Perceived Contingency of Skill and Chance Events: 
A Developmental Analysis

John R. Weisz, Keith Owen Yeates, Dan Robertson, and Jean C. Beckham
University of North Carolina at Chapel Hill

Outcomes in skill-controlled activities are contingent on people's attributes and behavior. Outcomes of chance activities are inherently noncontingent. In this study, we explored age differences in the recognition of this distinction between skill and chance. Kindergarteners, fourth graders, eighth graders, and college students took part in one game of chance and one game of skill. After each game, subjects predicted the winnings of other players who differed in certain attributes (e.g., intelligence) and behavior (e.g., effort) that would influence only skill outcomes. Two developmental trends emerged: (a) On both chance and skill tasks, older subjects expected the variations in attributes and behavior to have less impact on task outcomes than did younger subjects; and (b) older subjects were more adept at making predictions that reflected the contingency of skill and the noncontingency of chance. Kindergarteners showed no ability to make the skill-chance distinction. Fourth graders were aware of the distinction at a gross qualitative level, but their predictions showed that they were unaware of some of the most important logical implications of that distinction. Eighth graders and college students were aware of the skill-chance distinction and most of its logical implications; yet their predictions revealed a lingering belief that chance outcomes could be influenced slightly by variations in people's attributes and behavior. This last finding is interpreted in light of Rosch's (Rosch & Mervis, 1975) model of category formation.

To make accurate judgments about what we can and cannot control, we must accurately judge the contingency of outcomes. If an outcome (e.g., rain, a roll of the dice) is not contingent on people's behavior, then we will be unable to control it. (For more details on determinants of control, see Weisz & Stipek, 1982.) There is mounting evidence that such noncontingency is very difficult for young children to detect. Piaget (1930) reported that children younger than 6 or 7 years show "primitive psychological causality," that is, "the belief that any desire whatsoever can influence objects, the belief in the obedience of external things" (p. 303). The young child, as described by Piaget, has the illusion that diverse physical events in the natural world (e.g., movement of the wind or of heavenly bodies) are contingent on either the child's own behavior or on other people's behavior. In Piaget's (1932) writings on moral reasoning, he described a phenomenon called "immanent justice"; young children, he said, perceive people's physical accidents and misfortunes as being contingent outcomes, that is, consequences of a violation of moral precepts or social norms. In Piaget and Inhelder's (1975) writings on chance concepts, they reported that young children do not recognize the fortuitous nature of random events and thus often perceive even totally chance outcomes as contingent on variations in human behavior. The Piagetian literature describes illusory contingency in all three areas—causality, moral reasoning, and chance—as declining over the course of cognitive development.

Unfortunately, the Piagetian literature in all three areas is heavily anecdotal; it contains little evidence from carefully controlled experiments. Recently, though, more controlled experiments have been conducted to specifically explore judgments about chance.
events. In one of these experiments, Weisz (1980) used a card game designed to appear totally chance controlled. Kindergarteners and fourth graders drew cards blindly from a shuffled deck, and the color of the card drawn determined winnings. Children predicted the winnings of other players who differed from one another in age, intelligence, effort, and previous practice at the task. Predictions that differed as a function of any of these factors were interpreted as evidence of illusory contingency. Such illusory contingency was significantly more pronounced in younger than in older children. Younger children also interpreted very high or very low winnings by fictitious other children as evidence of high or low levels of concentration or other contingency-relevant factors. In contrast, older children interpreted high or low winnings as resulting from luck. Yet even older children showed signs of a lingering illusion of contingency. For example, they predicted higher winnings for a child who tried hard than for one who did not. A second study (Weisz, 1981) yielded a similar pattern of findings among youngsters playing games of chance at a state fair.

In both recent studies, when the older subjects predicted smaller differences as a function of the contingency factors (age, effort, etc.) than did younger subjects, this was interpreted as evidence that the older subjects were less susceptible to illusory contingency. An alternative interpretation is that older subjects were merely showing a more conservative, predictive style. Older individuals, compared to younger ones, may simply expect smaller differences as a function of age and the other contingency factors—regardless of whether the target activity is a chance or skill task. Thus, the key age group difference found in both studies (Weisz, 1980, 1981) may have actually resulted from a developmental difference in response style, not a developmental difference in recognition of noncontingency. This interpretation is given added credibility by research (e.g., Langer, 1975; Lerner, 1977) that suggests that beliefs in the controllability or contingency of purely chance events persist into adulthood. So the question remains: Does development really lead to a decline in illusory contingency?

The present study was designed to answer this question more definitively. People at four age levels made contingency judgments about chance and skill outcomes. We reasoned that if developmental effects were similar for chance and skill outcomes, then the effects could properly be regarded as evidence of response style differences, not differences in recognition of noncontingency. However, if the kinds of age group differences found by Weisz (1980, 1981) were more pronounced with chance tasks than with skill tasks, a developmental increase in recognition of noncontingency would be indicated.

Method

Subjects and Design

Subjects were kindergarten, fourth-grade, and eighth-grade public school students and introductory psychology students at the University of North Carolina at Chapel Hill. The younger groups had 10 boys and 10 girls each. The college sample had 16 men and 16 women. At each age level, half the male and half the female subjects received a chance task, then a skill task; the other half had the reverse order. The groups formed a 4 (age) X 2 (gender) X 2 (order) factorial design with proportional cell frequencies.

Chance Task

The chance task involved 12 white 7 X 12 cm cards. Six cards had a blue circular spot on one side; six had a yellow circular spot on one side. The experimenter said that he would place pairs of cards face down and that in each pair one card would have a blue spot, the other a yellow spot. The subject was to point to one card each time; when that card had a blue spot (yellow for half the subjects), the subject would be awarded one chip from a row of six that had been placed in front of the subject.

Six trials followed. For each trial, the experimenter shuffled the cards twice, then tossed two cards face down on the table. The cards were casually tossed into a variety of positions, never perfectly parallel. The intent was to avoid cues that might suggest a detectable pattern and, thus, a controllable task. By surreptitiously arranging for the two cards on each trial to always have a spot of the same color, the experimenter ensured that all subjects won four chips (win-loss order randomized for each subject). The procedure seemed easily understood, and no subject showed signs of suspicion.

Skill Task

The skill task was similar. Subjects were shown pairs of cards with colored spots face down; they pointed to one and won a chip when the designated color (blue or yellow) was chosen. But the skill procedure was designed to make skill factors, particularly memory and concentration, appear relevant to outcomes. The experimenter
spread out six white 7 × 12 cm cards. At the top of each was a line drawing adapted from Kagan's (1965) Matching Familiar Figures Test. Each drawing differed from the others in minor details. At the bottom of each card was a blue or yellow spot. Each subject was told to study the cards and to try to remember which pictures went with blue and which with yellow. A 30-sec study period was given for each of six sets of cards. After all six sets had been studied, six chips were spread in front of the subject, and six trials (one for each set of cards) were given. In each trial, the experimenter put two of the six cards from a set on the table. Line drawings were on the visible side; colored spots were on the reverse side. The subject's job was to pick the card that had a blue spot (yellow for half the subjects). The experimenter surreptitiously arranged for each pair to have both blues or both yellows; this ensured that every subject won on four trials in an order differing from the chance task. Subjects at all age levels seemed invested in concentrating and memorizing during the study periods and in searching their memories during the trials. We had tried to structure the task so that subjects would believe that their concentration and memory played a role in outcomes but would rarely be certain about any particular memory. To satisfy this latter condition, we required storage of 36 memories, all concerning minor details of complex pictures. The condition seemed to be satisfied; no subject showed signs of suspicion during or after the task.

**Posttask Questions**

After the skill task and again after the chance task, the experimenter asked,

(a) Now that you've finished this card game, why do you think you won four chips out of six tries? (b) Some of the times you got the right color in this card game, and some of the time you didn't. Why does that happen? (c) I once played this game with another student and she didn't win even one chip—she picked the wrong color every time. Why do you suppose that happened? (d) I once played this game with another student and he won all six chips—he got the right color every time. Why do you suppose that happened? (e) Suppose I put all the cards back together again and we played this game all over again... out of these six chips (line up on the table) show me how many you would win next time. (This question was included for experience in using the chips to make predictions.) (f) Suppose that a ______ (person one grade lower than the subject, e.g., "nursery school boy") and a ______ (person one grade higher than the subject, e.g., "first grade boy") were to play this card game. How many chips do you think the ______ (lower grade person) would win? Show me with this row (first row of chips). How many chips do you think the ______ (higher grade person) would win? Show me with this row (second row).

In similar fashion, the experimenter asked for predictions for a person who

(g) got to practice lots of times before trying to win chips vs. one who didn't get to practice any times; (h) was very careful and tried real hard to pick blue (or yellow cards) vs. one who didn't try hard at all; (i) was very smart vs. one who wasn't so smart.

Questions f-i were randomly ordered for each child.

After all questions had been asked both times, the experimenter placed stimulus materials for the two games side by side on the table. Then he said,

Now think of the two games you've played. In this game (pointing to chance game), you just picked cards. In this other game (pointing to skill), you had to pick cards with pictures. In which of the two games (a) Does it help to be smart? (b) Would you get better as you get older? (c) Does it help to practice a lot? (d) Does it help to try hard? (e) Is it just a matter of luck? (f) Does it help to remember well? (g) Is it just guessing?

After each question, the subject pointed to one game.

**Results**

**Unstructured Questions**

Subjects' responses to the first four questions (see Table 1) for both skill and chance activities were coded as either perceived contingency (e.g., "I tried hard"), perceived non-contingency (e.g., "It's just luck"), or not codable (e.g., "I don't know," "It just happened."). Two independent raters achieved 88% agreement in their judgments. As Table 1 shows, younger subjects gave more uncodable responses than older subjects; these data should, therefore, be viewed cautiously. For each question, responses for skill and chance combined were classified as either a logical pattern (i.e., skill activity contingent, chance activity noncontingent) or an illogical pattern (all other combinations). A 4 (age) × 2 (logical vs. illogical pattern) chi-square analysis was computed separately for each question. The results revealed no significant relation between age and pattern of judgment for the two questions about personal outcomes, both p s > .05 (see Table 1). The two questions concerning the outcomes of fictitious others, however, both showed significant relations between age and judgment pattern, with older subjects increasingly more likely to give logical patterns of causal judgment than younger subjects, both χ² s > 10, both p s < .01. Although judgments about self (Questions 1 and 2) did not show significant effects, whereas judgments about others (Questions 3 and 4) did, Table 1 shows that frequency of logical patterns tended to increase with age for all four questions; the
Table 1
Percentage of Children Making Logical Patterns of Contingency Judgments on Unstructured Questions

<table>
<thead>
<tr>
<th>Group</th>
<th>Kindergarten</th>
<th>4th grade</th>
<th>8th grade</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Why subject won 4 of 6?</td>
<td>33 (6)</td>
<td>67 (9)</td>
<td>69 (16)</td>
<td>79 (24)</td>
</tr>
<tr>
<td>2. Why subject guessed some right and some wrong?</td>
<td>50 (4)</td>
<td>82 (11)</td>
<td>69 (13)</td>
<td>73 (21)</td>
</tr>
<tr>
<td>3. Why another person won no chips?</td>
<td>22 (9)</td>
<td>62 (13)</td>
<td>79 (12)</td>
<td>87 (23)</td>
</tr>
<tr>
<td>4. Why another person won all six chips?</td>
<td>10 (10)</td>
<td>57 (14)</td>
<td>73 (15)</td>
<td>86 (21)</td>
</tr>
</tbody>
</table>

Note. Grade level differences were significant for Question 3 and 4 (both ps < .01) but nonsignificant for Questions 1 and 2 (both ps > .05). The number of subjects in each group who gave codable responses is shown in parentheses after each percentage.

The mean percentage of logical patterns across all four questions for the four age groups was 29%, 67%, 72%, and 81%, respectively. This suggests a developmental increase in the ability to make accurate contingency judgments about chance and skill outcomes.

**Paired-Prediction Questions: Magnitude of Perceived Effects**

Paired-prediction questions were analyzed using a 4 (age) X 2 (activity type, i.e., skill vs. chance) X 2 (sex) X 2 (order) X 4 (characteristic) X 2 (trial) repeated measures design. We used the multivariate approach to repeated measures that was described by McCall and Appelbaum (1973). Activity type, characteristic, and trial were all repeated measures. The four levels of the characteristic factor were the areas of variation (i.e., age, intelligence, effort, and practice); the two levels of the trial factor were the separate judgments in each characteristic area (e.g., smart vs. not smart). The large number of significance tests posed a substantial risk of chance findings. To avoid capitalizing on such findings, we used a Boneferroni correction (Neter & Wasserman, 1974) to set an alpha level appropriate to the number of tests performed in the initial analyses. In this case, the appropriate level was .001.

The analysis yielded a significant trial main effect. Subjects expected older, intelligent, effortful, and practiced players to win more than their less able or motivated counterparts F(1, 76) = 281.73, p < .001 (see Figures 1 and 2 and Table 2). In addition, a significant Trial X Characteristic interaction revealed that intelligence, effort, and practice...
were expected to be more influential than age, $F(3, 74) = 91.89, p < .001$ (M trial differences: 1.3 for age, 2.4 for intelligence, 2.6 for effort, 2.4 for practice). The analysis showed a significant Activity Type $\times$ Trial interaction; variations in contingency factors were perceived to be more important for skill than for chance, $F(1, 76) = 51.09, p < .001$. An Activity Type $\times$ Characteristic $\times$ Trial interaction indicated that effort, intelligence, and practice variations were expected to be particularly more important for the skill than for the chance activity, $F(3, 74) = 13.64, p < .001$; the Trial $\times$ Activity Type interaction was significant for practice, effort, and intelligence (all $p$s < .001) but nonsignificant for age ($p > .20$). The latter interactions are further evidence that the activities were perceived to represent skill and chance, respectively.

Further findings reflected the age differences in response style discussed earlier. A significant Age $\times$ Trial interaction indicated that younger subjects expected variations in contingency factors to be more influential than did older subjects, $F(3, 76) = 8.87, p < .001$ (M trial differences: 3.38 for kindergarteners, 1.8 for fourth graders, 1.35 for eighth graders, 1.13 for college students). A significant Age $\times$ Characteristic $\times$ Trial interaction showed that younger subjects expected intelligence and effort to be especially more important than did older subjects, $F(9, 180.427) = 8.46, p < .001$. To assess whether these age group differences in the expected influence of contingency factors existed for both chance and skill activities, further tests were conducted separately for both domains in each characteristic area. For the chance activity, the predicted differences decreased significantly and monotonically with increasing age in all four characteristic areas, all $F$s(1, 76) > 20, all $p$s < .001; for the skill activity, the predicted differences decreased significantly and monotonically with age for variations in age, practice, and intelligence, all $F$s(1, 76) > 5, all $p$s < .05, and monoton-

![Figure 2](image.png)

**Table 2**

Predicted Winnings on Paired Prediction Questions

<table>
<thead>
<tr>
<th>Characteristics of hypothetical child</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K</td>
</tr>
<tr>
<td>Chance tasks</td>
<td></td>
</tr>
<tr>
<td>Older</td>
<td>4.90</td>
</tr>
<tr>
<td>Younger</td>
<td>2.40</td>
</tr>
<tr>
<td>Smart</td>
<td>5.85</td>
</tr>
<tr>
<td>Not smart</td>
<td>1.65</td>
</tr>
<tr>
<td>Try</td>
<td>5.65</td>
</tr>
<tr>
<td>Not try</td>
<td>1.90</td>
</tr>
<tr>
<td>Practice</td>
<td>5.20</td>
</tr>
<tr>
<td>No practice</td>
<td>1.90</td>
</tr>
<tr>
<td>Skill tasks</td>
<td></td>
</tr>
<tr>
<td>Older</td>
<td>5.20</td>
</tr>
<tr>
<td>Younger</td>
<td>2.80</td>
</tr>
<tr>
<td>Smart</td>
<td>5.65</td>
</tr>
<tr>
<td>Not smart</td>
<td>2.05</td>
</tr>
<tr>
<td>Try</td>
<td>5.35</td>
</tr>
<tr>
<td>Not try</td>
<td>1.85</td>
</tr>
<tr>
<td>Practice</td>
<td>5.80</td>
</tr>
<tr>
<td>No practice</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Note. K = kindergarten; 4 = fourth grade; 8 = eighth grade. Numbers in brackets are paired prediction differences significant by Scheffé test, $p < .05$. 

---

Figure 2. Predicted winnings at the chance and skill tasks of (top) a person who tries hard versus one who does not try hard and (bottom) a person who has practiced versus a person who has not practiced.
ically but marginally for variations in effort, \( F(1, 76) = 3.56, p < .07 \). These findings suggest the very age group differences in predictive style discussed in the introduction. Older subjects predicted smaller differences than younger subjects on both chance and skill tasks.

The following analyses were conducted to determine whether this predictive style difference could fully account for all age group differences found here. Developmental differences in the ability to discriminate between contingent and noncontingent outcomes could be inferred only if older children predicted larger differences due to contingency factors on skill activities than on chance activities, whereas younger children predicted substantial differences due to contingency factors for both types of activities, that is, if the magnitude of the skill–chance difference was greater among older than younger subjects. There was evidence for such a pattern: The Age \( \times \) Activity Type \( \times \) Trial interaction was significant, \( F(1, 76) = 13.36, p < .001 \). As Figures 1 and 2 show, the age group differences in predictions based on variations in practice, intelligence, and effort were significantly larger for the chance activity than for the skill activity: All \( F_s(i, 76) > 10, all \ p_s < .001 \). The age variation showed no such difference, but this may well have resulted from the small differences predicted as a function of age variations in general; older subjects’ comments indicated that they did not expect such a small variation in age to influence performance on either chance or skill tasks very much. On further reflection, we agree with their comments. The age variation questions were problematic in two ways: By asking subjects to make predictions for players one grade level higher and one grade level lower than themselves, we were (a) asking subjects at different grade levels to respond to different questions and (b) presenting our younger subjects with age differences that could matter a great deal in skill tasks while presenting older subjects with relatively unimportant age differences.

To further examine the results, Scheffe tests were computed for each pair of trial means in each characteristic area separately for each age group in both chance and skill activities. Table 2 reveals the following pattern: Kindergarteners expected age, intelligence, effort, and practice to influence both chance and skill tasks (all \( p_s < .05 \)); fourth graders expected intelligence, effort, and practice to influence chance outcomes and all four factors to influence skill outcomes (all \( p_s < .05 \)); eighth graders and college students expected none of the factors to influence chance outcomes, but they expected intelligence, effort, and practice to influence skill outcomes (all \( p_s < .05 \)). These findings strongly suggest a developmental increase in the ability to distinguish between contingent and noncontingent outcomes.

**Paired-Prediction Questions: Incidence of Perceived Effects**

The findings thus far show a marked superiority of older subjects over younger ones. However, we must also ask whether even older subjects completely understood the nature of noncontingency, that is, did they carry their judgments to a logical endpoint and predict equal winnings for the chance activity regardless of variations in contingency factors while also predicting different winnings on the skill activity as a function of such variations? To answer this question, we combined subjects’ predictions in each characteristic area for chance and skill activities, and we classified their prediction patterns as either logical (i.e., different for skill, equal for chance) or illogical (all other combinations). A 4 (age) \( \times \) 2 (logical vs. illogical pattern) chi-square analysis for variations in age was not significant; but variations in intelligence, effort, and practice all showed significant relations between age and pattern of prediction: All \( \chi^2 s(3) > 25, all \ p_s < .001 \). Across these three characteristic areas, the mean percentage of logical prediction patterns for kindergarteners, fourth graders, eighth graders, and college students was 0%, 7%, 53%, and 53%, respectively. Thus, most of the older subjects were able to carry their judgments about contingency to a logical endpoint, but the size of the majority was modest indeed.

**Comparison Questions**

Finally, we contrasted the answers given by each of the age groups to the comparison
questions asked at the end (e.g., "In which of the two games does it help to remember well?"). The number of correct responses was summed to provide a score (possible range: 0 to 7). An analysis of variance (ANOVA), with age as an independent variable, yielded a significant main effect, $F(3, 79) = 42.82$, $p < .001$. Scores improved monotonically with age ($M$s: 3.3, 5.4, 6.8, 7.0). In fact, the percentage answering correctly increased monotonically for each of the seven questions (all seven $\chi^2$ values $> 19$, all $p$s $< .001$).

Discussion

By comparing people's reasoning about chance and skill outcomes, we were able to test for the existence of developmental differences in predictive style independently of differences in perceived contingency per se. The evidence pointed to both kinds of differences. As age level increased, predictive style grew more conservative; for both chance and skill tasks, older subjects expected age, intelligence, effort, and practice to have more modest effects than did younger subjects. But this developmental effect was markedly sharper for chance than for skill outcomes. In fact, younger subjects generally did not distinguish skill from chance outcomes in their predictions, whereas older subjects did. So, with increasing age, we saw both a general reduction in the perceived effects of contingency factors on both skill and chance tasks and a general sharpening of the distinction between chance and skill. This latter effect was seen clearly on the final questions, in which skill and chance outcomes were directly compared. So, over and above age differences in predictive style, there are evidently real age differences in recognition of noncontingency.

The results can be summarized in terms of grade levels. Kindergarteners showed little evidence of an ability to distinguish contingent (skill) from noncontingent (chance) outcomes in any way. Fourth graders could make gross qualitative distinctions between the two types of outcome; they answered some of the comparison questions correctly, and they could generally identify their chance outcomes as "luck." Yet, when asked to predict outcomes on the chance tasks, fourth graders failed to fully appreciate the quantitative implications of noncontingency; that is, they illogically expected variations in contingency factors to influence the noncontingent outcomes. Only at the eighth grade and college levels did we begin to see substantial numbers of people correctly identifying the causes of their outcomes, correctly distinguishing contingent from noncontingent tasks on all posttask comparison questions, and correctly predicting no effects of contingency factors on noncontingent outcomes.

In a notable departure from ideal information processing, however, more than 45% of both eighth graders and college students illogically predicted that at least one of the contingency factors would influence outcomes on a chance task. The persistence of such illusory contingency among adults is difficult to explain if we assume that adults define noncontingent events logically as, say, events not influenced by variations in human behavior and that they base their predictions on such a definition. In contrast to this rather traditional view of category definition, however, consider the perspective of Rosch (see, e.g., Rosch & Mervis, 1975). Building on the work of Wittgenstein (1953), Rosch has argued that people assign many events and objects to categories not by definitionally listing criterial attributes but rather by judging the similarity of these events and objects to an abstract representation, or "prototype," of the category. To explain our findings, one might argue that many adults carry in mind an abstract prototypical "contingent event" and a contrasting prototypical "noncontingent event" and that these adults base contingency judgments and predictions on the perceived similarity of a target event to the two prototypes. If the target event bears a marked similarity to the contingent-event prototype, then predictions will reflect a high level of perceived contingency (e.g., substantial perceived effects of an actor's level of effort). If the target event bears only a slight resemblance to the non-contingent-event prototype, then the adult involved may classify the
target event as caused by noncontingent forces yet still predict in a way that reflects a modicum of perceived contingency (e.g., predicting that an actor's level of effort would have a very slight effect on outcomes). Even a few inadvertent skill cues might be enough to trigger this latter process, and such cues are virtually always present in chance activities (see Langer, 1975). In the present study, for example, the very fact of being given information on effort, ability, and so forth and then being asked to predict may have suggested that such information was relevant to task outcomes; this, in turn, according to the model just presented, could have caused the chance task to resemble, however faintly, a contingent-event prototype that some of our adults relied on when making predictions (for related lines of reasoning, see Kahneman & Tversky, 1973, on the "representativeness heuristic" and the "availability heuristic").

Although traces of illusory contingency persisted into young adulthood, our data did reveal strong developmental effects. These effects suggest that young children show a much stronger tendency than older persons to overestimate the degree to which outcomes are contingent on human attributes and behavior. Other research (reviewed by Stipek, in press) indicates that young children also show a strong tendency to overestimate their own competence at various tasks. Judgments about the contingency of outcomes and about one's personal competence are both required for an accurate assessment of one's capacity to control a given outcome (see Weisz & Stipek, 1982, for a discussion of contingency and competence judgments and their role in perceived control). It seems likely, then, that young children will be relatively poor judges of their capacity to exercise control and that they will tend most often to overestimate that capacity. As a result, they might be expected to be less susceptible than older individuals to perceptions of a lack of control, as manifested in such states as "learned helplessness." Recent evidence (Rholes, Blackwell, Jordan, & Walters, 1980) supports that conclusion.

References


Received December 21, 1981