.001$].

Participants' scores were compared again using a 2 x 2 x 2 (Gender x Order x Condition) Between-Within ANCOVA with Age as the covariate. No main effects, nor significant interactions, were found, although a statistical trend was detected in the test for an interaction between Gender and Condition [$F(1,72)=3.05$, $p=.085$], and, as anticipated, Age was again a significant covariate [$F(1,72)=10.31$, $p=.002$]. Potential gender differences notwithstanding, the central finding in these data was that participants' level of *Emotion* understanding, just as their *Theory-of-Mind* scores, was strongly age-graded [$r(76)=.36$, $p=.001$ (one-tailed)]. Given such similarly patterned distributions, it was not surprising to find that children's overall scores on the *Theory-of-Mind* and *Victimizer* measures were significantly correlated [$r(76)=.463$, $p<.001$ (one-tailed)]. Critical to the study's main hypotheses, however, was the fact that this correlation was only modestly affected when the common factor of Age was partialled out [$r(73)=.33$, $p=.002$ (one-tailed)]. Finally, and supporting the argument that theory of mind constitutes a distinct cognitive factor in the determination of children's emotion understanding, the correlation between *Theory-of-Mind* and Age remained strong when controlling for participants' *Emotion* scores [$r(73)=.43$, $p<.001$ (one-tailed)], while the reverse (i.e., treating *Theory-of-Mind* as the control factor) was importantly weakened [$r(73)=.18$, $p=.06$ (one-tailed)]. This strong relationship between the level of children's emotion and theory-of-mind understanding is more clearly illustrated in Figure 2, which shows the break-down of multi-valence, or mixed, emotion attributions made by participants who were scored as having either an *Interpretive*, *Transitional*, or *Non-Interpretive* conception of mental life.

Victimizer Protocol

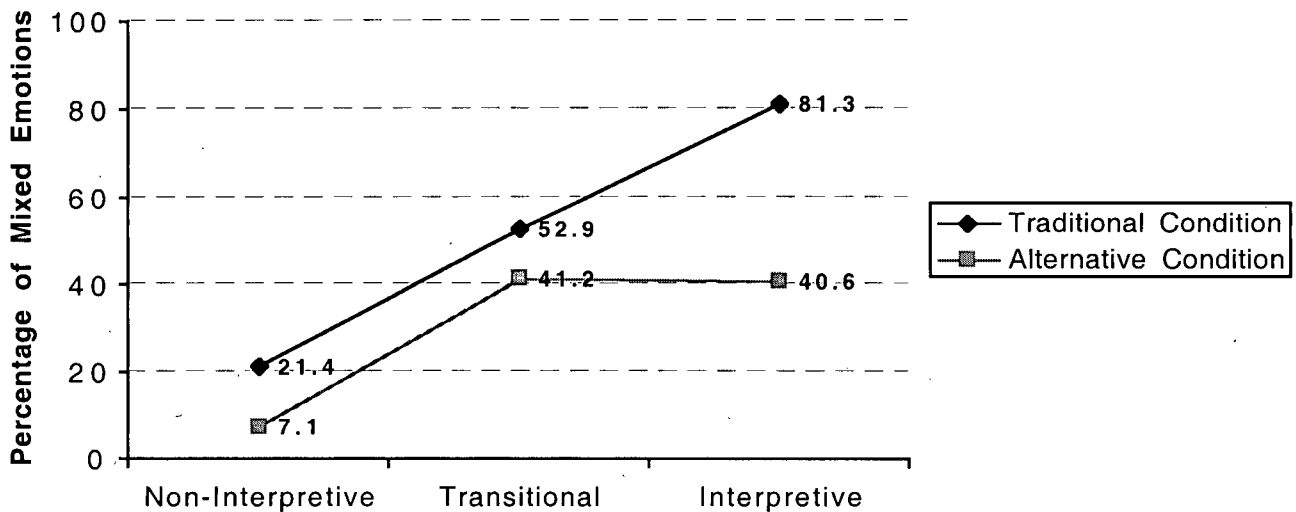


Figure 2: Mixed Emotion Attributions grouped by Theory-of-Mind Score

Although these summary findings suggest relatively few differences between the *Traditional* and *Alternative* portions of the protocol, a more detailed break-down of the individual emotion attributions offered in each case indicates otherwise. This more targeted analysis was accomplished by applying a strategy borrowed from Weiner and his colleagues (Graham & Weiner, 1986; Weiner & Graham, 1984; see Chapter II for further discussion) in which participants' response pairs were treated as separate or distinct *primary* and *secondary* appraisals of the stimulus conditions. Consistent with Yuill et al.'s (1996) findings, the way in which participants were led, or "primed," to attend to the stimulus events – in particular, whether they first received a relatively neutral question (*Traditional Condition*) or, instead, one that directed them to the morally relevant features of the situation (*Alternative Condition*) – appeared to have a systematic impact on how they ordered their individual responses. In particular, the happy victimizer phenomenon was most evident in the *Traditional Condition*. Here a simple

tally of responses to the first, or primary, test question of “*How does [the victimizer] feel at the end of the story?*” indicated that “*Happy*” (66.2%) and “*Good*” (10.4%) were by far the most frequent emotion attributions given, whereas responses reflecting evident moral considerations, such as “*Sad*,” “*Bad*,” “*Sorry*,” or “*Guilty*,” were much less common (18.2%). Participants’ reactions to the follow-up prompt (“*What else might [the victimizer] be feeling?*”) – that is, their secondary attribution – led to real, but still marginal changes in this predominantly positive trend. That is, even though the number of emotion attributions expressing remorse (e.g., “*Sad*” and “*Sorry*”) and negative self-evaluations (“*Upset*” and “*Foolish*”) doubled after prompting (up from 18.2% to 37.7%), a sizeable proportion of participants (35.1%) continued to see no reason to adjust their initial responses, or, if they did, to simply reinforce what they said before (e.g., to supplement “*Good*” with “*Glad*”).

In contrast, this pattern of individual responses was almost completely reversed in the *Alternative* portion of the interview. In this case, the moral concerns expressed in participants’ initial, or primary, attributions (i.e., their responses before the prompt) were three times greater than those found in the *Traditional Condition*: 59.8% versus 18.2%. Whereas, attributions like “*Happy*” and “*Good*” now accounted for no more than one quarter (24.7%) of participants’ responses [McNemar’s Test for Within-Subjects Proportions, $\chi^2(1, N=77)=36.1, p<.0005$]. Interestingly, the follow-up prompt in the *Alternative Condition* did not show the same effectiveness in eliciting oppositely valenced responses as the *Traditional* format. This was evident in the direct comparison of *Multi-Valence* attributions that were offered in each condition. Specifically, in the *Alternative Condition*, the proportion of such response pairs was nearly half that evident in the *Traditional Condition*: 28.6% versus 53.2% [McNemar’s Test for Within-Subjects Proportions, $\chi^2(1, N=77)=15.70, p<.0005$].

These differences in the pattern of responses for the two interview formats begin to shed some light on the statistical trend that was found when testing for a Gender x Condition interaction. That is, while *Multi-Valence* emotion attributions were generally less common in the *Alternative Condition* (see again Figure 2), this was especially the case for girls, whose responses tended to reflect a greater emphasis on moral considerations than boys. Such an emphasis had the effect of actually lowering girls' *Emotion* scores when compared to the boys: .26 versus .32, respectively. So, although girls may have been showing greater moral commitment in such responses, because emotional maturity was measured by participants' ability to balance multiple emotions, they received fewer points – perhaps somewhat ironically – for holding such commitments. On whole, these modest gender differences when coupled with the “priming effects” found in the two conditions, suggest that other important variables work in conjunction with children's theory of mind to constrain the emotion attribution process. Results reported in the next section (*Non-Victimizer Protocol*) support this conclusion as well.

Finally, it is worth highlighting the fact that overall developmental differences in the *Victimizer* data surfaced only after accounting for participants' secondary responses to the follow-up question or prompt. That is, no significant age effects, in either condition, were apparent from just a simple inspection of children's initial, primary attributions. In the *Alternative Condition*, for example, equally high numbers from each age group (5-, 6-, and 7-year-olds) indicated in their primary responses that the victimizer would display negative feelings following his actions: 70.8%, 84.0%, and 64.3%, respectively [$\chi^2(2, N=77)=2.65$, $p=.27$, N.S.]. In the *Traditional Condition*, on the other hand, just as many were committed in their primary attributions to the exact opposite sentiments and responded that the target protagonist would feel positive emotions: 83.3%, 68.0%, and 78.6% [$\chi^2(2, N=77)=1.70$, $p=.43$,

N.S.]. In other words, without the addition of a secondary prompt, which allowed some participants (generally the older, and particularly those scored as *Interpretive*) to temper their judgments about the target's feelings, the take-home message from this more restricted dataset would have been very much like Keller et al.'s (2003) claim that the happy victimizer phenomenon is largely an artifact of how experimenters pose their questions to young children. As it stands, not only do these broader findings speak to the dramatic effects a slightly altered test question can produce, they also lend critical support to the view that children's attitudes toward victimization show less of a simple *reversal* with age (i.e., from positive to negative feelings), than a growing *complexity* in which both instrumental and moral considerations are held in balance (Arsenio & Kramer, 1992).

Non-Victimizer Protocol

One 6-year-old child was unable to complete the *Non-Victimizer* interview procedure, so analyses in this section were conducted on the responses of 76, instead of 77, participants. Emotion Scores were determined by applying the same procedure used in the *Victimizer Protocol*, and so, they ranged again between 0 and 2. By comparison, the overall average score in the *Non-Victimizer Protocol*, 0.72 ($SD=.83$), was only marginally lower than that observed for the *Victimizer* condition, 0.83 ($SD=.82$) [$t(75)=.942$, $p=.35$, N.S.]. The inter-rater agreement, again calculated on roughly one-third of the protocols, matched that found in the *Victimizer* procedure: 96%, Cohen's kappa=.93, $p<.001$.

As a quick reminder, the reason for including a "non-victimization" scenario in Study One's battery of stimulus materials was to compare children's responses to events or circumstances that were, in many ways, similar to the *Victimizer Protocol*, but that omitted the intentional act of victimization, and introduced, instead, a much different set of moral

obligations. The plan, in other words, was not to simply duplicate the findings from the *Victimizer* procedure. Rather, because the normative, or moral, constraints present in prohibitive contexts (i.e., the “swing story” in the *Victimizer Protocol*) were imagined to be less salient than in discretionary contexts (i.e., the “race story” in the *Non-Victimizer Protocol*), the prediction was that the *Non-Victimizer* results would be like an “imperfect mirror.” That is, these results should reflect a similar general pattern to the *Victimizer* data, but be much less defined. Such a general pattern is exactly what the findings of the following 2 x 2 x 2 (Gender x Order x Trial) Between-Within ANCOVA and follow-up correlational analyses appeared to show.

Just as before, Age emerged as a significant covariate [$F(1,71)=9.06$, $p=.004$], and the test for an interaction between Gender and Condition approached conventional levels of significance [$F(1,71)=3.41$, $p=.069$]. Where the “imperfections” in these *Non-Victimizer* data began to appear, however, was with the diminished strength of the relation between children’s *Theory-of-Mind* scores and their *Emotion* understanding. That is, while *Theory-of-Mind* and *Emotion* scores in the *Non-Victimizer Protocol* were correlated [$r(75)=.31$, $p=.003$ (one-tailed)] at nearly the same level found in the *Victimizer* data, when Age was controlled in this calculation, the new correlation was reduced by roughly half, and statistically significant only at trend levels [$r(72)=.17$, $p=.078$ (one-tailed)]. In these new testing conditions, children’s *Theory-of-Mind* competence was not, then, the same strong predictor of *Emotion* understanding as before, again suggesting that other factors influence the emotion attribution process.

Although this latter finding certainly points to some differences between the *Victimizer* and *Non-Victimizer Protocols*, these would appear to be relatively minor. A closer comparison of the *Traditional* and *Alternative Conditions* within each procedure, however, reveals an important difference. By separating participants’ responses again into primary and secondary

appraisals (as was done with the analyses of the *Victimizer* data) a much different pattern of positive (e.g., “Happy,” “Good,” or “Proud”) and negative (“Sad,” “Bad,” or “Sorry”) emotion attributions emerges in children’s primary responses (i.e., attributions made before receiving the secondary prompt). Specifically, the reversal of emotion attributions (i.e., from predominantly positive to negative) found in the *Traditional* and *Alternative Conditions* of the *Victimizer* procedure was absent in the responses that children gave in the *Non-Victimizer Protocol*. In short, participants were essentially responding the same way for each condition of the *Non-Victimizer Protocol*, whereas opposing responses were elicited by the same interview conditions in the *Victimizer Protocol*. This can be seen more clearly in Figures 3 and 4, which show the proportion of children’s positive and negative emotion attributions in each condition of the two protocols.

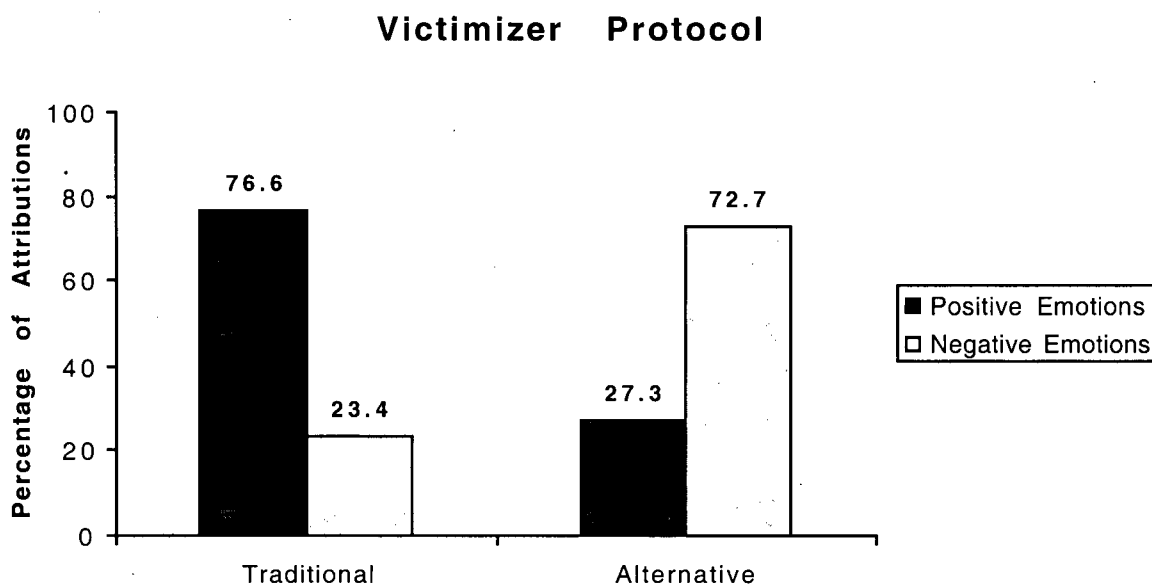


Figure 3: Primary Emotion Attributions in the Victimizer Protocol

Non-Victimizer Protocol

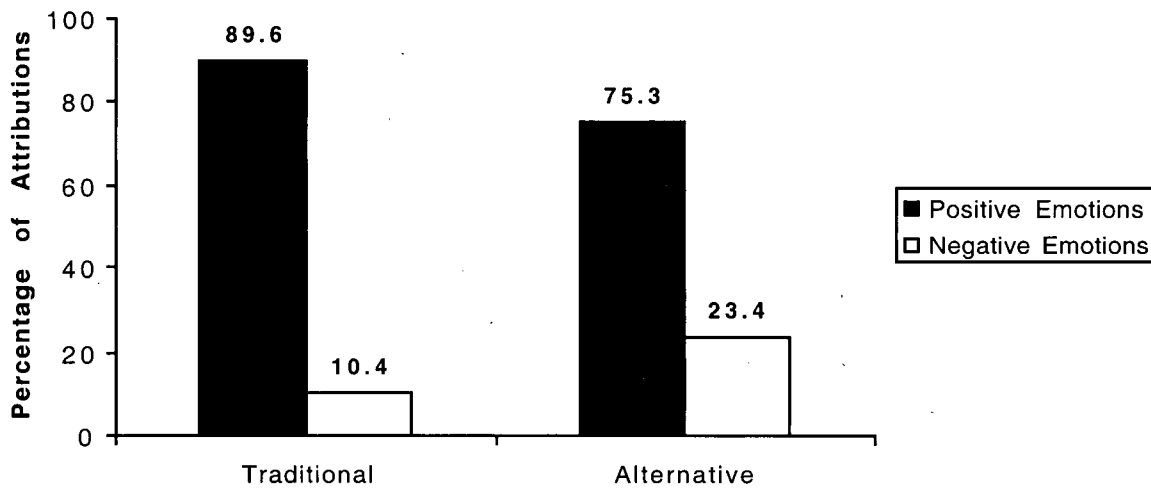


Figure 4: Primary Emotion Attributions in the Non-Victimizer Protocol

Instead of a reversal, the emotion attributions offered in the *Non-Victimizer Protocol* were generally positive, with only a 14.3% reduction of positive responses between the *Traditional* and the *Alternative* interview conditions [McNemar's Test for Within-Subjects Proportions, $\chi^2(1, N=76)=7.14, p<.05$]. By comparison, the *Victimizer Protocol* showed almost a 50% reduction. Clearly, using the alternative interview question to orient participants in the *Non-Victimizer Protocol* to the story protagonist's action – that is, to direct children's attention to the unusual means by which the racer won – had much less of an impact on their judgments than the same condition in the *Victimizer* procedure.

Such predominantly positive emotion attributions also highlight an important feature of the Gender x Condition interaction that emerged in the *Non-Victimizer Protocol*. Much like the results of the *Victimizer* interview, girls were again seen to have comparatively lower *Emotion* scores in the *Alternative Condition* than boys (.24 versus .35, respectively). This time, however, it was not, as before, because they exhibited a greater commitment to moral concerns. Rather, in

the *Non-Victimizer Protocol*, the girls saw nothing wrong with the protagonist's actions, and so, their tendency was to place greater emphasis on the positive emotions associated with the story's outcome – that is, the protagonist's victory. In the end, this suggests that the differences between boys and girls is not one of contrasting moral sensibilities, but of a tendency for girls to be generally more univocal than boys in their answers about others' feelings, moral or otherwise.

Conclusions

As described earlier, the impetus behind much of Study One grew out of two general observations regarding past research on the so-called "happy victimizer phenomenon." First, the majority of work in this area was seen to focus almost exclusively on documenting what are now fairly standard age-differences in the pattern of emotion attributions that younger versus older children make when asked to comment on the morally hazardous actions of others. Although there are certainly speculative accounts concerning how these differences emerge, few studies have actually been designed to test them directly, and none – except the present study – have explored specific cognitive factors that potentially shape (e.g., Arsenio & Lover, 1995) such "happy victimizer" emotion attributions.

The second observation that motivated the present research concerned a structural similarity between the vignettes used as stimulus materials in the typical happy victimizer interview procedure and some of the more central features of Piaget's (1954/1981) description of the "will." Specifically, these vignettes portrayed the "victimizer" as a character who suffered a conflict of opposing desires and the eventual collapse of his or her will. The argument that followed from this observation was that children who were tested with these stories could potentially view the emotional consequences of the victimizer's conflict differently depending on how they, themselves, understood the functioning of the will. In short, it was argued that only

children with an understanding of “autonomous agency” would be in a position to attribute mixed, or multiple, emotions to victimizers for their misdeeds, while children with a more “mechanized” view of human behavior would make predominantly single-valence, outcome-oriented attributions. Importantly, the difference between “mechanical” and “autonomous” conceptions of agency was hypothesized to turn on children’s developing theories of mind, and whether, in particular, they had achieved an interpretive understanding of mental life.

The results from the *Victimizer Protocol*, where this latter claim was tested, largely supported the study’s central hypothesis. Using procedures that closely matched the standard happy victimizer testing conditions (i.e., the *Traditional Condition*), children who were found to possess an interpretive theory of mind were, just as predicted, the most likely to attribute mixed emotions to the victimizer (see Figure 2). On the other hand, participants who fell short of this cognitive milestone – that is, children who were scored as *Non-Interpretive* – tended to offer only single-valence emotions. Finally, although children’s age was correlated with both interpretive status and emotion understanding, it failed to be the strong predictor that past research has assumed it to be. That is, while the present study also found similar age-differences in participants’ happy victimizer attributions, these were shown to turn more directly on children’s growing interpretive understanding of the knowing process.

Such support for the study’s main hypothesis did not, however, come without several important caveats about the happy victimizer phenomenon more generally. These followed, in particular, from the two procedural manipulations that were designed to explore the effects of breaking with the more traditional testing practices. First, with the use of an alternative interview question – one that specifically directed participants’ attention to the victimizer’s actions – it was shown that children’s tendency to provide happy victimizer attributions could be

greatly reduced. That is, even minor changes to the typical protocol, in this case just a few additional words in the interview procedure, were found to generate rather sizeable differences in children's responses.

At least in part, this particular break from the standard findings could be seen as lending support to Keller et al.'s (2003) contention that children often interpret the *Traditional* interview question ("*How does the victimizer feel at the end of the story*") as a request for more neutral, descriptive information about the victimizer's emotional state, rather than a "moral-prescriptive" account of how he or she should be feeling (see, again, Chapter II). Still, Keller et al.'s (2003) more general explanation for the occurrence of the happy victimizer phenomenon fails to fit with the present results. Specifically, Keller and her colleagues argue that children, who tend to offer positive emotion attributions for victimizers, do so because they fail to identify how they, themselves, would feel in similar circumstances. To explore this possibility, they asked children not only the standard question of how the victimizer would feel at the end of the story, but also an alternative question meant to encourage, what they called, "spontaneous identification" (Keller et al., 2003, p. 15) with the victimizer and his or her misdeeds (i.e., "*How would you feel, if you had done that?*"). Not unlike the present study's results, they found that such prompting led to fewer happy victimizer attributions. What Keller et al. overlooked, however, was the fact that their method of prompting directed children to reflect on essentially two issues: 1) how the participant might feel, and 2) how the victimizer had acted. Consequently, either of these two directives could be responsible for the reduction they report in children's positive emotion attributions. The present study helps to resolve this matter. Specifically, the results from the *Alternative* portion of the interview procedure (which used the less ambiguous question: "*How does the victimizer feel about acting like that?*") suggest that this reductive effect is more closely

related to children's notions of agency than, as Keller et al. claim, the result of increased self-other identification.

Beyond these results from the *Alternative Condition*, the second testing modification – the one that introduced a contrasting, *Non-Victimizer*, story condition into the mix of procedures – also produced a set of novel findings. More specifically, the *Non-Victimizer* condition was designed to explore the potential differences in children's emotion attributions between discretionary moral contexts and morally prohibitive contexts. Typically, only more prohibitive contexts, such as the “swing scenario” used in the present *Victimizer Protocol*, have been studied in past research. The results of this comparison between the situational differences in the *Non-Victimizer* and the *Victimizer* conditions suggested that, at least when it comes to discretionary matters, cognitive factors, such as having or not having an interpretive theory of mind, play a relatively minor role in the organization of children's emotional inferences. That is, while participants' *Theory-of-Mind* scores were found to be highly predictive of their *Emotion* understanding in the *Victimizer Protocol*, this was only marginally the case for the *Non-Victimizer* condition.

One possible explanation for this difference between the protocols concerns the fact that moral actions in each context were structured according to different sets of desires and obligations, which, in turn, had a differential impact on the story protagonist's experience of a “conflicted will.” Specifically, in the more prohibitive *Victimizer* scenario, the potential for conflict between opposing desires and the subsequent experience of mixed emotions was much greater than in the *Non-Victimizer* condition. This was because, in the *Victimizer* condition, there were clear-cut moral obligations about not harming others that competed with the protagonist's strongly felt desire to play on the swing. By contrast, in the more discretionary

backdrop of the *Non-Victimizer* story, such moral obligations were absent, or, at least, not as easily perceived. There was little basis, then, for children to offer mixed, or multiple, emotion attributions in these simpler circumstances. Accordingly, the empirical upshot of these differences was that the *Non-Victimizer* scenario placed fewer cognitive demands on participants, and so, children's theory-of-mind competence had only modest predictive value in determining how they would respond. For future research, a more targeted effort to manipulate the levels of internal conflict that story characters experience, perhaps by describing more individualistic qualities or character traits (e.g., one who always helps others), could potentially avoid this current design limitation. As it is, such personal conflict in the present study emerges because explicit societal norms stand against an individual's private desire, and not because an individual somehow deviates from his or her more personally held values or beliefs about him or herself – that is, the kind of person they imagine themselves to be. Addressing this last issue would also become a way to explore some of the recent theoretical connections that have been made between the happy victimizer phenomenon and the development of a “moral self” (e.g., Bergman, 2002, 2004; Blasi, 1995).

Although the present account generally bypasses such matters of the “self,” it seems unlikely that children's conceptions of agency, particularly as they evolve in older children, would not also implicate some kind of “theory of self” (Moshman, 1999). In fact, children's early interpretive achievements have been argued (Chandler & Carpendale, 1998) to constitute an essential part of young people's first insights into selfhood and self-awareness. Still, in terms of the present findings, there are some potential “self-explanations” that do not seem to fit the data. One of these has to do with the role that “self-presentational” concerns (Goffman, 1959; Schlenker, 1980) have in children's emotion attributions: That is, another possible explanation

for the happy victimizer phenomenon, and particularly its observed decline in older children, is that young people are becoming increasingly adept at "impression management," both for themselves and others, and so, better understand the detrimental interpersonal effects that experiencing (or, at least, displaying) positive emotions for doing wrong can have. While participants in the present study generally fell below the age when such concerns have been found to be active (Aloise-Young, 1993; Banerjee & Yuill, 1999; Bennett & Yeeles, 1990), some researchers (e.g., Krettenauer & Eichler, in press) studying the happy victimizer phenomenon in older children and adolescents have begun to argue for the importance of introducing more systematic controls for detecting potential self-preserving biases (e.g., social desirability) in children's responses. According to such an account, children's attributions in the happy victimizer scenario are determined by how actively they hold to the naïve, though socially desirable, belief that they and others are fundamentally good people and would never harm the innocent. At least in regards to the present study, if this alternative "self-presentational" account were true, then a different pattern of results should have emerged in the comparison of the *Victimizer* and *Non-Victimizer Protocols*. That is, the socially desirable response in both cases was a negative emotion attribution (i.e., feeling "bad" for hurting someone, and similarly, for capitalizing on others' pain or misfortune). By and large, however, most children in the *Non-Victimizer* story condition looked past the fact that the target protagonist capitalized on the other racers' misfortune, and attributed him or her with experiencing generally positive emotions.

While dismissing this "self-presentation hypothesis" helps to build the present argument that young people's conceptions of agency underlies the happy victimizer phenomenon, and that, more generally, studying children's developing theories of mind holds much promise for framing future research on this topic, it still relies quite heavily on an untested account linking agency,

and particularly children's views about action, to their understanding of interpretation. That is, if achieving an interpretive understanding of the knowing process is, in fact, part of a broader shift in children's conceptions of human agency, then other cognitive consequences, besides just changes in their views about victimizers' emotional lives, should also be evident. Accordingly, the purpose behind the next study was to explore some of the further consequences of attaining an interpretive theory of mind. In particular, it tackled the question of agency more head-on by studying children's understanding of action motivation and their growing ability to drive a conceptual wedge between others' "desires" and "intentions."

IV. STUDY TWO

Exploring the Building Blocks of Human Agency: Children's

Conceptions of Intention and Desire

Despite its linchpin role in day-to-day self-understanding, most ordinary talk of human agency has never found a workable place in the traditional explanatory frameworks of psychology. Within recent memory, even the slightest whiff of agentiveness was enough to evoke the discipline's signature fear of vitalism (Skinner, 1973) – the very thing that psychology as a science had arisen to defeat. Although such a reflexive aversion toward all matters agentic is now less automatic than was once the case, remnants of this collective phobia continue to persist throughout the field.

Notwithstanding such classical aversions, things are clearly changing (e.g., Martin, Sugarman, & Thompson, 2003). For many, Wittgenstein (1953) is credited as one of the first to give a more contemporary voice to the problems surrounding human agency. He asked, in particular, "What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?" This curious arithmetic, as it is typically understood, yields two kinds of potential answers. The first, historically ceded to philosophy, has to do with all those human activities that are guided by so-called higher-order faculties such as beliefs or intentions. Such purposeful actions, as they are commonly described, are generally construed as agentic in nature and contrasted with less heady, non-agentic behaviors. This second class of behaviors, on the other hand, has traditionally been seen as the proper subject matter for psychology, and is meant to capture all the remaining occurrences or happenings, open to scientific scrutiny, that are "passively" visited upon human bodies.

Although this “subtract and conquer” strategy has very long legs, there is, of late, less consensus about whether such a division of labor is the best way to achieve conceptual clarity concerning more robust forms of agency. Even Wittgenstein (see Hacker, 1996) seemed unsatisfied with this either-or distinction between actions and behaviors, and, in more contemporary circles, there are many (e.g., Frankfurt, 1988; Velleman, 2000) who have argued that expanding the number of categories from just two to at least three is a necessary step toward adequate self-understanding. Alternatively, and occupying the opposite camp, there are those who have tried to reduce the whole of human behavior to just one foundational category – a flatland in which all presumably dated talk of “agency” is relegated to the dustbin of mere superstition (e.g., Churchland, 1984). Anyone in search of better guidance about the proper place of agency in contemporary psychological thought will find little that does not belong to one or the other of these oppositional views.

As it stands, the particular tack to be taken up here has less to do with choosing sides in the foregoing debate than with appropriating some of the language that surrounds it, and, in turn, using it to elaborate a critical shortcoming in the literature on children’s developing theories of mind. In particular, I will argue that, in its push to characterize how children employ notions of belief and desire in their construction of folk psychological laws (e.g., Gopnik & Wellman, 1994), the theories-of-mind enterprise has inadvertently promoted an impoverished conception of agency – one that closely resembles the passive, copy theory of mind earlier described in Chapter II.

Although part of the purpose of this critique will be to deconstruct these truncated assumptions, it is also meant to illustrate other new possibilities for research. One of these – the one making up Study Two – concerns the empirical connections between children’s changing

views of intention and their newfound achievement of an interpretive theory of mind. More specifically, I mean to show that, with the arrival of such an interpretive view, children acquire the conceptual tools needed to understand the central differences between simpler wants, or *desires*, on the one hand, and more fully elaborated plans, or *intentions*, on the other.

The Inherent Shortcomings of Belief-Desire Psychology

It is now a commonplace in theory-of-mind quarters to hear human action described as the simple product of crossing an individual's desires and beliefs. Wellman (1990), for instance, argues that "belief-desire psychology" provides the basic organizing principles for understanding children's early conceptions of mind, and Kim (1996), in his influential text, Philosophy of Mind, has stated quite explicitly that "it seems essential to our concept of action that our bodies are moved in appropriate ways by our wants and beliefs" (p. 8). As it turns out, these are only some of the more recent expressions of a view whose intellectual history stretches at least as far back as the empiricist philosophies of Hobbes (Leviathan, Part 1, Chapter 6) and Hume (A Treatise of Human Nature; see McNaughton, 1988), if not Aristotle (see Smith & Rogers, 2003). In fact, it is precisely the mechanistic conceptual baggage inherent in such narrowly empiricist worldviews (Overton, 1991), and contained in the working tenants of belief-desire psychology, that becomes the main sticking point in the present discussion of agency and children's conceptions of intention. This problem becomes clearer upon consideration of what the separate components of belief and desire are generally seen to bring to the construction of actions.

The Mechanization of Human Action. Among most contemporary theorists of mind, beliefs and desires are considered contrasting mental states (e.g., Astington & Gopnik, 1991) that differ according to the opposing "directions of fit" understood to obtain between the mind and the world. Trading upon language popularized by Searle (1983), *beliefs* are said to "fit" the mind

to the world by reproducing or re-presenting some duplicate of the world's contents inside the mind. *Desires*, by contrast, reverse this relationship and, instead of representing things as they currently stand, work to transform the world in accordance with one's wishes – that is, by aspiring to make the world “fit” with one's own ambitions for it. On this account, it should be clear that beliefs are taken to serve as the quintessential representational state, while desires perform a more motivational function. This fundamental contrast is perhaps most evident in the distinct roles that these mental states are said to play in explanations of human action. That is, desires are typically seen as providing the motive force, or hydraulic push, that drives behavior; while beliefs are all about steerage, and taken as supplying the information that properly channels, or guides, this force (see, e.g., Harris, Johnson, Hutton, Andrews, & Cooke, 1989). Beliefs, here, as McNaughton (1988) has noted, are “motivationally inert; they are merely passive responses to the way the world is” (p. 107); while desires, by contrast, “are active; directed as they are towards obtaining something, they are intrinsically ‘pushy’” (p. 107). No matter how familiar all this may sound, there are, nevertheless, inherent difficulties with treating human beings like hydrostatic machines – the most obvious being that it throws agency out the window.

Laws of Behavior. The problem, then, with belief-desire psychology is that it too easily allows human action to be construed as the mechanized output of what Harré (1982) has called “subpersonal components” in the agent's mind. With these isolated parts taking center stage in action-explanations, the integrative role of the agent or person, to whom any such component beliefs and desires are assumed to belong, is reduced to that of a bit-part and effectively explained away. Beliefs, on this account, are merely pale imitations of reality, whereas desires,

as Dent (1984) has quipped, are like “weighty brick[s] ... [they] fall upon one and impart a certain push to one’s body” (p. 99).

To be clear, this argument should not be taken to mean that the attribution of beliefs and desires, as usual parts of everyday action-explanation, is founded on a mistake. Adults and children alike do quite regularly evoke such mental states in their commonsense accounts of behavior, and no doubt use them to reasonably explain and predict a subset of others’ actions. Rather, the current objection to belief-desire psychology arises primarily when the loosely held notions of “explain” and “predict” employed in our ordinary day-to-day accounting practices are co-opted by the scientific community and re-read as “causal” laws that mechanistically “determine” human action. No doubt, ordinary people are quick to agree that, in important ways, we all act as we do because of the beliefs we hold or the desires we have. They do so, however, without being burdened, as is the case with most members of the scientific community, with empiricist or reductionist ambitions to depopulate the mind of any and all remnants of will or agency. That is, freed from any professional obligations to adhere to Morgan’s Canon, the beliefs and desires of which they (your ordinary child- or person-on-the-street) speak are those closely held beliefs and heartfelt desires of active agents. When all such reference to agentiveness is dropped with extreme prejudice from our professional accounting practices, the result is a “theory-theory” that is only a pale shadow of those real flesh and blood accounts of mental life that such formal theories aim at, but regularly fail to describe. The end result, in fact, begins to take on an uncanny resemblance to the classical “black box” accounts that are so much a part of the professional histories of psychologists.

The overall posture of the current theories-of-mind literature seems to set up the conditions for this inadvertent maneuver. That is, by adopting the widely-held view that beliefs

and desires are analogous to “theoretical constructs” (e.g., Perner, 1991, p. 108) that, in turn, “provide causal explanations” (Gopnik & Wellman, 1994, p. 260, emphasis added) for one’s own and others’ behavior, it is all too easy to slip into making claims such as the following: it is “what actors think—their representation of the world...[that] inevitably determines their actions” (Gopnik & Wellman, 1994, p. 264, emphasis added); or, even worse, into imagining that a simple “practical syllogism ... works if no unusual circumstances intervene” (Perner, 1991, p. 108) to establish an inviolable, psychological law that is adequate enough to capture the essence of human action (e.g., If Homer wants the beer, and thinks it is in the fridge, then he’ll go there). Such a privileging of impersonal, causal language to describe everyday experience has been characterized by Dewey (1989/1925, p. 21) as the very sort of “intellectualism” that, when allowed to proceed unchecked (as it has been in the theories-of-mind literature), wrongly supplants our primary experience of agency as the basis for understanding human action. As a result, we are left, in Campbell’s (1995) words, with a “fragmented ... picture of human action that fits a machine better than an organism” (p. 34).

It is, perhaps, not the job of empirical scientists to attempt to sort out these questions of whether people really are or are not “agents.” Still, in order to adequately describe what children and adults ordinarily suppose happens in the exercise of mental life, researchers must avoid imposing the conceptual commitments of a “causal theory of mind” onto the folk accounts of study participants. Insofar as the theories-of-mind enterprise has allowed this to happen, it is now necessary to re-work the standard practice by which belief-desire psychology is applied to folk conceptions of action. More specifically, it is time for an account that, as Duff (1990) has suggested, sees “persons and actions ... [as] logically basic categories [that] cannot be explained [away] by an analysis which seeks to reduce them to supposedly simpler elements” (p. 130; see

also Taylor, 1958, p. 215). In other words, beliefs and desires need to be re-integrated into a conceptual framework that at least potentially treats human agency as the primary unit of analysis in folk psychological accounts, rather than something to be analyzed away.

The New Frontier of Intentions

Having strenuously objected to any account of children's theories of mind that makes beliefs little more than the "passive responses to the way the world is" (McNaughton, 1988, p. 107) and agents merely "the arena" meant to house such "physiological events" (Velleman, 1993, p. 189), the present deconstructive efforts would end on a dour note were it not also the goal to incorporate these foregoing ideas about agency into a program of empirical research. A good deal has already been said in Chapter II relating children's interpretive achievements to their broader conceptions of agency. Where the rationale for the present research begins, then, is with a page from Wren's (1974) argument that the "core of agency is intention" (p. 33, emphasis added). More specifically, the following study begins to explore how young persons' interpretive competence (as measured by the Droodles procedure described earlier) may be used as a marker for further advances in their growing understanding of others' intentions.

Although not alone in this newfound interest in intentionality (e.g., Malle, Moses, & Baldwin, 2001; Zelazo, Astington, & Olson, 1999), many of those contributing to this new literature appear to have brought along with them much of the "causal" language that previously characterized still earlier work on beliefs and desires. For instance, intentions are commonly said to "cause the actions they represent" (Astington, 2001, p. 88, emphasis added) and to "underlie and cause bodily movements" (Meltzoff, Gopnik, Repacholi, 1999, p. 24, emphasis added). The central problem with these causal accounts of intention is again to be found in what they leave out. That is, nothing about them suggests, as Kenny (1963) once remarked, that

forming “an intention (in the sense of making a decision) is itself a human action” (p. 94; emphasis added). The point here is that intentions are poorly understood if taken to be mere causal mental entities. Rather, intentions fundamentally represent an active process that is initiated by an autonomous agent. Failing to honor this division between entities and activities is perhaps justified within a system that makes mechanical agency the only live option, but this is clearly not the case for those who also take the possibility of autonomous agency seriously. The guiding thesis in this second study, then, went as follows: Children who could be credited with an interpretive theory of mind, and so also holding to an autonomous conception of epistemic agency, would be better situated than their non-interpretive counterparts to recognize the fundamentally active underpinnings of intention. In other words, interpretive children would recognize that intentions constitute a self-initiated, or autonomous, mental activity.

Operationalizing Children’s Understanding of Intentions. To explore this active dimension of intention, Study Two borrowed from Bratman’s (1987) characterization of human beings as fundamentally “planning agents” (p. 2). According to Bratman, we recognize in ourselves, not only the ability to act purposively, but more significantly as being constituted by our capacity to form and execute plans. On this view, then, although some of our actions may be described as fulfilling “present-directed” desires or wants, much of what we do follows from more “future-directed” prior intentions, or plans. In other words, an individual’s act of planful deliberation is said to distinguish prior intentions from simpler desires. As a result, even though both intentions and desires may be seen more generally as motivational attitudes, there is a great deal to suggest that collapsing them into a single category would be a serious mistake (Astington, 2001). Specifically, intentions implicate a kind of autonomous epistemic agency that desires, by themselves, do not. Desires refer only to specific outcomes or end-states of actions (see Malle &

Knobe, 2001) that exist quite apart from, or external to, an agent's deliberations. To say, for instance, that an individual's desires are fulfilled in no way requires reference to his or her planning or deliberative activities. This is not the case, however, when describing an individual's intentions. To illustrate, consider Searle's (1983) account of a "deviant causal chain" (p. 82; see also Astington, 1993; Chisholm, 1976) in which a person's desires are ultimately met, but not in the particular way in which he or she intended or planned.

According to this oft-cited philosophical example, a young man, who wants to kill his rich uncle so that he can collect his inheritance, carelessly drives his car over a pedestrian while thinking about how to enact his malicious scheme. The dead pedestrian, as it happens, is the young man's uncle. Now while it is safe to say that the young man's desire is fulfilled in this unusual (or "deviant") chain of events, it cannot also be said that he followed through with his intentions. This is because intentions are "self-referential" (Searle, 1983) and so, in order to be fulfilled, require that an action, or sequence of actions, occur under the description in which they were originally conceived.

As it turns out, such deviant causal chains provide the necessary test-case to investigate the hypothesis that interpretive and non-interpretive children differ in their conceptions of intention. More specifically, in the present study, children were presented with a more child-friendly variation of two such causal chains: a) one in which a girl story protagonist's overall goal, or desire, was achieved, but not according to her plan (i.e., the traditional version of a deviant causal chain); and b) another where the same girl character initiated all the steps in her plan, but nevertheless failed to obtain her desired goal. Together, then, these contrasting story conditions portrayed the protagonist as meeting the satisfaction requirements for either a) her desires, or b) her intentions, but not both. A recent study by Schult (2002) has successfully used

a similar methodology with early school-aged children, although its purpose was only to compare the age-graded differences of participants' intention understanding. Because this earlier research found that children as young as 4 could accurately comment on others' desires and intentions when no disjunction, or mismatch, in the satisfaction requirements occurred (i.e., when both had been fulfilled, or when both had been left unfulfilled), the present study targeted only those remaining disjunctive instances where clear mismatches between intentions and desires could be seen to arise.

Method

Participants

Sixty-six children (31 boys and 35 girls) between the ages of 5 and 7 years old were recruited from a local parochial school for the study. Of these, 23 were 5-year-olds ($M = 5$ years, 8.6 months, $SD = 2.1$ months, range = 65.4 – 71.3 months); 23 were 6-year-olds ($M = 6$ years, 5.2 months, $SD = 3.4$ months, range = 72.4 – 83.5 months); and, 20 were 7-year-olds ($M = 7$ years, 6.2 months, $SD = 4.5$ months, range = 84.0 – 95.9 months). As with Study One, the majority of the children were again from white, middle-class families. All had volunteered to participate in the study, and their parents or guardians had provided informed consent.

Stimulus Materials and Procedures

Participants' understanding of the interpretive nature of knowing was again measured using the *Theory-of-Mind Protocol* detailed in Study One, and scored using the same procedures. In addition to this protocol, children also completed a *Intention-Desire Mismatch (IDM)* procedure designed to explore their ability to distinguish between intentions and desires. Finally, a subset of 39 children, whose parents had provided special permission to allow additional testing, received the *Peabody Picture Vocabulary Test – 3rd Edition* (Dunn & Dunn, 1997).

These procedures were individually administered in a 20-25 minute interview. The order in which participants received each measure was counterbalanced, as were the trials and interview questions within the two experimental protocols.

Verbal Intelligence. When possible, participants' general verbal competence was assessed using the *Peabody Picture Vocabulary Test – 3rd Edition*, or PPVT-III (Dunn & Dunn, 1997), a standardized measure of receptive language abilities. The PPVT has been found to correlate highly with a wide range of intelligence tests, including the Wechsler Preschool and Primary Scale of Intelligence (Carvajal, Parks, Logan, & Page, 1992), the Weschler Intelligence Scale for Children (Hodapp & Gerken, 1999), and the verbal subscale of the Stanford-Binet (Hodapp, 1993). It was used in the present study as both a measure for children's receptive language skills and, more generally, as a way to estimate their "mental age."

Intention-Desire Mismatch Protocol. Two illustrated story conditions (see Appendix D) were designed to gauge children's understanding of the different satisfaction criteria that apply for intentions and desires. In the *Inadvertent Success Condition*, which employed all the elements of a classical deviant causal chain, the target protagonist, Alice (from Lewis Carroll's popular book series), is shown devising a plan to catch the white rabbit – a short, two-step procedure of putting carrots in a cage and hiding behind a tree. Before enacting any part of her plan, however, the rabbit serendipitously hops into the cage, thereby satisfying Alice's desire, though not as she had originally planned. The story ends with Alice taking the rabbit home.

The second, *Foiled Plan Condition* was designed to create a similar mismatch between the story protagonist's intentions and desires, but reversed the critical elements of the *Inadvertent Success* story. That is, in the counterpart *Foiled Plan* vignette, Alice, who is now shown trying to catch a dormouse, actually follows-through with the steps of her plan (to put cheese in a cage

and hide behind a tree), but then fails to lure the mouse into her trap. The final frame, in this contrasting scenario, shows Alice returning home empty-handed, despite carefully doing everything she had intended. Arguably, this *Foiled Plan Condition* falls short of being a “true” deviant causal chain (i.e., there is nothing especially “deviant” about an unsuccessful plan), but it was nevertheless included to provide symmetry in the study design and to allow comparisons to other similar research (e.g., Schult, 2002).

After each condition, the storybook was closed and participants were asked to repeat the central elements of the vignette, particularly what the target protagonist was trying to accomplish (i.e., Alice’s desire) and the steps she imagined were necessary to meet with success (her intention). If children failed these control questions the first time (which was rare), they were re-told the story and again asked to describe Alice’s goals and plans (“*What did Alice want to catch?*” and “*How is she planning to do that?*”). Finally, the two test questions that immediately followed these controls were: 1) *Did Alice get what she wanted?* and 2) *Did Alice do what she planned?* Children were expected to respond with a simple “yes” or “no” to each question, however, in order to be scored as passing the condition, both had to be answered correctly. A single point was awarded for each correct response-pair.

Results

The pattern of children’s *Theory-of-Mind* scores closely paralleled the findings of Study One. Specifically, participants’ Age and level of interpretive understanding were again highly correlated [$r(65)=.55$, $p<.0005$ (one-tailed)]. Similar age-graded results were also found, as expected, for children’s PPVT scores [$r(38)=.60$, $p<.0005$ (one-tailed)] and their performance on the *Intention-Desire Mismatch (IDM) Protocol* [$r(65)=.30$, $p=.008$ (one-tailed)].

Similar to previous work on children's understanding of intentions (Schult, 2002), only the oldest participants (7-year-olds, in this case) were consistently able to tell in the *IDM* procedure whether the story protagonist's intentions or desires had been satisfied. Beyond such age-graded results, however, was the fact that children's *Theory-of-Mind* scores were also predictive of success on the *IDM* interview [$r(65)=.44$, $p<.0005$ (one-tailed)]. Importantly, this relationship held up when controlling for participants' Age [$r(63)=.34$, $p=.003$ (one-tailed)], and, although computed for a smaller subset of children, when Age was substituted with participants' standardized PPVT *Age-Equivalents* [$r(36)=.32$, $p=.025$ (one-tailed)]. Finally, even when using a more conservative analysis strategy of controlling for both Age and PPVT raw scores – a test which essentially factors out children's Age in two different ways – the correlation between *Theory-of-Mind* and *IDM* scores was still found to be marginally significant at trend levels [$r(35)=.25$, $p=.07$ (one-tailed)]. It is likely that this analysis would have achieved standard significance levels with a larger sample size.

Interestingly, Schult's (2002) research reports that children performed differently across the two disjunctive conditions. She found, in particular, that 5-year-olds matched the performance-levels of 7-year-olds and adults in her equivalent of the *Inadvertent Success Condition* (or traditional deviant causal chain story), but not in the *Foiled Plan Condition*, where they did more poorly. Schult speculated that the relative ease that children experienced with the former condition – that is, the standard deviant causal chain – was the result of more salient cues regarding the story protagonist's actions to satisfy her intentions. Specifically, because the target protagonist in this deviant causal chain was shown to take no action at all, Schult argued that it was easier for children to respond negatively to the question of whether the story character “did what she planned to do” (after all, she did nothing). While such speculation seems reasonable,

the theoretical framework for the present research, coupled with the findings from Study One, would predict the opposite pattern of results. That is, providing children with more details about an individual's agency (such as simply directing children's attention to a character's action as was done in the *Alternative Condition* of Study One) should have the general effect of enhancing their ability to accurately comment on others' intentions. In the context of the present study, this means that children should have done better overall in the *Foiled Plan Condition*, where the story protagonist was explicitly shown to enact the various steps of her unsuccessful plan, rather than (as Schult found) in the *Inadvertent Success Condition*, which should actually be the harder test case for children to pass. As the preceding also suggests, this means that children's *Theory-of-Mind* competence should be a weaker predictor of success in the easier *Foiled Plan Condition*, while a stronger one for the *Inadvertent Success* (or traditional deviant causal chain) *Condition*.

Although initially no such differential predictions were made regarding these two conditions (i.e., they were treated as roughly equivalent in the first round of analyses), follow-up analyses found support for the preceding claims. As the following graph (Figure 5) comparing children's success in each condition shows, performance in the *Foiled Plan Condition* was roughly the same for all participants, regardless of their *Theory-of-Mind* competence [$\chi^2(2, N=66)=1.37, p>.05, N.S.$]. By contrast, children's performance in the *Inadvertent Success Condition* varied systematically according to their *Theory-of-Mind* understanding [$\chi^2(2, N=66)=15.16, p=.001$]. In the end, it appears that the correlation between participants' *Theory-of-Mind* and overall *IDM* scores was driven largely by children's performance in the *Inadvertent Success* story condition – that is, the condition with the most authentic rendering of a “deviant” causal chain.

Intention-Desire Mismatch Protocol

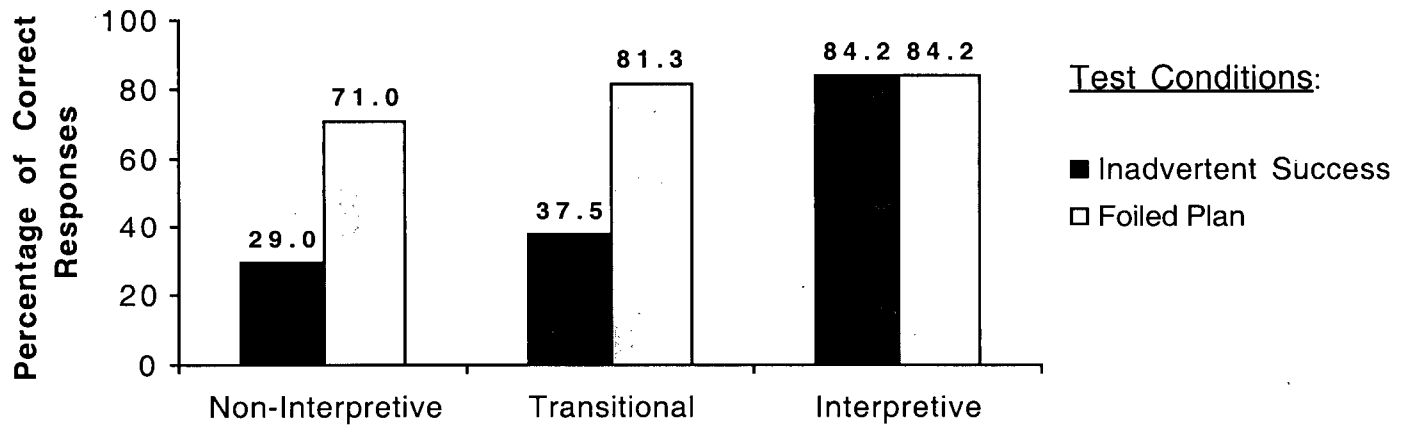


Figure 5: Participants' Success in the Separate Conditions of the *IDM Protocol*

While these differences between Schult's (2002) earlier findings and the present results are perhaps surprising, it is important to note that procedurally the two studies were not identical, and so, some caution is required in making comparisons. There are, for instance, differences in the content of the stories each study used, as well as differently worded test questions.¹ Finally, and perhaps the most critical difference of all, Schult introduced her test stories with a "warm-up" phase in which participants discussed the meaning of plans. This could have easily made matters of planning very salient in the following test procedure, and ultimately, provide participants with a task-specific advantage in the story condition where no plans were executed. The upshot here is that more careful follow-up work is necessary before any definitive conclusions can be drawn about the differing patterns of results between these studies.

¹ The strategies for analyzing the two data sets were also different, although similar results to the present ones were still found even when using Schult's analysis of variance method. Because the present data were treated as dichotomous in nature, such a method was judged inappropriate.

Conclusions

Although certainly one of the purposes behind Study Two was to explore potential differences between interpretive and non-interpretive children's understanding of intentions, the more general impetus for this research was to find converging evidence for Study One's conclusion that an interpretive theory of mind reflects a broader shift in children's views about human agency – one that is expressed not only in their conceptions of others' emotional lives, but also in other areas, such as how they construe mental states like intentions and desires. The present results generally support the notion of such a broader shift. In particular, children's overall understanding of intentions – again, what Wren (1974) has argued to be the “core” of human agency – was strongly related to their interpretive status. What is more, even the differences that emerged between the two test conditions of the *IDM Protocol* fit with the present theoretical account. That is, while both conditions involved some sort of mismatch between an individual's intentions and desires, only the *Inadvertent Success Condition*, based on the standard deviant causal chain, highlighted how intentions carry an added sense of ownership over one's actions that is absent in the case of simpler desires. On this account, when individuals achieve their goals (i.e., fulfill their desires) without also enacting their intentions, they are generally not entitled to claim such ownership – they must credit someone or something else for their unanticipated good fortune. Still, to appreciate such a nuance – that is, to recognize that the circumstances underlying a traditional deviant causal chain are actually devoid of agentiveness – requires that one hold a well-developed understanding of agency. Not surprisingly, then, it was in the *Inadvertent Success* story condition where children's interpretive abilities gained them the greatest purchase in differentiating intentions from desires.

Taken as a whole, these findings from Studies One and Two lend important support to the present claims regarding children's conceptions of agency and the emergence of an interpretive theory of mind. Still, there are potentially simpler explanations – ones that make no reference to human agency – that might also account for these results. In particular, a common dimension shared by both studies is that the procedures they employ require children to reason about two terms of a problem simultaneously – whether they be mental states like intentions and desires in Study Two, or emotions like happiness and sadness in Study One. One reductionistic reading of these findings, then, is that success on any of the present tasks could be little more than an indication that children are capable of “holding two things in mind at once.” Exploring this particular issue was the purpose of pilot work described in the next chapter.

V. STUDY THREE

Agency or Otherwise? Exploring Alternative Explanations of an

Interpretive Theory of Mind

Although the results from Studies One and Two are framed as evidence for a broader shift in 5- to 7-year-olds' conceptions of agency, there remain potentially simpler explanations that might also account for children's performance on the experimental procedures used in these studies. In particular, a more generic ability to coordinate multiple features of a single problem or task is clearly an important component skill for success on the various measurement tools employed in Studies One and Two. For instance, the criteria that children needed to meet in order to be credited with an interpretive theory of mind – the central marker in these studies for children's understanding of "autonomous agency" – involved recognizing "that one and the same object or event can sometimes afford multiple meanings" (Carpendale & Chandler, 1996, p. 1687), or that "two views can be properly said to be *different* views of *one and the same* thing" (Lalonde & Chandler, 2002, p. 166). In the "Doodles" procedure, more particularly, this meant that children had to provide at least two different descriptions of the ambiguous stimulus objects to which they were exposed. Similarly, in the *Victimizer* and *Non-Victimizer Protocols* of Study One, children's level of emotion understanding was assessed by whether or not they recognized that the target protagonist would experience multi-valenced, or mixed, emotional reactions to a situation. This required that children respond to the test questions by again providing two different descriptions, this time, dealing with particular emotional terms, such as feeling happy and sad at the same time. Finally, Study Two's *Intention-Desire Mismatch Protocol* required that children differentiate between two mental states and respond according to whether an individual's "wants" or "plans" had been fulfilled. All of these procedures, in other words, share

the common structural, or design, characteristic of requiring children to reason simultaneously about two aspects of a given stimulus event. Any systematic failures or successes on these procedures could, then, potentially be traced to an inability to coordinate multiple features of anything whatsoever, rather than because these measures tapped a common, but deeper, conceptual structure responsible for shaping children's views on human agency.

Obviously, such a reductive reading presents a critical challenge to my claims that the present research makes manifest various related aspects of children's working conceptions of agency. The pilot work detailed in this chapter was begun as an attempt to weigh these two competing explanations – or, what I will call, the “simultaneity hypothesis” and the “agency hypothesis.” More specifically, this work paired the key *Theory-of-Mind*, or “Doodles,” *Protocol* from Studies One and Two with a new measure that, unlike the others, was chosen because it was conceptually unrelated to matters of agency, but still shared the same structural characteristic of requiring children to coordinate multiple parts or aspects of an individual problem. I reasoned that there would be little basis for accepting the reductionistic implications of the “simultaneity account” if performance on the Doodles task could be shown to vary independently of success or failure on this new measure. All that was needed to sort out these competing claims, then, was to find an appropriate measure – one, in other words, that did not appear to depend upon an appreciation of human agency.

Measuring Simultaneity. As it turns out, in addition to the theoretical work on agency or the “will” already described in Chapter II, Piaget and his colleagues also explored children's understanding of simultaneity. Much of what Piaget had to say about this topic can be found in his work with Bärbel Inhelder involving their classic studies on the seriation of length (Inhelder & Piaget, 1969; Inhelder, Sinclair, & Bovet, 1974, p. 192). Together, they observed three

distinct developmental levels, or “stages,” in the strategies that children used when attempting to arrange a series of rods from shortest to tallest: 1) “first the markers [objects] are separated into groups of two or three (one short, one long, etc.), each seriated with itself but incapable of being coordinated into a single series;” 2) “next, a construction by empirical groping in which the child keeps rearranging the order until he finally recognizes he has it right;” and, finally 3) “a systematic method consist[ing] in seeking first the smallest element, then the smallest of those left over, and so on” (Inhelder & Piaget, 1969, pp. 101). According to Piaget and Inhelder, only this final “operational” method or strategy was seen to demonstrate a “true” understanding of simultaneity. More specifically, it was only at this level, as they described, that “a given element E [was] understood in advance to be simultaneously larger than the preceding elements ($E > D, C, B, A$) and smaller than the following elements ($E < F, G$, etc.), which [they argued was] a form of reversibility by reciprocity” (Inhelder & Piaget, 1969, pp. 101-102).

Elkind (1964) describes how these three strategies led to the development of several different procedures for measuring children’s seriation abilities. The first of these was a straightforward *construction task* which simply required participants to “make a staircase” by placing differently sized blocks in order from smallest to largest. The second, *insertion task*, involved a slightly more complicated procedure of “fixing the staircase” by inserting a second set of blocks into an otherwise already completed series. Finally, the third, *blind construction task*, was designed to be the most demanding of all and used a cardboard screen to block children’s view, and thereby limit visual feedback, during the construction of the series. Here children were required to pass the blocks, one by one, to the experimenter, who then assembled them behind the screen in the order they were received. Of these three tasks, only the third “blind” procedure, which removed all visual feedback and the opportunity for “empirical groping” (i.e.,

trial-and-error placement of the blocks), ensured that children's performance was based on an understanding of simultaneity. Nevertheless, all three tasks were included in the seriation battery used in the present study.

Method

Participants

To begin piloting these seriation measures, 31 children (16 boys and 15 girls) were recruited from four private elementary schools and a university-based after-school program. Of these, 9 were 4-year-olds ($M = 4$ years 6 months, $SD = 3.85$ months), 12 were 5-year-olds ($M = 5$ years, 6 months, $SD = 2.77$ months), and 10 were 6-year-olds ($M = 6$ years, 5.5 months, $SD = 3.33$ months). The majority of the children in the sample were from white, middle-class families.

Seriation Battery

The seriation procedure consisted of three separate tasks, completed in the following fixed order: *construction task*, *insertion task*, and *blind construction task*. A single point was awarded for the successful completion of each individual task, allowing children to score a total of 3 points on the entire battery. All the tasks were embedded in an interactive story in which children participated. The story began with the character, Alice (again, from the *Wonderland* series), trapped in a rabbit hole. After listening to Alice's dilemma, participants were given eight wooden blocks of increasing size (each was .5 inches longer than the other) and asked to construct a staircase for Alice to use to climb out of the hole. Children were instructed to begin with the smallest and proceed to the biggest of the blocks.

The second, *insertion task* was presented as a new dilemma for Alice to solve. In this portion of the story, children were shown that Alice's staircase had shaken apart before she could

escape. They were asked to help her fix it by inserting five additional blocks into the series. The experimenter demonstrated with one of the blocks and the participants inserted the remaining four.

Finally, children completed the *blind construction task*. For this phase of the procedure, the initial eight blocks were spread out randomly and a cardboard screen was placed between the experimenter and the participant. Children were instructed to make another staircase, just as before, but this time to hand the blocks in order to the experimenter who assembled them behind the screen. Participants were reminded to begin with the smallest block and proceed with progressively larger blocks. They were only allowed to see their work once all the blocks were assembled.

Results

As Table 1 shows, roughly half of the sample successfully passed all three parts of the *Seriation Battery*, and the mean score was relatively high: 2.03 out of 3 ($SD=1.14$). As expected, children's performance was strongly age-graded [$r(30)=.76$, $p<.0005$ (one-tailed)], with all but one 4-year-old failing the procedure, while the opposite pattern held for 6-year-olds.

Seriation Battery		
	Fail	Pass
Age 4-year-olds	8	1
5-year-olds	7	5
6-year-olds	1	9
Total	16	15

Table 1: Children's Seriation Performance

Further analysis of children's performance on the separate parts of the procedure suggested that the *blind construction task* was the most difficult. Only 58% of participants passed this task as compared to the 71% and 74% of children who passed the *construction* and *insertion tasks*. The relative ease that children had with the *insertion task* proved inconsistent

with earlier research, but was likely due to methodological differences. In particular, in the present study, children were shown how to complete the insertion of the initial block, whereas, in previous work, such a demonstration was absent. In fact, Elkind (1964) describes children's confusion with earlier versions of this task as one of bafflement: "as if [the child] had just completed a jigsaw puzzle and were given additional pieces to put within the finished picture" (p. 289).

Participants' overall performance on the *Theory-of-Mind Protocol* was relatively poor, with the mean score only 1.84 out of 4 ($SD=1.37$). Nearly two-thirds (see Table 2) of the sample was coded as *Non-Interpretive*, suggesting that this procedure posed significant difficulties for the children. Still, just as before, success on the protocol was closely related to participants' Age [$r(30)=.44$, $p=.007$ (one-tailed)].

		Level of Interpretive Understanding		
Age		Non-Interpretive	Transitional	Fully Interpretive
	4-year-olds	7	2	0
	5-year-olds	8	3	1
	6-year-olds	5	2	3
	Total	20	7	4

Table 2: Children's Theory-of-Mind Performance

Finally, children's *Theory-of-Mind* and *Seriation* scores were found to be moderately correlated [$r(30) = .37$, $p=.021$ (one-tailed)]. Critical to rejecting the "simultaneity hypothesis," however, was the fact that this correlation was reduced to nearly zero when the common factor of Age was controlled [$r(28)=.06$, $p=.38$ (one-tailed), N.S.]. A Chi-square analysis comparing children's *Theory-of-Mind* performance to their success on even just the most difficult, *blind construction task* showed similar null results [$\chi^2(2, N=31)=4.07$, $p=.13$, N.S.].

Conclusions

Despite common design characteristics – that is, a shared requirement that children have some understanding of simultaneity – no empirical connections between the *Interpretation* and *Seriation* procedures in Study Three were found. While such null results are typically nothing to brag about, in the present circumstances, these findings actually help to rule out a simpler, alternative to my “agency hypothesis” for Studies One and Two. In other words, these pilot data suggest that the connections found earlier between children’s interpretive understanding, their reasoning about “happy victimizers,” and the ability to distinguish between intentions and simpler desires are all unlikely to be the result of some simpler cognitive process, such as “being able to hold two things in mind at once.” Rather, I would argue, there is a credible basis in the present research for accepting the account that 5- to 7-year-old children are undergoing a major conceptual re-structuring in their views about human agency – one that extends broadly to their reasoning about others’ beliefs, emotions, and intentions.

VI. GENERAL DISCUSSION

Limitations and Future Directions

Although the work reported here initially began as an effort to better understand findings in the “happy victimizer” literature, the hypothesis that it generated regarding children’s conceptions of agency proved testable in areas beyond the study of young persons’ reasoning about moral emotions. In particular, Studies Two and Three, although not directly about matters of victimization, shared with Study One a common focus on children’s developing appreciation of the place of agency in orchestrating their understanding of mental life. The thesis, then, that emerged from this ongoing program of research is that sometime in their earliest school years (usually between 5 and 7) young people acquire a general “agentive” framework that allows them to join their older counterparts in the common conviction that their own and others’ actions arise out of plans or intentions owned and authored by willful human agents. The research presented here is not understood to firmly establish this bold claim, but it does go some distance toward demonstrating that emerging ideas about agency provide a plausible connection between such diverse content as children’s notions of interpretation and intention understanding.

To quickly recap, then, the present findings, Study One demonstrated that only children holding to an interpretive understanding of the knowing process – a marker here for children’s notions of “epistemic agency” – went on to reliably attribute morally ambivalent feelings to story characters who were portrayed as victimizers. While framed in this first study as a cognitive constraint underlying children’s reasoning about emotions, such an interpretive theory of mind was argued to have wider application for exploring matters of agency. In support of this claim, Study Two broadened the focus of Study One’s research questions by examining the relation between young people’s conceptions of interpretation and intention. This second study provided

important converging evidence that children who had achieved an interpretive understanding of mental life also conceived of others' agentiveness in a new light – one that allowed them to see more clearly the differences between simpler desires and more elaborate intentions. Finally, Study Three represented an attempt to rule out potential reductionistic explanations of the present “agency account.” In particular, a successful effort was made to distinguish the simpler ability to simultaneously hold any two things in mind, from the later-arriving agentive abilities measured in Studies One and Two. Because each of these separate studies already ends with its own set of general conclusions, the special focus of these final remarks will be on how the “agency account” I am advocating might be used to better guide my continuing program of research.

Future Directions. The findings from Study One show that the happy victimizer phenomenon, while relatively robust under some measurement situations, is more or less absent in others. This was shown using two experimental manipulations – one that made minor wording changes to the otherwise “standard” test question, and the other that introduced a new set of story conditions involving discretionary moral matters. The upshot of the latter manipulation, in particular, was that, although situational constraints in prohibitive moral contexts served as a prompt for conflicting or mixed emotional responses, more discretionary moral contexts did not. The open-question prompted by this finding concerns those conflicts of will that arise, not between normative values and personal desires, but between more self-selected values and one's immediate impulses. What, for instance, would young children understand about the emotional reactions of a committed health food advocate who caves to her craving for McDonald's french fries, or the monk who breaks his vows of silence to speak with a close friend? Obviously, broader society places no restrictions on such common activities, but

an individual, for personal or prudential reasons, may choose to do so. Matters of the will, almost by definition, suggest such individuality – of sticking to one’s personal principles, of resisting the “done thing,” of being “true” to one’s self. While contemporary research dealing with children’s understanding of personal choice (e.g., Nucci, 1981, 1996) and the strategies involved in delaying gratification or resisting temptation (e.g., Mischel & Mischel, 1983; Mischel, Shoda, & Rodriguez, 1992) touch on some of these aspects of personal discretion and the will, there is little on children’s affective or emotional understanding of such matters, and certainly nothing yet linking them to a broad-based shift in children’s conceptions of human agency. A useful first step, then, in expanding the current program of research would be to further explore the potential relations between such discretionary matters and the interpretive measures employed here.

Another critical dimension of Study One needing further investigation concerns the prospect that other cognitive processes, besides an understanding of agency, may conceivably account for some, or even all, of children’s performance on the particular test procedures adopted in this suite of studies. One such alternative explanation – the “simultaneity hypothesis” – was explored in Study Three, but others clearly remain. Of these, one of the more compelling follows from the view that, at a certain developmental moment (but not before), children become capable of simultaneously accessing two different “modes of thought” or reasoning systems (Nisbett, Peng, Choi, & Norenzayan, 2001; Norenzayan, Smith, Kim, & Nisbett, 2002). That is, children who succeed in offering mixed emotions in the happy victimizer task may be drawing on some more “holistic,” or relational, reasoning strategies that strike a balance between the various situational factors presented – for instance, the victim’s “pain” and the victimizer’s “gain.” By contrast, other less integrative, or “analytic,” reasoning strategies may lead children

to focus on only one of these aspects, and therefore, offer only single-valence emotions. Recent research (e.g., Choi, Nisbett, & Norenzayan, 1999) on these two systems of reasoning has begun to show that systematic cultural variations exist in how accessible they are to individuals, leading members of Western cultural groups to reason more analytically and East Asian groups to think more holistically about a wide range of issues. Although there are few such cultural comparisons of the happy victimizer phenomenon, one study (Kim, 1993, cited in Arsenio & Lover, 1995) with a Korean sample of 4- to 10-year-olds has found comparable results to those of Arsenio and his colleagues (Arsenio & Kramer, 1992) in North America. Still, this work was limited by the fact that participants did not receive a follow-up prompt (i.e., "how else might the victimizer feel?") to the standard test question. While it is perhaps premature to launch a full-blown program of cross-cultural research, the ethnic and immigrant groups available more locally in the Vancouver metropolitan area are a reasonable place to begin exploring potential cultural variations in the happy victimizer phenomenon.

Finally, and related to this matter of cultural variation, "agency" has many meanings, and may even be viewed in more or less holistic or analytic ways. Certainly, the notion of "autonomous agency" used throughout this thesis to describe the shift in children's reasoning once they have achieved an interpretive theory of mind suggests a potential Western bias. This terminology, however, was chosen only to provide a clearer contrast to the more "mechanical" views of agency available to non-interpretive children. In fact, Bandura (1989) has suggested that a more sophisticated view of agency beyond either of these mechanical or autonomous forms is "interactive agency" – a kind of synthesis of the previous two. This raises the question of whether children's understanding of autonomous agency is the last transition of its kind in the developmental course, or if other more mature, and perhaps "interactive," forms may also exist.

Here, research on adolescents' views of personhood and "self-continuity" (Chandler, Lalonde, Sokol, & Hallett, 2003) suggests that an understanding of agency undergoes further developments well beyond middle childhood.

Closing Remarks – A Return to the Beginning. Psychology, as suggested in the opening pages of this thesis, may well have gotten its good start by choosing to collectively turn a blind eye toward matters of the "will," and, for certain scientific purposes, may have been wise to do so. The idea never really caught on, however, with the less tough-minded general public, and if it is the views and self-understanding of such rank-and-file persons that happens to be the focus of one's research (as it is with mine), then it is almost certainly too "catholic" to go on refusing to speak of will or agency while attempting to bring their "folk" or "common sense" psychology to light. All of us, professional psychologists included, could hardly accomplish our usual business without constant reference to our own and others' plans and intentions and interpretive acts. What we know much too little about is how young persons come, in the course of their early development, to join us in sharing this agentive framework. What my work suggests is that there is an early time before they ordinarily subscribe to such notions, and because of this they behave and reason in otherwise strange and childish ways. They can't tell an intention from a desire, they imagine an emotional flatland where everyone feels only one way at a time, and they don't appreciate that we all make an active (interpretive) contribution to our own knowing process. Working out the possible relations between these and other superficially separate matters has been, and will into the future be, the focus of this program of research.

References

- Aloise-Young, P. A. (1993). The development of self-presentation: Self-promotion in 6- to 10-year-old children. *Social Cognition, 11*, 201-222.
- Arsenio, W. (1988). Children's conceptions of the situational affective consequences of sociomoral events. *Child Development, 59*, 1611-22.
- Arsenio, W. A., & Kramer, R. (1992). Victimiziers and their victims: children's conceptions of the mixed emotional consequences of moral transgressions. *Child Development, 63*, 915-927.
- Arsenio, W., & Lover, A. (1995). Children's conceptions of sociomoral affect: Happy victimizers, mixed emotions, and other expectancies. In M. Killen & D. Hart (Eds.), *Morality in everyday life: Developmental perspectives* (pp. 87-128). Cambridge, UK: Cambridge University Press.
- Astington, J. W. (1993). *The child's discovery of the mind*. Cambridge, MA: Harvard University Press.
- Astington, J. W. (2001). The paradox of intention: Assessing children's metarepresentational understanding. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Understanding intentions and intentionality: Foundations of social cognition* (pp. 85-103). Cambridge, MA: MIT Press.
- Astington, J. W., & Gopnik, A. (1991). Developing understanding of desire and intention. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development, and simulation of everyday mindreading* (pp. 39-50). Oxford: Basil Blackwell.
- Australia wins first ever gold. (2002). Retrieved September 7, 2004, from http://news.bbc.co.uk/winterolympics2002/hi/english/skating/newsid_1825000/1825339.

- Bandura, A. (1969). Social learning of moral judgments. *Journal of Personality and Social Psychology, 11*, 275-279.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*, 1175-1184.
- Bandura, A. (1990). Mechanisms of moral disengagement. In W. Reich (Ed.), *Origins of terrorism: Psychologies, ideologies, theologies, states of mind* (pp. 161-191). Cambridge: Cambridge University Press.
- Bandura, A. (1991). Social cognitive theory of moral thought and action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development: Vol. 1 Theory* (pp. 45-104). Hillsdale, NJ: Lawrence Erlbaum.
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3*, 193-209.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26.
- Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education, 31*, 101-119.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*, 364-374.

- Bandura, A., & McDonald, F. J. (1963). The influence of social reinforcement and the behavior of models in shaping children's moral judgments. *Journal of Abnormal and Social Psychology, 67*, 274-281.
- Banerjee, R., & Yuill, N. (1999). Children's explanations for self-presentational behaviour. *European Journal of Social Psychology, 29*, 105-111.
- Barden, R. C., Zelko, F. A., Duncan, S. W., & Masters, J. C. (1980). Children's consensual knowledge about the experiential determinants of emotion. *Journal of Personality and Social Psychology, 39*, 968-976.
- Bennett, M., & Galpert, L. (1993). Children's understanding of multiple desires. *International Journal of Behavioral Development, 16*, 15-34.
- Bennett, M., & Hiscock, J. (1993). Children's understanding of conflicting emotions: A training study. *Journal of Genetic Psychology, 154*, 515-525.
- Bennett, M., & Yeeles, C. (1990). Children's understanding of the self-presentational strategies of ingratiation and self-promotion. *European Journal of Social Psychology, 20*, 455-461.
- Bergman, R. (2002). Why be moral? A conceptual model from developmental psychology. *Human Development, 45*, 104-124.
- Bergman, R. (2004). Identity as motivation: Toward a theory of the moral self. In D. K. Lapsley & D. Narvaez (Eds.), *Moral development, self, and identity* (pp. 21-46). Mahwah, NJ: Lawrence Erlbaum Associates.
- Blasi, A. (1984). Autonomie im gehorsam: Die entwicklung des distanzierungsvermögens im sozialisierten handeln (Autonomy in obedience: The development of distancing in socialized action). In W. Edelstein & J. Habermas (Eds.), *Soziale interaction und soziales verstehen* (pp. 300-347). Frankfurt an Main: Suhrkamp.

- 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 & Bacon.
- Bratman, M. E. (1987). *Intention, plans, and practical reason*. Cambridge: Harvard University Press.
- Bratman, M. E. (1997). Responsibility and planning. *Journal of Ethics*, 1, 27-43.
- Campbell, J. (1995). *Understanding John Dewey: Nature and cooperative intelligence*. Chicago, IL: Open Court.
- Carroll, J. J. & Steward, M. S. (1984). The role of cognitive development in children's understanding of their own feelings. *Child Development*, 55, 1486-1492.
- Carvajal, H. H., Parks, J. P., Logan, R. A., & Page, G. L. (1992). Comparisons of the IQ and vocabulary scores on Wechsler Preschool and Primary Scale of Intelligence – Revised and Peabody Picture Vocabulary Test – Revised. *Psychology in the Schools*, 29, 22-24.
- 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. (2001). Perspective taking in the aftermath of theory-theory and the collapse of the social role-taking literature. In A. Tryphon & J. Voneche (Eds.), *Working with Piaget: Essays in honour of Barbel Inhelder* (pp. 39-63). Hove, UK: Psychology 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. J. Sternberg (Eds.), *Self-awareness: Its nature and development* (pp. 148-190). New York: The Guilford Press.

- Chandler, M. J., & Lalonde, C. (1996). Shifting from 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, IL: University of Chicago Press.
- Chandler, M. J., Lalonde, C. E., Sokol, B. W., & Hallett, D. (2003). Personal persistence, identity development and suicide. *Monographs of the Society for Research in Child Development, Vol. 68, No. 2, Serial No. 273.*
- Chandler, M. J., & Sokol, B. W. (1999). Representation once removed: Children's developing conceptions of representational life. In I Sigel's (Ed.), *Development of mental representation: Theories and applications* (pp. 201-230). Mahwah, NJ: Erlbaum.
- Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development, 60*, 1263-1277.
- Chisholm, R. M. (1976). *Person and object: A metaphysical study*. London: Allen & Unwin.
- Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin, 125*, 47-63.
- Churchland, P. M. (1984). *Matter and consciousness*. Cambridge, MA: MIT Press.
- Colby, A., & Damon, W. (1992). *Some do care: Contemporary lives of moral commitment*. New York: Free Press.
- Colby, A., & Damon, W. (1993). The uniting of self and morality in the development of extraordinary moral commitment. In G. G. Noam & T. W. Wren (Eds.), *The moral self*. Cambridge, MA: MIT Press.

- Damon, W. (1984). Self-understanding and moral development from childhood to adolescence. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Morality, moral behavior, and moral development* (pp. 109-127). New York: John Wiley & Sons.
- Davidson, P., & Youniss, J. (1995). Moral development and social construction. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Moral development: An introduction* (pp. 289-310). Boston: Allyn & Bacon.
- Denham, S. A. (1998). *Emotional development in young children*. New York: The Guilford Press.
- Dent, N. J. H. (1984). *The moral psychology of the virtues*. Cambridge: Cambridge University Press.
- Dewey, J. (1989). *Experience and nature*. La Salle, IL: Open Court. (Original work published 1925)
- Donaldson, S. K., & Westerman, M. A. (1986). Development of children's understanding of ambivalence and causal theories of emotion. *Developmental Psychology*, 22, 655-662.
- Duff, R. A. (1990). *Intention, agency, and criminal liability: Philosophy of action and the criminal law*. Oxford: Basil Blackwell.
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test – Third edition: Form A*. Circle Pines, MN: American Guidance Service.
- Eisenberg, N. (1986). *Altruistic emotion, cognition, and behavior*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Eisenberg, N., Shea, C. L., Carlo, G., & Knight, G. P. (1991). Empathy-related responding and cognition: A "chicken and the egg" dilemma. In W. M. Kurtines & J. L. Gewirtz (Eds.),

- Handbook of moral behavior and development (Vol. 2: Research)* (pp. 63-88). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Elkind, D. (1964). Discrimination, seriation, and numeration of size and dimensional differences in young children: Piaget replication study VI. *Journal of Genetic Psychology*, 104, 275-296.
- Frankfurt, H. G. (1988). *The importance of what we care about: Philosophical essays*. Cambridge, UK: Cambridge University Press.
- Frankfurt, H. G. (1999). *Necessity, volition, and love*. Cambridge: Cambridge University Press.
- Gnepp, J., & Klayman, J. (1992). Recognition of uncertainty in emotional inferences: Reasoning about emotionally equivocal situations. *Developmental Psychology*, 28, 145-158.
- Gnepp, J., McKee, E., & Domanic, J. A. (1987). Children's use of situational information to infer emotion: Understanding emotionally equivocal situations. *Developmental Psychology*, 23, 114-123.
- Goffman, E. (1959). *The presentation of self in everyday life*. New York: Doubleday.
- Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257-293). New York: Cambridge University Press.
- Graham, S. (1988). Children's developing understanding of the motivational role of affect: An attributional analysis. *Cognitive Development*, 3, 71-88.
- Graham, S., & Weiner, B. (1986). From an attributional theory of emotion to developmental psychology: A round-trip ticket? *Social Cognition*, 4, 152-179.
- Hacker, P. M. S. (1996). *Wittgenstein: Mind and will. An analytical commentary on the Philosophical Investigations, vol. 4*. Oxford: Basil Blackwell.

- Hala, S., & Chandler, M. J. (1996). The role of strategic planning in accessing false-belief understanding. *Child Development*, 67, 2948-2966.
- Harris, P. L. (1983). Children's understanding of the link between situation and emotion. *Journal of Experimental Child Psychology*, 36, 490-509.
- Harris, P. L. (1989). *Children and emotion*. Oxford: Blackwell.
- Harris, P. L. (1995). Children's awareness and lack of awareness of mind and emotion. In D. Cicchetti & S. Toth (Eds.), *Rochester symposium on developmental psychopathology: Vol. 6 Emotion, cognition, and representation* (pp. 35-57). Rochester, NY: University of Rochester Press.
- Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. *Cognition and Emotion*, 3, 379-400.
- Harré, R. (1982). *Personal being: A theory of individual psychology*. Oxford: Basil Blackwell.
- Hart, D., & Fegley, S. (1995). Prosocial behavior and caring in adolescence: Relations to self-understanding and social judgment. *Child Development*, 66, 1346-1359.
- Hart, H. L. A. (1968). *Punishment and responsibility: Essays in the philosophy of law*. Oxford University Press.
- Harter, S. (1983). Children's understanding of multiple emotions: A cognitive-developmental approach. In W. F. Overton (Ed.), *The relationship between social and cognitive development* (pp. 147-194). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Harter, S., & Buddin, B. J. (1987). Children's understanding of the simultaneity of two emotions: A five-stage developmental acquisition sequence. *Developmental Psychology*, 23, 388-399.

- Harter, S., & Whitesell, N. R. (1989). Developmental changes in children's understanding of single, multiple, and blended emotion concepts. In C. Saarni & P. L. Harris (Eds.), *Children's understanding of emotion* (pp. 81-116). Cambridge: Cambridge University Press.
- Hodapp, A. F. (1993). Correlation between Stanford-Binet IV and PPVT-R scores for young children. *Psychological Reports*, 73, 1152-1154.
- Hodapp, A. F., & Gerken, K. C. (1999). Correlations between scores for the Peabody Picture Vocabulary Test—III and the Wechsler Intelligence Scale for Children—III. *Psychological Reports*, 84, 1139-1142.
- Hoffman, M. L. (1984). Empathy, its limitations, and its role in a comprehensive moral theory. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Morality, moral behavior, and moral development* (pp. 283-302). New York: John Wiley & Sons.
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge: Cambridge University Press.
- Inhelder, B. & Piaget, J. (1969). *The early growth of logic in the child*. New York: Norton.
- Kahn, Jr., P. H. (1992). Children's obligatory and discretionary moral judgments. *Child Development*, 63, 416-430.
- Keller, M., Lourenço, O., Malti, T., & Saalbach, H. (2003). The multifaceted phenomenon of 'happy victimizers': A cross-cultural comparison of moral emotions. *British Journal of Developmental Psychology*, 21, 1-18.
- Kenny, A. (1963). *Action, emotion, and will*. London: Routledge and Kegan Paul.
- Kim, J. (1996). *Philosophy of mind*. Boulder, CO: Westview Press.
- Kim, S. (1993). *Children's understanding of emotions in moral transgressions*. Unpublished doctoral dissertation, Yonsei University, Seoul, Korea.

- Krettenauer, T., & Eichler, D. (in press). Moral emotions in adolescence: An extension of the happy victimizer research paradigm. *British Journal of Developmental Psychology*.
- Koestler, A. (1967). *The ghost in the machine*. London: Hutchinson.
- Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology*, 20, 163-198.
- Lightfoot, C., & Bullock, M. (1990). Interpreting contradictory communications: Age and context effects. *Developmental Psychology*, 26, 830-836.
- Lourenço, O. (1997). Children's attributions of moral emotions to victimizers: Some data, doubts, and suggestions. *British Journal of Developmental Psychology*, 15, 425-438.
- Malle, B. F., & Knobe, J. (2001). The distinction between desire and intention: A folk-conceptual analysis. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Understanding intentions and intentionality: Foundations of social cognition* (pp. 45-67). Cambridge, MA: MIT Press.
- Malle, B. F., Moses, L. J., & Baldwin, D. A. (2001). Introduction: The significance of intentionality. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 1-24). Cambridge: MIT Press.
- Martin, J., Sugarman, J., & Thompson, J. (2003). *Psychology and the question of agency*. Albany: SUNY Press.
- McNaughton, D. (1988). *Moral vision: An introduction to ethics*. Oxford: Basil Blackwell.
- Meltzoff, A. N., Gopnik, A., & Repacholi, B. M. (1999). Toddler's understanding of intentions, desires, and emotions: Exploring the dark ages. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 17-41). Hillsdale, NJ: Erlbaum.

- Mischel, H. N., & Mischel, W. (1983). The development of children's knowledge of self-control strategies. *Child Development*, 54, 603-619.
- Mischel, W., Shoda, Y., & Rodriguez, M. L. (1992). Delay of gratification in children. In G. Loewenstein & J. Elster (Eds.) *Choice over time* (pp. 147-164). New York: Sage.
- Mitchell, R. W. (1986). A framework for discussing deception. In R. W. Mitchell & N. S. Thompson (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 3-31). Albany: SUNY Press.
- Moshman, D. (1999). *Adolescent psychological development: Rationality, morality, and identity*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Moya, C. J. (1990). *The philosophy of action: An introduction*. Cambridge, UK: Polity Press.
- Murgatroyd, S. J., & Robinson, E. J. (1993). Children's judgments of emotions following moral transgression. *International Journal of Behavioral Development*, 16, 93-111.
- Murgatroyd, S. J., & Robinson, E. J. (1997). Children's and adults' attributions of emotion to a wrongdoer: The influence of the onlooker's reaction. *Cognition and Emotion*, 11, 83-101.
- Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review*, 108, 291-310.
- Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science*, 26, 653-684.
- Nucci, L. (1981). The development of personal concepts: A domain distinct from moral or societal concepts. *Child Development*, 52, 114-121.
- Nucci, L. (1996). Morality and the personal sphere of actions In T. Brown, E. Turiel, & E. Reed (Eds.), *Knowledge and values* (pp. 41-60). Hillsdale, NJ: Erlbaum.

- Nunner-Winkler, G. (1999). Development of moral understanding and moral motivation. In F. E. Weinert & W. Schneider (Eds.), *Individual development from 3 to 12: Findings from the Munich longitudinal study* (pp. 253-290). Cambridge: Cambridge University Press.
- Nunner-Winkler, G., & Sodian, B. (1988). Children's understanding of moral emotions. *Child Development, 59*, 1323-38.
- Overton, W. F. (1991). Historical and contemporary perspectives on developmental theory and research strategies. In R. Downs, L. Liben, & D. Palermo (Eds.), *Visions of aesthetics, the environment, and development: The legacy of Joachim Wohlwill* (pp. 263-311). Hillsdale, NJ: Erlbaum.
- Peng, M., Johnson, C., Pollock, J., Glasspool, R., & Harris, P. (1992). Training young children to acknowledge mixed emotions. *Cognition and Emotion, 6*, 387-401.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: The MIT Press.
- 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)
- Piaget, J. (1981). *Intelligence and affectivity: Their relationship during child development*. (T. Brown & C. Kaegi, Trans. and Eds.). Palo Alto, CA: Annual Reviews Monograph. (Original work published 1954)
- 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.
- Price, R. (1953). *Doodles*. New York: Simon & Schuster.
- Reissland, N. (1985). The development of concepts of simultaneity in children's understanding of emotions. *Journal of Child Psychology and Psychiatry, 26*, 811-824.
- Rich, J. M. (1986). Morality, reason, and emotion. In S. Modgil & C. Modgil (Eds.), *Lawrence Kohlberg: Consensus and controversy* (pp. 209-222). London: Falmer Press.
- Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton: Princeton University Press.
- Rottschaefer, W. A. (1991). Social learning theories of moral agency. *Behavior and Philosophy, 19*, 61-75.
- Rottschaefer, W. A. (1998). *The biology and psychology of moral agency*. Cambridge: Cambridge University Press.
- Saarni, C. (1999). *The development of emotional competence*. New York: The Guilford Press.
- Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations*. Belmont, CA: Wadsworth, Inc.
- Schult, C. A. (2002). Children's understanding of the distinction between intentions and desires. *Child Development, 73*, 1727-1747.
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.
- Selman, R. L. (1971). Taking another's perspective: Role-taking development in early childhood. *Child Development, 42*, 1721-1734.

- Selman, R. L. (1976). Social-cognitive understanding: A guide to educational and clinical practice. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp. 299-316). New York: Holt, Rinehart, and Winston.
- Selman, R. L. (1980). *The growth of interpersonal understanding*. New York: Academic Press.
- Selman, R. L. (1981). The development of interpersonal competence: The role of understanding conduct. *Developmental Review*, 1, 401-422.
- Skinner, B. F. (1973). Answers for my critics. In H. Wheeler (Ed.), *Beyond the punitive society: Operation conditioning: Social and political aspects* (pp. 256-266). San Francisco: W.H. Freeman and Co.
- Smetana, J. (1995). Morality in context: Abstractions, ambiguities, and applications. *Annals of Child Development*, 10, 83-130.
- Smith, L., & Rogers, C. (2003). Introduction. *British Journal of Educational Psychology Monograph Series II* (Number 2 - Development and Motivation), 1-11.
- Sodian, B., Taylor, C., Harris, P. L., & Perner, J. (1991). Early deception and the child's theory of mind: False trails and genuine markers. *Child Development*, 62, 468-483.
- Sokol, B. W., Chandler, M. J., & Jones, C. (2004). From mechanical to autonomous agency: The relationship between children's moral judgments and their developing theories of mind. In J. A. Baird & B. W. Sokol (Eds.), *Connections between theory of mind and sociomoral development* (pp. 19-36). San Francisco: Jossey-Bass Inc.
- Stein, N. L., & Trabasso, T. (1989). Children's understanding of changing emotional states. In C. Saarni & P. L. Harris (Eds.), *Children's understanding of emotion* (pp. 50-77). Cambridge: Cambridge University Press.

- Stipek, D. J., & DeCotis, K. M. (1988). Children's understanding of the implications of causal attributions for emotional experiences. *Child Development*, 59, 1601-16.
- Taylor, R. (1958). Determinism and the theory of agency. In S. Hook (Eds.), *Determinism and freedom in the age of modern science* (pp. 211-218). New York: New York University Press.
- Taylor, R. (1966). *Action and purpose*. Englewood Cliffs, NJ: Prentice-Hall.
- Thompson, R. A. (1987). Development of children's inferences of the emotions of others. *Developmental Psychology*, 23, 124-131.
- Thompson, R. A. (1989). Causal attributions and children's emotional understanding. In C. Saarni & P. L. Harris (Eds.), *Children's understanding of emotion* (pp. 117-150). Cambridge: Cambridge University Press.
- Thompson, R. A., & Hoffman, M. L. (1980). Empathy and the development of guilt in children. *Developmental Psychology*, 16, 155-156.
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge: Cambridge University Press.
- Turiel, E. (1998). The development of morality. In W. Damon (Series Ed.) and N. Eisenberg (Vol. Ed.), *The handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 863-932). New York: Wiley.
- Turiel, E., & Smetana, J. G. (1984). Social knowledge and social action: The coordination of domains. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Morality, moral behavior, and moral development* (pp. 261-282). New York: John Wiley & Sons.
- Turiel, E., Hildebrandt, C., & Wainryb, C. (1991). Judging social issues: Difficulties, inconsistencies, and consistencies. *Monographs of the Society for Research in Child Development*, Vol. 56 (Serial 224).

- Turiel, E., Killen, M., & Helwig, C. C. (1987). Morality: Its structure, functions, and vagaries. In J. Kagan & S. Lamb (Eds.), *The emergence of moral concepts in young children* (pp. 155-244). Chicago: University of Chicago Press.
- Velleman, J. D. (1993). What happens when someone acts? In J. M. Fischer & M. Ravizza (Eds.), *Perspectives on moral responsibility* (pp. 188-210). Ithaca, NY: Cornell University Press.
- Velleman, J. D. (2000). *The possibility of practical reason*. Oxford, U.K.: Clarendon Press.
- Wainryb, C. (2000). Values and truths: The making and judging of moral decisions. In M. Laupa (Ed.), *Rights and wrongs: How children and young adults evaluate the world* (pp. 33-46). San Francisco: Jossey-Bass.
- Walker, L. J., & Hennig, K. H. (1997). Moral development in the broader context of personality. In S. Hala (Ed.), *The development of social cognition* (pp. 297-327). Hove, UK: Psychology Press.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.
- Weiner, B., & Graham, S. (1984). An attributional approach to emotional development. In C. E. Izard, J. Kagan, & R. B. Zajonc (Eds.), *Emotion, cognition, and behavior* (pp. 167-191). Cambridge: Cambridge University Press.
- Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.
- Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of belief and desires. *British Journal of Developmental Psychology*, 9, 191-214.

- Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, 35, 245-275.
- Wiersma, N., & Laupa, M. (2000). Young children's conceptions of the emotional consequences of varied social events. *Merrill-Palmer Quarterly*, 46, 325-341.
- Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). New York: Macmillan.
- Wintre, M., Polivy, J., & Murray, M. A. (1990). Self-predictions of emotional response patterns: Age, sex, and situational determinants. *Child Development*, 61, 1124-1133.
- Wintre, M. G., & Vallance, D. D. (1994). A developmental sequence in the comprehension of emotions: intensity, multiple emotions, and valence. *Developmental Psychology*, 30, 509-514.
- Wren, T. E. (1991). *Caring about morality: Philosophical perspectives in moral psychology*. London: Routledge.
- Wren, T. E. (1974). *Agency and urgency: The origin of moral obligation*. New York: Precedent.
- Wren, T. E., & Noam, G. G. (Eds.) (1993). *The moral self*. Cambridge, MA: MIT Press.
- Yeates, K. O., Schultz, L. H., & Selman, R. L. (1991). The development of interpersonal negotiation strategies in thought and action: A social-cognitive link to behavioral adjustment and social status. *Merrill-Palmer Quarterly*, 37, 369-406.
- Yuill, N., Perner, J., Pearson, A., Peerbhoy, D., & van den Emde, J. (1996). Children's changing understanding of wicked desires: From objective to subjective and moral. *British Journal of Developmental Psychology*, 14, 457-475.
- Zelazo, P. D., Astington, J. W., & Olson, D. R. (Eds.) (1999). *Developing theories of intention: Social understanding and self-control*. Hillsdale, NJ: Erlbaum.

Appendix A

Theory-of-Mind Protocol ("Doodles" task)

Introduction.

Here are TweedleDee and TweedleDum. I want you to pretend that these dolls are real people just like you and me. That means that they see and hear and know things just like real people. Okay? They live together in this box—their house. When TweedleDee and TweedleDum are inside their house, they can't hear what we're saying, and they can't see what we're doing. [The dolls are then placed inside the box]

--*Can TweedleDee hear us talking right now?*
--*Can TweedleDum see us right now?*

Okay, we're going to look at some pictures together. TweedleDee and TweedleDum have never seen these before.

Guess Trial. [Leave Doodle inside envelope and present to subject]

--*What is this a picture of?* [direct attention to relevant parts]
--*That's a good guess, have another guess.* [Record at least two different guesses]

Good. Here's the whole picture. Now can you tell me what is it? [direct attention to relevant parts]

[Put Doodle back in envelope] Now, I'm going to cover the picture again so we can see only this very small part of it, okay? Remember, TweedleDee and TweedleDum have never seen this picture before. We're going to show it to them, but all they're going to see is this right here [point to restricted view].

Let's get TweedleDee out of the house and show him the picture first.

--*So TweedleDee has never seen this picture before. What will TweedleDee say this is?*
[Probe, if necessary: What will TweedleDee think this part is?]

[Put TweedleDee to the side] Now let's get TweedleDum out of the house and show him the picture.

--*So TweedleDum has never seen this picture before. What will TweedleDum say this is?*
[Probe, if necessary: What will TweedleDum think this part is?]

Recall Phase.

--*Remember when I first showed you this [restricted view of picture], what was your first guess?*
--*What was your second guess?*

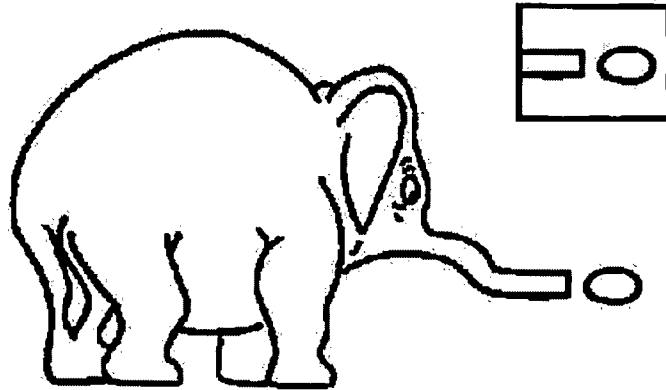
Standard Trial. [Take Droodle out of envelope and present to subject]

--*What is this a picture of?* [direct attention to relevant parts]

[Put Droodle back in envelope] Now, I'm going to cover the picture so we can see only this very small part of it, okay? Remember, TweedleDee and TweedleDum have never seen this picture before. We're going to show it to them, but all they're going to see is this right here [point to restricted view].

Repeat questions from the previous trial for each puppet.

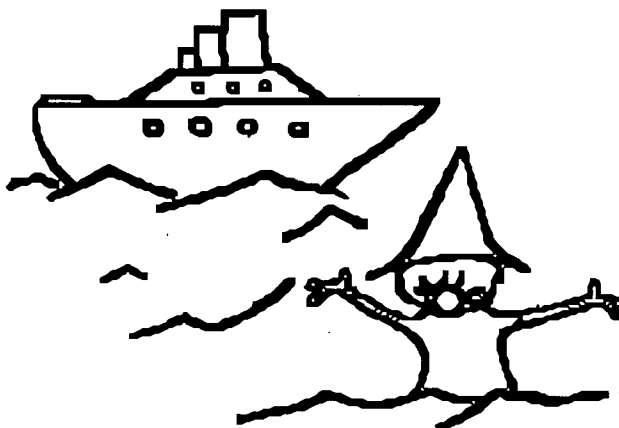
Restricted View



Full Picture

“An elephant smelling a grapefruit”

Restricted View



Full Picture

“A ship arriving too late to save a drowning witch”

Appendix B

Victimizer Protocol ("Swing" task)

Control Questions: Story comprehension and rule understanding

--*Tell me how the story went. Start at the beginning.*

[The most important features here are Jesse's desire and how s/he achieved it: Use additional questions below if necessary]

--*When Jesse saw his/her friend playing on the swing, what did Jesse want?*

[Additional probe, if necessary: What did Jesse want to play on?]

--*After Jesse saw his/her friend on the swing, what did Jesse do to get to play on the swing?*

[Additional probe, if necessary: Did s/he push his/her friend onto the ground?]

--*Is what Jesse did to play on the swing okay or not okay?*

--*Why is it [subject's response]?*

--*What else could Jesse have done to get to play on the swing?*

--*Would that have been okay or not okay to do?*

Test Questions: Emotion attributions

Traditional Condition

--*I would like to talk about how the boys/girls in this story are feeling at the end. Tell me how the boy/girl on the ground, Jaime, feels?* (use the last page to point to Jaime)

[Additional probe, if necessary: Is Jaime happy or sad?]

--*How does Jesse feel at the end of the story?*

--*Why does s/he feel like that?*

--*What else might Jesse be feeling?* (repeat as often as necessary)

[Additional probe, if necessary: Do you think Jesse is feeling anything else besides (subject's response)?]

Alternative Condition

--*What did Jesse do to get to play on the swing?*

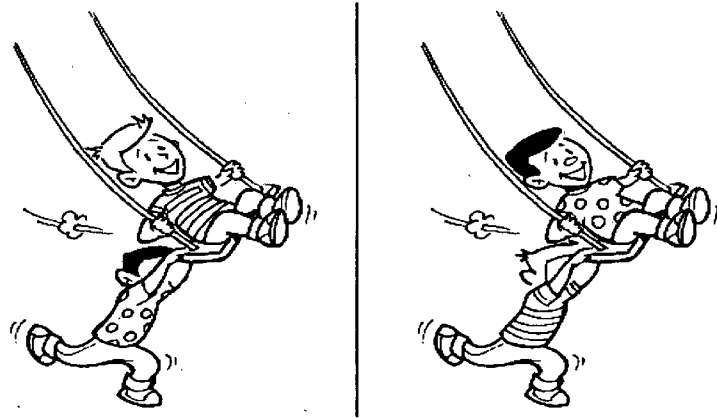
[Additional probe, if necessary: Did s/he push his/her friend onto the ground?]

--*How does Jesse feel about acting like that?*

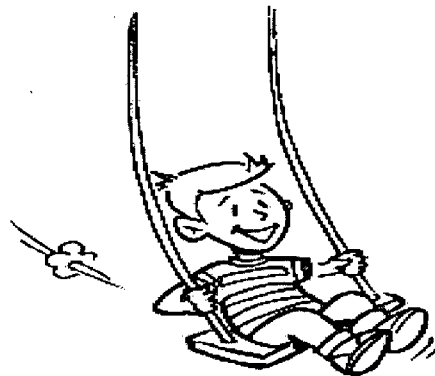
--*Why does s/he feel like that?*

--*How else might Jesse feel about acting like that?* (repeat as often as necessary)

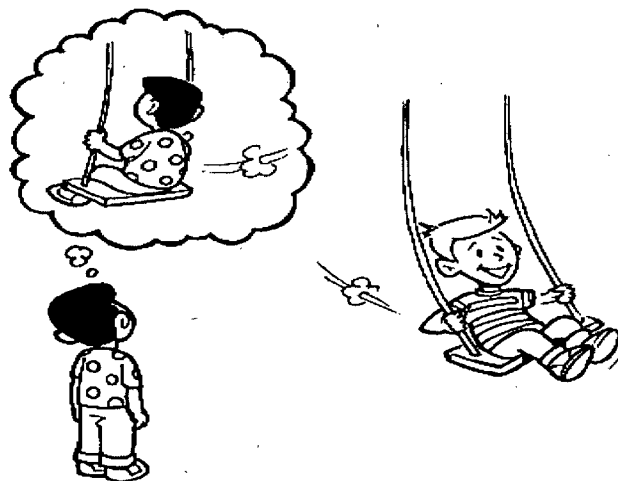
[Additional probe, if necessary: Do you think Jesse is feeling anything else besides (subject's response)?]



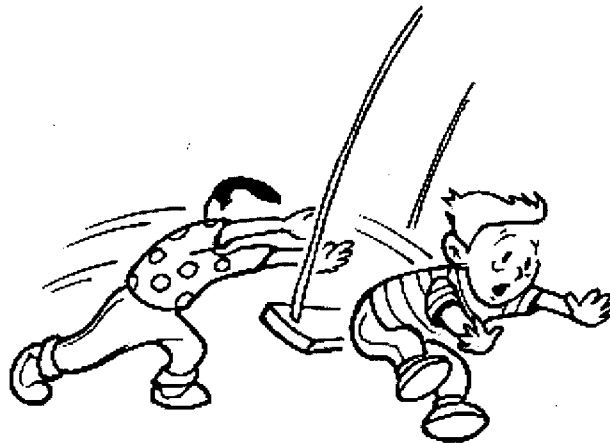
Panel 1: Jaime and Jesse are best friends. They play almost every day together taking turns on this swing.



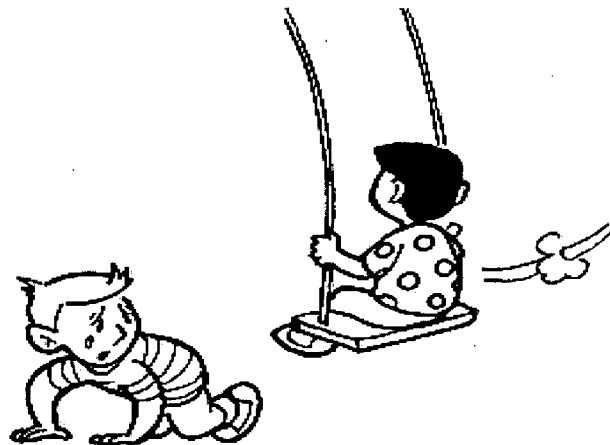
Panel 2: One day Jaime is playing on the swing alone.



Panel 3: When Jesse arrives, he sees Jaime and thinks to himself: "I want to be playing on the swing."



Panel 4: So what does Jesse do? He acts mean and pushes Jaime to the ground!



Panel 5: Now Jesse is playing on the swing and Jaime is hurt on the ground.

Appendix C

Non-Victimizer Protocol ("Race" task)

Control Questions: Story comprehension and rule understanding

--*Tell me how the story went. Start at the beginning.*

[The most important features here are Racer 3's desire and how s/he achieved it: Use additional questions below if necessary]

--*Before the race, how was Racer 3 planning to win?*

[Additional probe, if necessary: What sorts of things did Racer 3 try to do to win the race?]

--*During the race, what happened so that Racer 3 could win?*

[Additional probe, if necessary: What happened during the race that really let Racer 3 win?]

--*Is the way that Racer 3 won the race okay or not okay?*

--*Did Racer 3 do anything wrong to win the race?*

--*Why is it [subject's response]? (if necessary)*

Test Questions: Emotion attributions

Traditional Condition

--*I would like you to talk about how the boys/girls in this story are feeling at the end. Tell me how the boys/girls on the ground, Racers 1 and 2, feel?* (use the last page to point to characters)

[Additional probe, if necessary: Are Racers 1 and 2 happy or sad?]

--*How does Racer 3 feel at the end of the story?*

--*Why does s/he feel like that?*

--*What else might Racer 3 be feeling?* (repeat as often as necessary)

[Additional probe, if necessary: Do you think Racer 3 is feeling anything else besides (subject's response)?]

Alternative Condition

--*What happened so that Racer 3 could win?*

[Additional probe, if necessary: What happened during the race that really let Racer 3 win?]

--*How does Racer 3 feel about winning like that?*

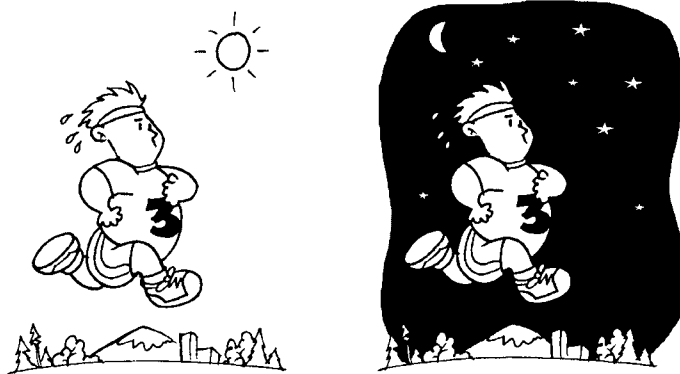
--*Why does s/he feel like that?*

--*How else might Racer 3 feel about winning like that?* (repeat as often as necessary)

[Additional probe, if necessary: Do you think Racer 3 is feeling anything else besides (subject's response)?]



Panel 1: Three friends are getting ready to run in a race against each other. Two of them have won lots of times, but one has never been in a race before. Can you tell which one that is?



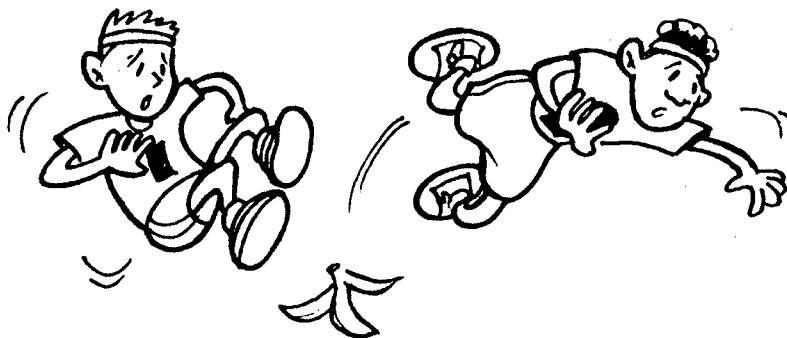
Panel 2: Racer 3 really wants to win, but he knows that his friends are great runners. It won't be easy to beat them! So he practices hard, day and night, to become a great runner too.



Panel 3: On race day, Racer 3 is ready. He plans to run as fast as he can and beat the others. Will he do it? On your marks, get set, GO!



Panel 4: Whoosh! Racers 1 and 2 are really fast. And even though Racer 3 is trying hard, he just can't keep up with them. They are too fast!



Panel 5: But what's this? Racers 1 and 2 slip and fall before they reach the finish line. Ouch! And now Racer 3 is the only runner left.



Panel 6: Racer 3 runs past his friends on the ground and crosses the finish line. He wins his very first race.

Appendix D

Intention-Desire Mismatch Protocol

Foiled Plan Condition ("Rabbit" story)

Introduce Alice (the doll) and the story by saying: "This story is about Alice and a white rabbit"

[Note: Close picture book when asking questions.]

Control Questions.

--*What did Alice want to catch in this story?*

--*How is she planning to do that?* [Emphasize only the TWO steps of plan; refer to book if necessary & repeat story]

Good. Now listen carefully to these next questions.

Want Question.

--*Did Alice get what she wanted?* (YES or NO)

Plan Question.

--*Did Alice do what she planned?* (YES or NO)

Inadvertent Success Condition ("Dormouse" story)

Introduce Alice (the doll) and the story by saying: "This story is about Alice too, but this time she is trying to catch a dormouse"

[Note: Close picture book when asking questions.]

Control Questions.

--*What did Alice want to catch in this story?*

--*How is she planning to do that?* [Emphasize only the TWO steps of plan; refer to book if necessary & repeat story]

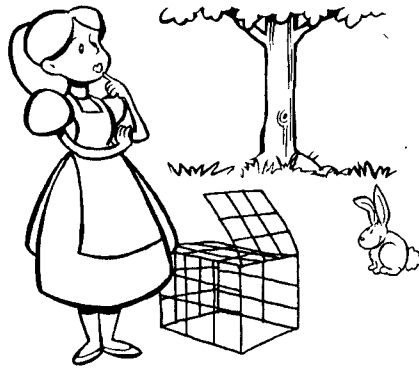
Good. Now listen carefully to these next questions.

Want Question.

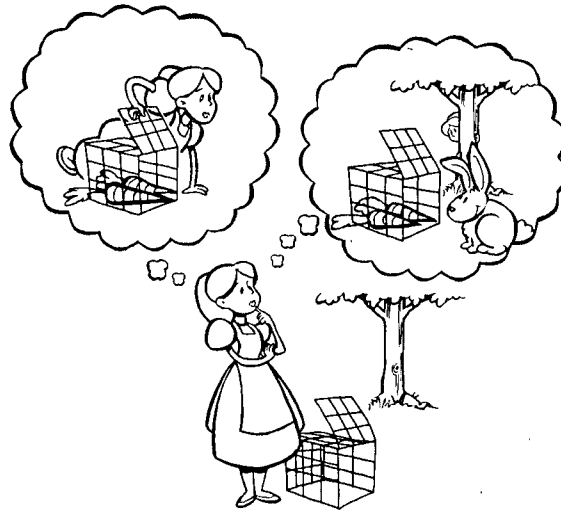
--*Did Alice get what she wanted?* (YES or NO)

Plan Question.

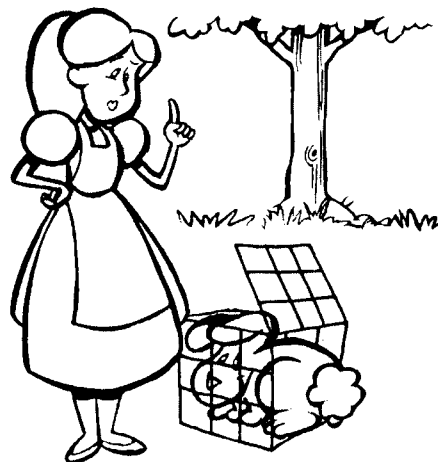
--*Did Alice do what she planned?* (YES or NO)



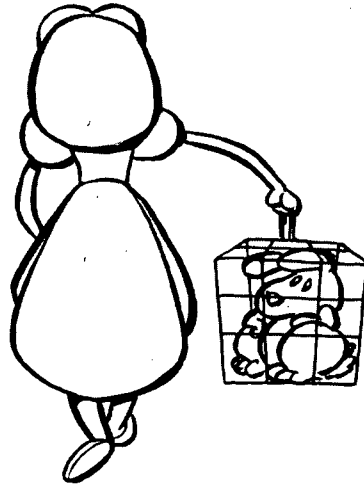
Panel 1: Alice wants to catch the white rabbit and take him home with her. Here she is planning how to catch him.



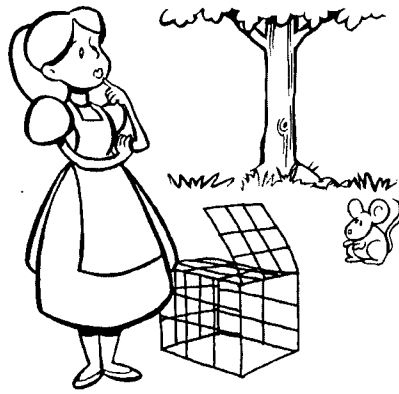
Panel 2: First, she will put some carrots in the cage, because rabbits love to eat carrots. Then, she will hide behind the tree and wait for the rabbit to go into the cage and eat them. That's Alice's plan, and when the white rabbit goes into the cage, she'll close the cage door and take him home with her.



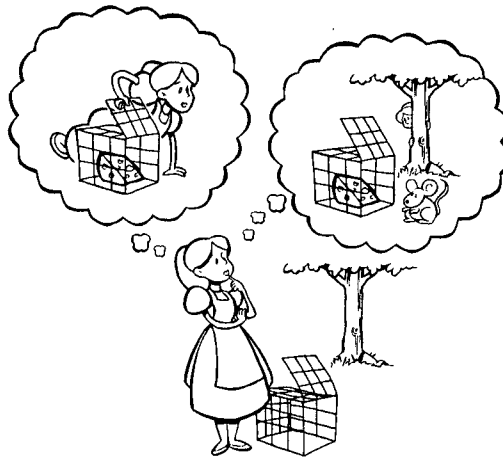
Panel 3: While Alice is planning to do these things, though, the white rabbit hops into the cage all by himself. So before she even gets the carrots and before she even hides behind the tree, Alice has the white rabbit right where she wants him.



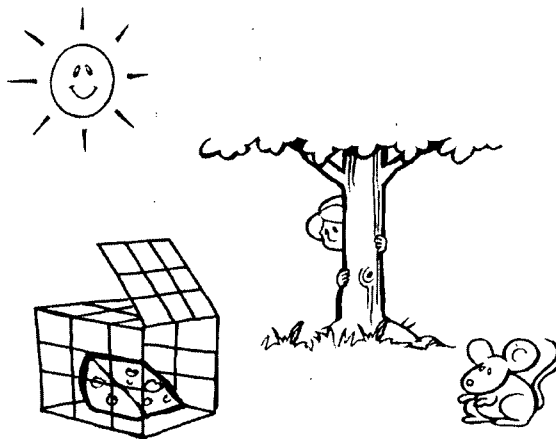
Panel 4: *Alice closes the cage door and takes the white rabbit home.*



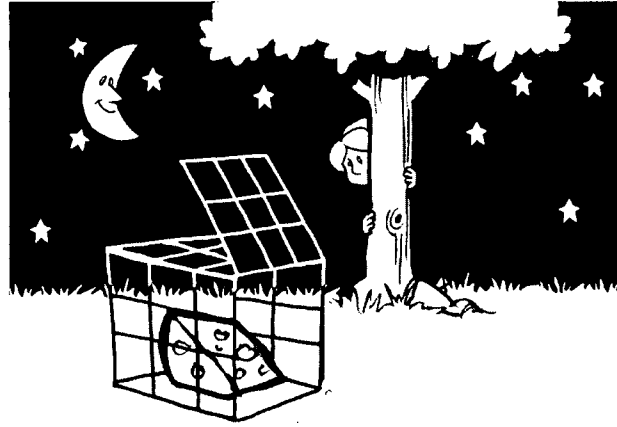
Panel 1: Alice wants to catch the dormouse and take him home with her. Here she is planning how to catch him.



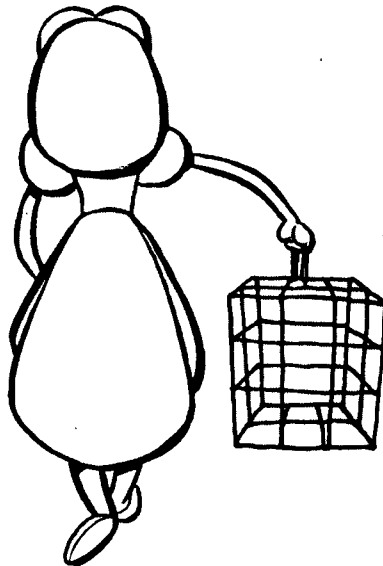
Panel 2: First, she will put some cheese in the cage, because mice love to eat cheese. Then, she will hide behind the tree and wait for him to go into the cage and eat it. That's Alice's plan, and when the dormouse goes into the cage, she'll close the cage door and take him home with her.



Panel 3: So Alice does just what she has been planning. She gets the cheese and puts it in the cage. Then, she hides behind the tree and waits.



Panel 4: Alice waits for a very long time, but the dormouse does not go inside the cage.



Panel 5: Finally, she takes the cage and goes home without the dormouse.