Young Children's Vulnerability to Self-Blame and Helplessness: Relationship to Beliefs about Goodness

Gail D. Heyman
University of Illinois at Urbana-Champaign

Carol S. Dweck
Columbia University

Kathleen M. Cain
Gettysburg College

Heyman, Gail D.; Dweck, Carol S.; and Cain, Kathleen M. Young Children's Vulnerability to Self-Blame and Helplessness: Relationship to Beliefs about Goodness. CHILD DEVELOPMENT, 1992, 63, 401–415. Motivational helplessness, linked to conceptions of intelligence, has been well documented in older children. While some researchers have reported that children just starting school are motivationally invulnerable, others have found evidence of helplessness when these children encounter failure. The present study seeks to determine whether the reactions associated with helplessness can be identified in a new context, that of criticism, and whether any such responses are related to the child's conceptions of goodness. Subjects were 107 5- and 6-year-old children who enacted achievement situations in which teacher criticism was presented. The 39% of children whose own assessments were undermined by criticism exhibited the affect, task choices, and nonconstructive problem-solving strategies characteristic of helplessness. They were also more likely to make global negative self-judgments following criticism, including negative judgments of their goodness. Finally, these children were more likely to endorse stable and global beliefs about goodness.

Many prominent researchers have recognized the importance of mastery motivation to development (see, e.g., Erikson, 1963; White, 1959). However, it is clear that by middle childhood, mastery striving is often abandoned in contexts in which negative outcomes are encountered (Diener & Dweck, 1978, 1980; see also Covington & A. N. V. A. U. V. M. A. 1. -4. u 1. -4. affect, to engage in negative self-cognitions, and to give up easily. Other children tend to show a more mastery-oriented pattern of response. These children tend to experience negative affect, to engage in negative self-cognitions, and to give up easily. Other children tend to show a more mastery-oriented pattern of response. These children are more likely than others to maintain positive affect and to remain actively engaged in problem solving.

By the time children reach the age of 9 or 10 there are strong individual differences in responses to failure. Some children tend to show a helpless pattern of response. These children tend to experience negative affect, to engage in negative self-cognitions, and to give up easily. Other children tend to show a more mastery-oriented pattern of response. These children are more likely than others to maintain positive affect and to remain actively engaged in problem solving.

This study was submitted in partial fulfillment of the requirements for the Masters of Arts degree in psychology at the University of Illinois and was supported by a grant from NICHD (HD07025-10). The authors wish to thank Brian Compton for his extensive help throughout all phases of this project. The authors also wish to thank Iris Heyman for her assistance in data collection, Jennifer Babula for her assistance in coding data, Mary Sue Compton for her drawings which were used in the study and Renee Baillegeon, Cindy Erdley, Valanne Henderson, Catherine Smith, and Robert Wyer for their suggestions and helpful comments on earlier versions of the manuscript. Requests for reprints should be addressed to Gail Heyman, Department of Psychology, University of Illinois, 603 E. Daniel Street, Champaign, IL 61820.

[Child Development, 1992, 63, 401–415. © 1992 by the Society for Research in Child Development, Inc. All rights reserved. 0009-3920/92/6302-0011$01.00]
While it is clear that many older children experience motivational helplessness, the case is less clear for younger children. By applying the experimental methodologies that have been used to study achievement motivation in older children to younger children, researchers have discovered that in many of the situations in which older children exhibit helplessness, younger children do not tend to do so. In these studies, young children as compared to older children are found to maintain an adaptive motivational stance after failure (Rholes, Blackwell, Jordan, & Walters, 1980; Ruble, Parsons, & Ross, 1976; see also Parsons & Ruble, 1977).

Although young children appear to be motivationally invulnerable when performing certain tasks, it is possible that when failure is clear and relevant to them, they, like older children, are susceptible to motivational difficulties (see Stipek & Daniels, 1988; Stipek, Roberts, & Sanborn, 1984, concerning the impact of salient feedback for young children). This seems particularly plausible in light of evidence that children show concerns with meeting standards even by the age of 2 (Kagan, 1981). In several studies, researchers have begun to test this hypothesis by providing preschoolers and first graders with salient failure experiences and then assessing their motivation in age-appropriate ways (Cain, 1990; Hebert & Dweck, 1985; Smiley, 1989; see Dweck, 1991, for summary). In these studies, children were given four jigsaw puzzles to do, but were able to solve only one of them. Later they were given a chance to rework the puzzle of their choice. Children were considered to be “persisters” if they expressed the desire to continue working on one of the puzzles that they had not yet mastered, and “nonpersisters” if they chose to rework the puzzle that they had already successfully completed. With this behavioral criterion established, these researchers then looked for the affective and cognitive concomitants characteristic of the helpless and mastery-oriented motivational patterns. In each of these studies, there was a considerable proportion of nonpersistent children who showed a pattern that parallels the pattern of helplessness evident in studies of older children. For example, these nonpersisters reported more negative affect than persisters and were less likely to perceive their failures as surmountable. In short, these studies provide evidence that some young children exhibit motivational responses similar to the helpless responses of older children.

Is there a link between young children’s responses to difficulty in achievement situations and beliefs they hold outside of immediate situational contexts? For older children, one such link has been found: the degree to which individuals think intelligence is malleable (Bandura & Dweck, 1985; Cain, 1990; Leggett, 1985; see also Dweck & Leggett, 1988, for a summary of this work). Specifically, these studies suggest that when children do not view intelligence as a malleable quality, performance may serve the function of diagnosing enduring intellectual capacities. Under this framework, poor performance implies low ability and often leads to helplessness. In contrast, when children view intelligence as a malleable quality that can be increased through effort, they tend to maintain adaptive motivational patterns, presumably because they view performance outcomes as reflecting only effort, strategy, or current ability level.

Clear links between motivation and cognition have been described for older children, but not for children younger than age 9. Many researchers have argued that young children do not understand the concept of ability as a capacity that is distinct from effort, and as a result do not view failure as reflecting upon ability. In contrast, when children view intelligence as a malleable quality that can be increased through effort, they tend to maintain adaptive motivational patterns, presumably because they view performance outcomes as reflecting only effort, strategy, or current ability level.

A definitive answer concerning the degree to which young children understand ability has not yet been reached. Although there is evidence of difficulty in reasoning about ability in certain contexts, there does appear to be some basic understanding of ability (see Cain, 1988; Cain & Dweck, 1989).
1984), it is possible that some form of trait thinking may be in operation when salient information is available.

What sorts of motivationally relevant trait concepts are children likely to hold early in life? There is evidence to suggest that children's trait concepts begin by being rather broad, and become differentiated over the course of development (Benenson & Dweck, 1986; Heller & Berndt, 1981; Mullener & Laird, 1971; Stipek & Mac Iver, 1989). Teaching children what is right or wrong, good or bad, is a major aim of socialization. In addition, children's books and movies are replete with characters that are "good" and "bad." As a result, young children are likely to be familiar with concepts of goodness and badness. They may apply these concepts to a variety of situations even before they have fully developed formulations of specific social or intellectual traits. If such a general evaluative trait is indeed primary in young children, it may encompass what older children and adults would consider to be a variety of different traits. Thus, while older children are using information from the achievement domain to figure out how smart they are, younger children may be using this same information to make inferences about how good or bad they are.

There is some evidence that children just starting school have a tendency to view information about their achievement as reflecting upon the domain of goodness. Hebert and Dweck (1985) were interested in the reactions children expected from social agents for their performance on the puzzle task that was previously described. In this study, children were asked to role-play the ensuing interactions of their teacher and parents. Children's responses contained frequent references to how good or bad they had been. Interestingly, nonpersisters were much more likely to role-play a punitive response to their failure. For example, one nonpersister portrayed her mother as saying, "Daddy's gonna be very mad and spank her," because she could only complete one out of four jigsaw puzzles. This work suggests that young children may have conceptions of their goodness that are tied to their performance.

What aspects of thinking about goodness might be linked to individual differences in motivational responses? One important aspect may be the type of inferences children make when they are faced with limited information. For example, some children may make broad inferences across time; they may witness a small set of "good" or "bad" behaviors and assume that the person cannot or is unlikely to change. For these children, negative achievement outcomes may pose a special threat, just as they do for older children who think about intelligence in a nonmalleable way. In addition, some children may be inclined to make broad trait inferences on the basis of limited information. Specifically, these children may be inclined to view a few instances of misbehavior (or even mistakes on a task) as diagnostic of general badness. Thus, it may be that when children have a tendency to make global inferences concerning goodness from specific information, negative outcomes take on broader meaning and lead to maladaptive motivational responses.

In short, recent evidence suggests that young children can show helpless motivational responses after a salient failure. However, for young children, differences in motivational patterns may be closely related to conceptions of goodness and badness.

One goal of this study was to confirm previous findings which indicate that some young children show the affective reactions and make the types of self-evaluations that have been associated with helplessness. The domain of adult criticism was chosen for investigation because it is meaningful and relevant for young children. The investigation employed interactive scenarios in which 5- and 6-year-old children pretended to create a product, and then were presented with a response of criticism by an adult. Following each scenario, subjects made evaluations of the products they had pretended to make. Children were divided into two groups for analysis based on these evaluations following criticism. The groups were then compared in terms of their affect, future task choices, and problem-solving strategies to examine whether they exhibited other differences that map onto the helpless and mastery-oriented motivation patterns. The groups were also compared in terms of their self-evaluations and general conceptions of goodness and badness. In summary, the overall purpose of the study was to determine whether some young children exhibit patterns of reactions associated with helplessness, and, if so, how such reactions relate to conceptions of goodness and badness.
Method

Subjects
Subjects in the present study were 107 5- and 6-year-old kindergartners (52 boys and 55 girls) drawn from a school district in an upper-middle-class neighborhood. All children whose parents gave consent to their participation were included.

Design and Procedure
Overview of procedure.—Children were interviewed individually in their schools by a female experimenter. First, children were asked about their general beliefs concerning goodness. Next, the experimenter read aloud three stories and directed children to simultaneously act out the role of the main character. The theme in all of the stories was the same: a child works hard on a task and then makes a small error. One of the stories presented ends at this point, and the other two stories end with the main character receiving criticism. Following the stories, children responded to a series of questions related to their motivation.

Design.—Subjects were presented with three analogous stories: story A, about painting; story B, about building a house out of blocks; and story C, about writing numbers. All subjects heard two stories ending in criticism and one story without such feedback. The story without feedback was presented first to half of subjects (assigned randomly) in order to establish a prejudgment baseline. The other half of the subjects encountered the no-feedback story last (postjudgment condition), and their evaluations on this story were used as a measure of the generalization of the reaction to criticism across situations.

Procedure.—A special procedure was devised in order to elicit meaningful responses and limit alternative explanations for any findings. First, to maximize the likelihood that reactions associated with helplessness would be exhibited, children engaged in familiar activities, and mistakes were made salient by criticism. Second, to maximize the likelihood of eliciting meaningful responses, questions were asked in vivid and concrete ways. Third, to maximize the likelihood that children would feel comfortable expressing their thoughts and emotions, questions were answered through role-play (see Mize & Ladd, 1988, concerning the advantages of enactive procedures). Finally, to ensure that any individual differences in motivation were not merely reactions to actual performance differences, children pretended to but did not actually create products.

Each child was individually brought from the classroom into the experimental area by one of two female experimenters, either the first author or a research assistant (both of whom, as will be explained later, remained blind to children's classifications during the procedure). Children were introduced to the experimenter and were informed that they were going to do some activities, and after that answer some questions that had no wrong answers. They were then interviewed to assess their beliefs about goodness.

After the initial interview, children were instructed to pick a toy person to represent themselves and were introduced to a female toy person who was to be their 'pretend teacher.' The children engaged in several familiarization activities with the toy people. Next, as the experimenter read aloud, children were directed to act out the three stories with their toy person using appropriate props. The following is the text of story A:

You spend a lot of time painting a picture of a family to give to your teacher. You pick out colors you think are nice and carefully draw each person. As you are about to give it to your teacher you say to yourself, "Uh oh, one of the kids has no feet." But you worked really hard on the picture and want to give it to her. You say, "Teacher, here's a picture for you."

After the experimenter finished reading the story, the child was directed to act out presenting the product to the teacher. When this story was administered without feedback, it ended at this point. When administered with criticism, the story continued, with the teacher responding as follows: "There are no feet on that child. That's not what I call drawing people the right way. I'm disappointed in you."

The other stories and corresponding feedback were analogous. In story B the child was building a house with blocks but forgot to put in windows. In story C, the child was writing out the number series from 1 to 10 but forgot to write the numeral 8. (See the Appendix for complete list of stories and feedback.)

After hearing each story, children were asked forced-choice questions about their affect and task preferences, as well as rating questions concerning their product and self-evaluations. In addition, a series of
open-ended questions assessed children’s recall of the criticism, their expected parent responses, appropriate teacher responses after the second feedback story, and story resolutions. Some of the open-ended questions were only asked after one story in order to maintain children’s attention in the experiment.

A great deal of care was taken to ensure that all children left the experimental session feeling good about themselves. In order to provide an opportunity for the toy teacher to offer praise, children were instructed to have their toy person help another toy person with a puzzle. The toy teacher informed all children that they did a great job and then offered a gracious apology for her earlier criticism. In cases where children had seemed at all upset during any part of the procedure, apologies were elaborated, more extensive praise was given, and the pretend nature of the interactions was emphasized. The experimenter gave all children stickers and thanked them for their help in the study.

General Beliefs
At the beginning of the interview, several questions were asked to assess general beliefs about goodness, and a question was asked about the validity of child versus adult opinions. When the questions referred to other children, the gender of the child in the question was matched to the gender of the subject. The general belief measures will be described for boys.

First, to examine whether children would make a negative assessment of other children as a result of an instance of poor academic performance, the experimenter asked, “Imagine a new boy is in your class. You look over at his schoolwork and see that he got lots and lots wrong and has a big frown on his paper. Does this mean that he is bad?”

Next, to learn whether children tended to see patterns of misbehavior as likely to persist over time, they were asked: “Imagine a new boy is in your class. You look over at his schoolwork and see that he got lots and lots wrong and has a big frown on his paper. Does this mean that he is bad?”

When subjects were asked, “Are there things you can do to make you more good than you are?” almost all said they could get better in some way. This, in conjunction with work done in a pilot study, suggests that when questions about goodness are asked in an abstract way, or when questions are phrased in terms of improvement, meaningful individual differences do not emerge. In contrast, children seemed to have no trouble thinking about the stability of a constellation of concrete behaviors.

Finally, to test whether vulnerability to criticism is merely a tendency to always accept adults’ opinions as being more valid than children’s opinions, children were asked: “You are having fun listening to some new music you really like. You tell your father, ‘I really like that music.’ Your father says, ‘I don’t like that music one bit.’ Who is right?”

Measures of Motivational Response
The following measures were administered after the stories, as explained in the procedure. The measures are described for story A for girls.

Product rating.—Children were asked to rank what they had done on a scale from 1 (a big frown) to 6 (a big smile). Children were told: “Really think again about everything that happened with the painting. Should you get a smile or a frown for what you did?” They were asked not to say their answer but instead to pick one of two envelopes with a smile or a frown on the front. (The experimenter could not see which envelope the child chose.) The envelope with the frown contained a big frown, a medium frown, and a little frown. The envelope with the smile contained a big smile, a medium smile, and a little smile. Children were then asked to choose the appropriate face within the envelope and to put it face down in a box. In this way the experimenter remained blind to the child’s product rating following criticism.

Task choice.—As an index of children’s willingness to continue to pursue the activities in the stories, children were asked, “Think again about everything that happened with the painting. If someone else asked you to paint a picture of a family, would you do it or would you do something else nice instead?” Children were asked to choose between a card depicting a child painting a picture and a card depicting the same child with a present behind his or her back.

The task-choice measure was scored from 0 to 2. Children were given a score of 0 if they did not choose to do the task in either feedback story, 1 if they chose to re-
peat the task once, and 2 if they chose to repeat the task in both feedback stories.

**Affect.**—To obtain a very general assessment of children’s affect during the stories, they were told, “I want to know about how you feel about what happened with the painting.” Children were shown a smiling face and asked if they felt happy about what had happened. Next, children saw a frowning face and were asked if they felt sad about what had happened. Finally, children saw an angry face and were asked if they felt mad about what happened. Children could select any number of these emotions. The rationale for allowing several emotions to be chosen was that during a pilot study, children often insisted that they were experiencing more than one of these feelings.

The affect measure was scored from 0 to 2 for each emotion. Children were given a score of 0 if they did not acknowledge experiencing the emotion in either feedback story, 1 if they reported experiencing the emotion in one of these stories, and 2 if they reported experiencing the emotion in both feedback stories.

**Self-ratings.**—To examine the degree to which children saw their performance as reflecting negatively on their traits and abilities, children were asked a series of four questions related to different aspects of themselves, in the following order: “Think about everything that happened with the painting. Did everything that happened make you feel like you were good or not good at painting?” “Did everything that happened make you feel like you were a good or not good girl?” “Did it make you feel like you were a nice or not nice girl?” “Did it make you feel like you were a smart or not smart girl?” The sum of the number of times children rated themselves positively on these four items on each of the two feedback stories was used to form a personal adequacy score for each story. Thus a child’s personal adequacy score could range from 0 to 8.

**Recall.**—To ensure that potential differences in responses to criticism were not merely the result of differences in the tendency to remember the criticism, children were asked to describe what they remembered about the story in this way: “I’d like you to tell me all you remember about everything that happened with the painting.” (This was asked only after the first feedback story.)

**Appropriate teacher response.**—To examine how the children thought they should have been responded to, the following question was asked: “Really think again about everything that happened with the painting. If you were the teacher, what would you say to [child’s name]?” (This was asked only after the second feedback story.)

**Expected parental response.**—To learn how the children expected their parents to respond in the situations presented in the stories, children were asked, “If you went home and your parents knew about what happened at school with the painting, what would they do or say?” (This was asked only after the first feedback story.)

**Story resolution.**—To assess what kinds of solutions children imagined for the stories, children were told, “Please hold your teacher [doll] and finish off what would happen next in the story.” Initial responses were probed. (This was asked after both feedback stories.)

**Teacher ratings.**—To confirm that children’s reaction to criticism in the stories was an ecologically valid way to assess vulnerability to criticism, teachers were asked to rank how true the following statement was for each child: “S/he becomes upset when told s/he made a mistake.” The scale was anchored from 1 (not true) to 5 (very true). In addition, to ensure that vulnerability to criticism was not merely a reflection of a child’s academic competence, teachers were asked to rate each child’s overall academic ability from 1 (very low) to 5 (very high).

**Results**

**Overview of Findings**

Children were assigned to categories for analysis on the basis of whether or not they tended to downgrade products after they had received criticism. The children who did so exhibited many important aspects of helplessness, as assessed through measures of affect, task choice, and problem-solving strategies. In addition, they were much more likely to make general negative self-judgments after they had received criticism, including negative judgments of their goodness. Moreover, low product raters were more likely than other children to make general inferences from specific information within the domain of goodness.

**Preliminary Analyses**

Preliminary analyses indicated that the gender of the child and the specific story presented did not influence the results, and therefore these variables were not included in subsequent analyses.
ses also revealed that children responded in the same manner to the feedback stories regardless of whether the no-feedback story was presented first or last. For this reason children's responses to the feedback stories were combined across both presentation orders for further analyses.

**Forced-Choice and Rating Measures**

**Product rating.**—As indicated in Table 1, almost all children (94.4%) made product ratings of 5 or 6 in the no-feedback story when it came first. Since these ratings almost never fell below 5 in the absence of criticism, such ratings were considered to be downgraded as a result of criticism. For this reason, children were divided into groups on the basis of whether or not their average rating in the feedback conditions was less than 5. Thus, the 60.7% of children who gave mean product ratings of 5 or more were categorized as **high product raters** and the 39.3% of children with mean scores below 5 were categorized as **low product raters**.

Overall, there was a great deal of variation in children's product ratings after criticism, ranging from two big frowns on both stories (a mean rating of 1) to two big smiles on both stories (a mean rating of 6). The mean for the high product raters was 5.62 and the mean for the low product raters was 3.36.

Interestingly, an examination of the product ratings of the high and low product raters in the no-feedback judgments revealed that these groups gave themselves equally high ratings ($M = 5.58$ for high product raters and $5.57$ for low product raters), $t(52) = .026$, $p = $ N.S. This indicates that the low product raters, like high product raters, overwhelmingly defined what they had done as a success before they were criticized.

It is clear that when children were criticized, there were strong individual differences in their tendencies to downgrade products. Do such differences map onto (a) subsequent task preferences, (b) affect, and (c) self-judgments? The following analyses were conducted to address these questions.

**Task choice and affect ratings.**—Results indicated that high product raters were more likely than low product raters to agree to persist in the type of activity for which they had been criticized ($M = 1.17$ on a 2-point scale for high product raters and .76 for low product raters, $t(105) = 2.318$, $p < .05$). The groups also responded differently to the affect measure in which children chose faces to represent their feelings. High product raters were more likely to report feeling happy ($M = 1.47$ and .91, respectively, for high and low product raters, $t(105) = 3.69$, $p < .001$), and were less likely to report negative affect. They were less likely to report experiencing sadness ($M = .563$ and 1.00, respectively, for high and low product raters, $t(105) = 2.91$, $p < .005$), and less likely to report feelings of anger$^3$ ($M = .42$, and .81, respectively, for high and low product raters, $t(105) = 2.61$, $p = .01$).

**Self-ratings.**—Since low product raters were particularly likely to make choices and express feelings consistent with helplessness, the generality of their negative thinking was of interest: did they merely make negative product evaluations, or did they call major aspects of themselves into question? Responses to the adequacy measure, involving four ability and trait evaluations (task ability, intelligence, goodness, and niceness), were analyzed in order to answer this question. Each of these four evaluations was made after each feedback story, and children were given 1 point for each question on which they expressed positive evaluations, with a resulting range of scores from 0 to 8. As predicted, high product raters diverged greatly from low product raters; they had an average score of 7.40, whereas low product raters had an average score of 4.66, $t(105) = 6.39$, $p < .001$.

In order to look further at the types of negative inferences that were being made, separate analyses were performed for each of the four components of the measure. Figure 1 depicts the percentage of children who expressed negative self-judgments in response to each of the ability and trait questions in at least one of the two feedback stories. As can be seen from this figure, for every self-rating, low product raters re-

---

$^3$ Anger can be viewed as one measure of negative affect. Caution should be taken in interpreting this result as it is not clear whether the child felt angry at him- or herself or at the teacher.
sponded more negatively than did high product raters. Specifically, they were more likely to say in at least one of the two feedback stories that they felt that they were not good at the task, $\chi^2(1) = 12.78, p < .001$, that they were not smart, $\chi^2(1) = 28.95, p < .001$, that they were not a good person, $\chi^2(1) = 24.59, p < .001$, and that they were not nice, $\chi^2(1) = 19.58, p < .001$.

The proportion of high product raters giving themselves positive versus negative ratings depended on the specific question, $\chi^2(3) = 11.12, p < .025$. When high product raters reported negative feelings about themselves, it tended to be at the most specific level, referring to their ability at the task. In contrast, low product raters were as likely to blame general traits such as lack of smartness or goodness as they were to indict specific abilities. In fact, over half of low product raters did so for smartness (52.5%) and for goodness (61.9%). Over a third (38.1%) of the low product raters reported feeling that they were “not nice” even though the goal of creating something nice for the teacher was made explicit in the story. The differences between the percentages of low product raters giving themselves positive versus negative ratings in the four categories did not differ significantly, $\chi^2(3) = 5.35, p = N.S.$

**Generalization of ratings to a new situation.**—Did the negative ratings of low product raters generalize to new products that had never been criticized? In order to answer this question, product ratings on the pre- and postjudgment stories were compared for high and low product raters. As previously noted, high and low product raters gave nearly equal product ratings on the prejudgment story ($M = 5.57$ for low product raters and $5.58$ for high product raters). In contrast, on the postjudgment story the low product raters made much lower assessments than the high product raters ($M = 3.86$ and $5.13$, respectively). An ANOVA yielded a main effect for order, $F(1,103) = 20.63, p < .001$, a main effect for product rating category, $F(1,103) = 7.12, p < .01$, and an interaction effect between these variables, $F(1,103) = 7.03, p < .01$. These results indicate that for low product raters, but not for high product raters, criticism was associated with substantially more negative product ratings in a new, similar situation in which no feedback was given.

In summary, results of the analyses of the quantitative measures demonstrate that when criticized, a substantial number of children evaluated products in a relatively negative manner. These children were more likely than others to report negative affect and to choose not to participate in the activity for which they had received criticism. In addition, they were more likely to report feeling badly about important aspects of themselves after receiving criticism. Finally, they tended to generalize their negative feelings about the products to a new situation in which no criticism was given.

**Open-ended Questions**

**Coding and reliability.**—Two coders, blind to the classification of the subjects, independently coded children’s responses to each of the questions. Reliabilities for each set of categories were determined by dividing the number of agreements by the number of agreements plus disagreements. Reliabilities ranged from .80 to 1.00. All disagreements were resolved by discussion between the coders.

**Recall.**—Children’s responses to what they remembered about the story were coded as to whether negative feedback was explicitly mentioned. Approximately the
same percentage of high and low product raters reported this feedback (68.3% vs. 66.7%). Thus, it does not appear that differences in responses resulted from differences in the recall of criticism.

Evaluation of the situation.—The questions concerning children's expected parental responses and appropriate teacher responses were coded into mutually exclusive categories based on the evaluative information they contained. Responses were coded as (a) negative if they included any kind of negative feedback and no positive feedback about the child or what the child had done, (b) positive if they included any kind of positive and no negative feedback, (c) mixed if they included both positive and negative feedback, or (d) nonevaluative if there was no positive or negative feedback. Children who gave no answer or said they didn't know were not included in the analysis. This amounted to two children for the parent question and three children for the teacher question. All of the children who did not respond were low product raters.

On the expected parental response measure, the responses of high and low product raters tended to fall in different categories, $\chi^2(3) = 10.47$, $p < .025$. Specifically, high product raters were more likely than low product raters to enact positive evaluations (46.2% vs. 25.0%), $\chi^2(1) = 4.70$, $p < .05$. In addition, about a quarter of high product raters (24.6%) as opposed to almost half of low product raters (47.5%) reported that their parents would respond with a negative evaluation, $\chi^2(1) = 5.84$, $p < .025$. Note that for this measure, as well as for measures of appropriate teacher response and story resolution, overall and then specific chi-square tests were performed. The specific chi-square tests were used to identify group differences in responses to particular categories. These tests of each category were not independent. For example, in the case of the expected parental response measure, the comparison of positive responses from each group is not independent of the comparison of negative responses.

The following are examples of positive and negative parental responses that were enacted:

Positive responses:
- "You did nice."
- "Did you do those terrific numbers?"
- "Good boy."

Negative responses:
- "You did a very bad job."

An analysis of children's perceptions of the appropriate teacher response revealed differences in the responses of high and low product raters, $\chi^2(3) = 9.58$, $p < .025$. Whereas almost two-thirds of the high product raters (64.6%) would have given themselves positive evaluations, only one-third of the low product raters would have done so (33.3%), $\chi^2(1) = 9.57$, $p < .005$. About twice as many low product raters as high product raters would have given negative evaluations (41.0% vs. 21.5%), $\chi^2(1) = 4.51$, $p < .05$.

The following are examples of positive and negative ways in which children felt the teacher should have responded:

Positive responses:
- "That is a good girl."
- "You did a good job. I'm proud of you."
- "That's a very nice picture. Why don't you put some feet on that picture?"

Negative responses:
- "You are very bad."
- "I hate those numbers."
- "How could you do this? This is not a very good building. I don't like it."

These results demonstrate that when children were given a chance to be the evaluators, many low product raters did not respond as though they felt they had been unfairly criticized, but instead reaffirmed the validity of the criticism.

Punishment.—As noted above, following criticism, many low product raters reported feeling that they were bad and deserving of negative evaluations. An analysis of all open-ended questions was done to examine whether low product raters were especially likely to make references to punishment. Although few children explicitly acted out or talked about being punished as a result of what they had done, those who did so were more likely to be low product raters than high product raters. Only 4.6% of high product raters referred to punishment by parents or the teacher, while 16.6% of low product raters did so, $\chi^2(1) = 4.37$, $p < .05$.

Story resolution.—Children's resolutions to the stories in which they received criticism were coded into one of three categories. If children resolved the story by saying they would work to fix or redo what they had done, and saw this as a means to an interpersonal solution with the teacher, this
High Product Raters
Low Product Raters

FIG. 2.—Solutions generated by high and low product raters

was called a constructive solution (considered the highest level solution).\(^4\) If children did not do anything to improve upon what they had done but attempted to deal with the teacher’s disapproval by a task-relevant explanation or apology, this was called an interpersonal solution (considered the intermediate level solution). If children did not address the task-related issues presented in the story, this was called no solution (considered the lowest level solution). This category included children who did not propose any solution and children who acted out bribing the teacher or enacted fantasies that did not address the task-related issues. Because the variable of interest was children’s problem-solving ability in the situation, children were assigned to the category of the highest level of problem solving that they had exhibited on either of the two feedback stories. Results of this coding for high and low product raters are presented in Figure 2. As indicated by these results, high and low product raters were not equally likely to fall into each of the categories, \(\chi^2(2) = 7.70, p < .025\). Specifically, high product raters were more likely to come up with constructive solutions than were low product raters (58.5% vs. 35.7%), \(\chi^2(1) = 5.28, p < .025\), and low product raters were more likely to come up with no solutions or situation-irrelevant solutions (26.2% vs. 52.4%), \(\chi^2(1) = 7.58, p < .01\).

The following are examples that highlight the differences in suggested strategies for resolving the story:

High product raters: One girl acted out her teacher telling her, “You are a good girl but you missed the 8.” In response she said, “That’s O.K. because I didn’t have time,” and said that she would finish it for homework. She then brought the completed work to a pleased teacher.

Another girl acted out finishing her work in the extra time her teacher gave her. She acted out her teacher telling her, “I really like your painting and I know you worked very hard at it, and I’m glad I gave you extra time, and it’s fun having paintings even when they aren’t already done, and you have to spend some extra time, but I know that I still really like your painting.”

Low product raters: One girl said she would go to her room to get a Kleenex (to go cry). A boy enacted himself going home to throw away the numbers he had done.

Stories were also coded for whether they had an unambiguously happy ending. Whereas 69.2% of high product raters presented a story with such an ending, only 40.4% low product raters gave an unambiguously happy ending, \(\chi^2(1) = 8.66, p < .005\). Taken together, these results suggest that high product raters were more likely than low product raters to view the situation as something that could be resolved in a positive and constructive manner.

In summary, analyses of the open-ended responses confirmed and extended the findings from the quantitative measures. Results suggested that following criticism, low product raters were particularly likely to define what they had done as a failure that could not be constructively resolved.

\(^4\)Children did not produce constructive solutions that were not also interpersonal. This is because children who reached these solutions consistently did so in the presence of the teacher or else reported back to the teacher.
TABLE 2
PERCENTAGE OF LOW AND HIGH PRODUCT RATERS ENDORSING THE VIEW THAT (a) NEGATIVE BEHAVIOR PATTERNS ARE LIKELY TO BE MAINTAINED AND (b) THAT ACADEMIC MISTAKES REFLECT BADNESS

<table>
<thead>
<tr>
<th>Product Raters</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2 (a):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes (always bad)</td>
<td>50.0</td>
<td>24.6</td>
</tr>
<tr>
<td>No (not always bad)</td>
<td>50.0</td>
<td>75.4</td>
</tr>
<tr>
<td>Table 2 (b):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes (bad)</td>
<td>47.6</td>
<td>18.5</td>
</tr>
<tr>
<td>No (not bad)</td>
<td>52.4</td>
<td>81.5</td>
</tr>
</tbody>
</table>

General Beliefs

Should the responses of the low product raters be seen as context-specific response patterns, or might they be linked to more general beliefs about goodness? Table 2a contains the proportion of high and low product raters who predicted that a child who exhibits several negative behaviors is likely to always engage in such behaviors. As indicated in the table, 50% of low product raters endorsed this belief, compared to only 24.6% of high product raters, $\chi^2(1) = 7.27, p < .01$.

Table 2b contains the proportion of high and low product raters who reported that poor academic performance by another child means that the child is "bad." Whereas 47.6% of low product raters made this inference, only 18.5% of high product raters did so, $\chi^2(1) = 10.35$, $p < .005$. These results suggest that there is a positive association between responses characteristic of helplessness and the tendency to view "badness" in a stable and global manner.

Did low product raters have a general tendency to view the opinion of adults as being more valid than their own opinion? Might such a tendency explain why they differ from high product raters? In order to address these questions, children were directed to imagine a situation in which they liked some music and their father did not, and were asked who had the appropriate judgment of the music. In this case there was no significant difference between the responses of the two groups: 69.2% of high and 76.2% of low product raters trusted their own judgment, $t(104) = .744, p = N.S.$

Teacher Ratings

Analyses of teachers' reports suggest that low product raters were perceived as somewhat more likely to get upset when criticized than were high product raters. The average score for this measure for low product raters was 2.31 (out of a maximum score of 5), whereas the average score for high product raters was 1.89, $t(104) = 1.97, p = .05$. This provides some evidence that behavior in the experimental setting was related to the child's typical reactions in real-life situations.

It could be argued that the helpless reactions of the low product raters reflected feelings of incompetence that have some basis in fact. The data, however, offer no support for this hypothesis. Low product raters had a mean teacher-rated ability score of 4.19 (out of a maximum score of 5); high product raters had a mean ability score of 4.06, $t(104) = 1.33, p = N.S.$

Discussion

Some of the 5- and 6-year-olds in this study exhibited motivational patterns in response to criticism that were consistent with young children's responses to failure in other studies (Cain, 1990; Hebert & Dweck, 1985; Smiley, 1989). Children in these studies expressed thoughts and feelings and enacted behaviors characteristic of the helpless pattern in older children. Parallel to findings with older helpless children, a group of kindergartners in this study expressed a relatively high degree of negative affect and were relatively unlikely to generate constructive problem solving strategies after encountering obstacles. Also, like older helpless children, young low product raters often viewed their difficulty as reflecting upon their traits and abilities. In fact, these negative evaluations were often even more broad than those that have been reported in studies of older helpless children.

Did the negative judgments of low product raters merely reflect a general tendency to be negative? This does not appear to be the case; before criticism, the groups rated products in an equally positive manner. Alternatively, might the negative judgments reflect a greater acceptance of the views of adults in all contexts? Again, the results suggest otherwise; when an adult held a conflicting opinion concerning the evaluation of music, the high and low product raters were equally likely to trust their own opinions.

In the present study, the toy teacher ex-
pressed disapproval of the products children pretended to make, and disappointment in the child. Consequently, one might argue that the striking pattern of response seen in low product raters in unique to situations in which children receive criticism, and perhaps only highly explicit criticism. Two pieces of evidence call this explanation into question. First, as has already been mentioned, children of this age have shown similar responses when they failed at puzzle tasks but were not criticized (see Dweck, 1991, for summary). Second, preliminary analyses of a follow-up study of 89 subjects suggests no lower incidence of negative reactions when disapproval was expressed for products but not for the child. However, the extent to which adult presence or reaction is necessary for young children (or older children) to show the affective and behavioral responses associated with helplessness is an important issue that remains to be addressed.

The differences between high and low product raters were not limited to situation-specific tendencies to become mired in the negative feedback. Rather, the groups also differed in their judgments of goodness and badness. Specifically, low product raters were more likely to view mistakes in the classroom as evidence that a child is bad. In addition, low product raters were more likely to expect a child who exhibits negative behaviors to always persist in such behaviors. It may be that these children tend to view badness in a nonmalleable way, just as older children who exhibit the helpless response pattern tend to view intelligence in this way. It should be noted, however, that this study does not distinguish between the child's notion that a person's status as good or bad is fixed and unchangeable versus merely stable and unlikely to change. In addition, it is not known whether children tend to interpret the word "bad" differently than do adults, and whether high and low product raters think about the word differently. Further research is needed to clarify what specific meanings children's evaluative terms carry. Nevertheless, striking differences were found in the tendencies of high and low product raters to make broad inferences from limited information.

Do general beliefs about goodness actually influence motivational response? This question remains to be addressed in future research. It is plausible to hypothesize that when children have a tendency to make broad trait inferences from a small set of behaviors, or when they infer stable behavior patterns from a small sample of outcomes, they may be vulnerable to sweeping inferences when they themselves experience these outcomes or exhibit these behaviors.

An important question for future research concerns what aspects of the child's social world lead to enhanced or diminished ability to cope with criticism and other negative outcomes. The parent-child relationship would be a good starting point for further investigation of this issue, since low product raters were relatively more likely to enact parental criticism. These findings suggest that the extent to which parents are judgmental and critical play a role in creating maladaptive motivational patterns. The quality of the parent-child relationship in general may also play a role. Children may come to know what negative outcomes mean about themselves from their understanding of important relationships. For example, if a child feels the love they receive from their parents is based on always being "good," negative outcomes may take on a powerful meaning. Some support for this possibility is provided in the attachment literature, which indicates that 6-year-old children who are "securely" attached tend to give positive, realistic, and flexible responses to self-evaluation questions, whereas children who are "insecurely" attached tend to give either negative evaluations or defensive evaluations in which they do not acknowledge their faults (Cassidy, 1988). Under this framework, one could ask whether low product raters correspond to the insecurely attached children who give negative self-evaluations and whether the subset of high product raters who did not come up with constructive solutions to the story dilemmas correspond to insecurely attached children who give defensive self-evaluations.

This study raises many questions about the development of children's motivation. Does a tendency to be vulnerable to criticism early in life predispose a child to be self-critical or adopt overly harsh standards later in the absence of external criticism? The young children in this study who gave themselves low ratings in the face of criticism were not likely to be critical of themselves prior to receiving criticism. However, it may be that vulnerability to the criticism of others is a precursor to being self-critical in other contexts.

Another developmental question that emerges from this study concerns how older
and younger children differ in their tendency to become helpless. As is evident from several studies comparing the ways in which children of different ages respond to failure (see Parsons & Ruble, 1977; Rholes et al., 1980; Ruble et al., 1976), there are contexts in which younger children are less motivationally vulnerable than older children. For example, it may be that when tasks failures are not salient, young children exhibit more motivational resiliency than older children; they may only rarely have internalized the stringent performance standards of many older children and adults. Although in the present study no direct comparison with older children was made, young children were in fact only rarely bothered by product flaws prior to the criticism, whereas older children might well have been more self-critical.

Although young children may be less motivationally vulnerable than older children in some ways, they may be more vulnerable in other ways. The present research suggests that for young children negative outcomes are often associated with negative feelings about global aspects of the self. Further development and differentiation of trait concepts may allow older children to compartmentalize different aspects of the self (see Benenson & Dweck, 1986), and therefore see failure or criticism as implicating more specific aspects of the self, rather than the self as a whole. It may be that in the face of negative achievement outcomes, older helpless children tend to blame their intellectual abilities, but are less likely to feel that they are "bad." In addition, young children may be more willing to accept a negative judgment, even if this judgment conflicts with other sources of feedback, because of a special concern with social reinforcement and a less developed understanding of other potential standards of judgment (see Stipek, 1984). Research examining children's changes in self-judgments within achievement contexts would be especially useful in answering some of these questions.

Although concerns of low product raters surrounding the domain of goodness have been emphasized, it is important to point out that this domain was salient to children in both groups; a large proportion of high and low product raters made multiple references to goodness or badness. However, the nature of these references tended to differ. Whereas the high product raters seemed to take the criticism they had received for a flawed product to mean they were bad. The following is an example role-play of a boy who seemed to be in conflict over these two points of view, locating these different points of view within each of his parents:

Mother: Hello. What are you sad about?
Subject: I gave my teacher some numbers and I skipped the number 8 and now I'm feeling sad.
Mother: Well, there's one thing that can cheer you up.
Subject: What?
Mother: If you really tell your teacher that you tried your best, she wouldn't be mad at you. We're not mad, aren't [sic] we?
Father: Oh [yes] we are!

Unfortunately, this subject appeared to accept the view he attributed to his father, as his product ratings and self-ratings were extremely low.

In summary, the present study provides evidence that after receiving criticism, some kindergartners show affective reactions and self-evaluations associated with motivational helplessness, and that these reactions are related to conceptions of goodness. The ways in which ideas about goodness develop and are implicated in motivational processes point to interesting directions for future research.

Appendix

Text of Stories Presented to Children

Story A: You spend a lot of time painting a picture of a family to give to your teacher. You pick out colors you think are nice and carefully draw each person. As you are about to give it to your teacher you say to yourself, "Uh oh, one of the kids has no feet." But you worked really hard on the picture and want to give it to her. You say, "Teacher, here's a picture for you."

Feedback: There are no feet on that child. That's not what I call drawing people the right way. I'm disappointed in you.

Story B: You spend a lot of time building a house out of blocks to give to your teacher. You carefully pick out which blocks you should use and how to fit them together to make the house look really nice. After it is built you say to yourself, "Uh oh, there are no windows in the house."

Feedback: There are no windows on that
house. That's not what I call building a house the right way. I'm disappointed in you.

Story C: You are learning how to write your numbers and want to write out the numbers 1 to 10 to show your teacher. You carefully write each number and think about what number should go next. After you finish the numbers you say to yourself, "Uh oh, I skipped the number 8." But you worked really hard on the numbers and want to have your teacher see them. You say, "Teacher, come look at the numbers I did."

Feedback: The number 8 is missing. That's not what I call writing numbers the right way. I'm disappointed in you.

References


Stipek, D., & Daniels, D. (1988). Declining per-
ceptions of competence: A consequence of changes in the child or in the educational environment? *Journal of Educational Psychology, 80*, 352–356.


Heyman, Dweck, and Cain 415


This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.