Developmental Psychology
1995, Vol. 31, No. 5, 866-876

0012-1649/95/$3.00

Relations Between Children's Overestimation of Their Physical Abilities and Accident Proneness

Jodie M. Plumert
University of Iowa

Two experiments examined developmental changes in the accuracy of children's judgments about their physical abilities. Experiment 1 showed that 6- and 8-year-olds overestimated their ability to perform tasks just beyond and well beyond their ability. Adults only had difficulty making judgments about tasks just beyond their ability. Experiment 2 investigated how experience with performing activities influences judgments about physical abilities. Six-year-olds again overestimated their ability to perform tasks just beyond and well beyond their ability. Eight-year-olds were more accurate about tasks well beyond than just beyond their ability. In both experiments, overestimation of ability was associated with accidental injuries for 6- but not for 8-year-olds. The discussion focuses on children's overestimation of physical abilities and the relation between overestimation and accident proneness.

Promoting children's safety and health are concerns shared by pediatricians, developmental psychologists, and educators. Accident prevention clearly plays an integral role in children's health, as accidents are the leading cause of death in children under age 18 (Rodriguez & Brown, 1990). The alarming statistics on children's accidents have led to investigations by researchers on issues such as pedestrian safety (Christoffel et al., 1986; Lee, Young, & McLaughlin, 1984), childhood drowning (Nixon, Pearn, Wilkey, & Corcoran, 1986), bicycling safety (Langley, Silva, & Williams, 1987), and children's ability to operate motorized vehicles (Pick, Plumert, & Arterberry, 1987). Although recent overviews of strategies for reducing childhood injury have called for a better understanding of the underlying factors that contribute to the occurrence of injuries (Brooks & Roberts, 1990; Peterson & Mori, 1985; Roberts, 1986), little is yet known about how developmental changes in cognitive and perceptual skills contribute to unsafe behavior (for an exception, see Coppens, 1986).

One perceptual-cognitive skill that may play an important role in children's safety is the ability to evaluate one's level of skill in relation to the demands of the task (see also Lee et al., 1984). When deciding whether it is safe to cross a street, for example, children must take into account both the speed of oncoming cars and how quickly they can walk or run. According to J. J. Gibson (1979), adaptive behavior within the environment depends on perceiving *affordances,* or the fit between one's own physical characteristics and the properties of the environment in which actions take place. Although accidents are complex phenomena and undoubtedly have several root causes, errors in judging the relation between one's physical abilities and the demands of the situation may be one important factor contributing to accident risk. For example, although some pedestrian accidents may result when children fail to follow simple rules like looking both ways when crossing a street, others may result when children make errors in judgment about their ability to run through traffic gaps.

Studies of the ability to make judgments about the fit between one's own skills and the characteristics of the environment have shown that even infants adjust their actions in response to changing environmental circumstances. For example, visual cliff studies have shown that crawlers refuse to cross over the deep side, but readily venture out over the shallow side (Gibson & Walk, 1960). Later studies demonstrated that walking infants shift from walking to crawling when presented with a nonrigid surface such as a waterbed (Gibson et al., 1987). Adolph, Eppler, and Gibson (1993) also found that walking infants changed their means of locomotion from walking to climbing or sliding as the slope of the surface increased. Likewise, McKenzie, Skouteris, Day, Hartman, and Yonas (1993) found that infants exhibited progressively more leaning during their reaches as objects became increasingly distant. Thus, when environmental circumstances change, it appears that even infants readily modify their actions to reach a goal.

Despite the fact that infants and children show remarkable skill in adjusting their actions to provide a better fit with the demands of the situation, studies also have shown that they often overestimate what they can accomplish with those actions (Adolph, 1995; Adolph et al., 1993; McKenzie & Forbes, 1992). For example, Adolph et al. (1993) and Adolph (1995) found that although toddlers were less likely to go down than up slopes, they consistently overestimated their ability to ascend and descend slopes that were beyond their ability. Likewise, McKenzie et al. (1993) found that even when objects were well out of reach, infants attempted to grasp the objects. They noted, in fact, that many infants had to be rescued from falling out of the infant seats while trying to grasp objects that were well out of

This research was supported in part by a University of Iowa Junior Faculty Seed Grant. I thank the following individuals for their assistance in data collection and coding: Christopher Carswell, Kathy DeVet, Susan English, Kimberly Ewert, Christopher Hulleman, Damien Ihrig, Kari Krenz, and David Strahan. Thanks also go to Lisa Oakes and Grazyna Kochanska for their helpful comments on the manuscript.

Correspondence concerning this article should be addressed to Jodie M. Plumert, Department of Psychology, 11 SSH East, University of Iowa, Iowa City, Iowa 52242.

reach. McKenzie and Forbes (1992) also found that 9- and 12-year-old boys consistently overestimated the height of steps that they could climb (for another viewpoint see Pufall & Dunbar, 1992). Moreover, although adults generally are quite accurate at judging reachability or climbability, studies have shown that they too, have a tendency to overestimate their abilities. For example, Carello, Grosofsky, Reichel, Solomon, and Turvey (1989) found that adults overestimated their ability to reach objects across a variety of postures. Thus, it appears that even adults do not always accurately perceive the boundary between actions that are within and beyond their ability.

Confidence ratings about judgments and latencies to begin movements also suggest that children and adults may perceive the boundary between actions that are within and beyond their ability as fuzzy. For example, Warren (1984) found that adults' confidence ratings about whether steps were climbable plunged sharply at the point at which steps shifted from being climbable to not climbable. Thus, adults were less confident about their judgments at distances closer to the boundary than at distances farther from the boundary. Similarly, Adolph et al. (1993) found that 14-month-olds hesitated longest before starting down 20° slopes presumably because they were unsure of whether to walk or slide down the slopes. In support of this interpretation, approximately half of the infants chose to walk and half chose to slide down slopes of 20°. In contrast, almost all infants walked down slopes of 10° and slid down slopes of 30°.

Studies also suggest that experience with performing actions may influence children's ability to make judgments about whether an activity is within or beyond their ability. For example, Bertenthal and Campos (1984) found that experience with crawling played a major role in infants' avoidance of the deep side of the visual cliff. McKenzie and Forbes (1992) also found that 12-year-old boys given experience with climbing real steps overestimated the height of the steps they could climb less than did their counterparts given no practice with climbing steps. These findings suggest that experience with acting within the environment is necessary for accurate perception of the fit between one's physical characteristics and the properties of the environment. However, it is not yet known whether there are developmental differences in children's ability to take advantage of information gained through experience. That is, do older children make more accurate perceptual judgments about their abilities if given experience with performing actions than do younger children?

Taken together, these results suggest that individuals may perceive the boundary between actions that are within and beyond their ability as ambiguous. Moreover, when faced with ambiguity, it appears that adults and children are more likely to overestimate than underestimate their abilities. One hypothesis this suggests about childhood accidents is that children are most likely to make errors in judgment when confronted with activities that are just beyond their ability. For example, children may be more likely to run out in front of cars when unsafe gaps between cars are larger and hence more ambiguous than when gaps are smaller and presumably less ambiguous. Thus, when children are confronted with situations that are beyond their ability, accident risk should peak in the range just beyond their ability and decline steadily thereafter.

The aims of the present investigation were threefold. The aim of Experiment 1 was to examine the accuracy of 6- and 8-year-olds' and adults' perceptual judgments about their ability to perform activities that were near their maximum ability and hence ambiguous. The ages of 6 and 8 were chosen because little is known about developmental changes in judgments of physical abilities between infancy and adulthood (see also Rochat, 1993). The aim of Experiment 2 was to examine possible developmental differences in the extent to which experience with performing activities facilitates 6- and 8-year-olds' perceptual judgments about their abilities. The third aim was to examine how individual differences in perceptual judgments relate to accident history.

Both studies involved a laboratory assessment in which participants were asked to judge whether or not they could perform activities of varying levels of difficulty before they attempted to perform each activity. For example, participants stood with their hands at their sides and were asked to look up at an object on a shelf and decide if they could reach it standing on their tip toes. Participants made judgments about their ability to perform four activities at four levels of difficulty. Each activity was scaled at difficulty levels of well within, just within, just beyond, and well beyond individual participants' ability. By comparing participants' judgments of their ability to perform tasks within and beyond their ability, the accuracy of their judgments was determined. Underestimation of ability was inferred when participants judged that they could not perform tasks that they actually were able to perform, and overestimation of ability was inferred when participants judged that they could perform tasks that they actually were unable to perform. Latencies to make judgments also were measured to provide information about the relation between task difficulty and decision time. It was expected that children would overestimate their abilities more than would adults, but that older children would benefit more from experience with performing activities than would younger children. It also was expected that children and adults alike would perceive tasks that were just beyond their ability as most ambiguous and therefore would exhibit the longest decision times for such tasks.

Finally, both experiments examined whether the accuracy of children's judgments in the laboratory setting was related to real-world accidental injuries. The measure of accidents used here was parental reports of the number of accidents their children had experienced requiring medical attention. These types of accidents were selected for study because they have the most serious consequences for children and families. Moreover, because accidents requiring medical attention presumably are salient events for parents, they may be more likely to report them accurately than less serious accidents.

Experiment 1

Method

Participants

Twenty 6-year-olds, 20 8-year-olds, and 20 adults participated. The mean ages of the children were 6 years and 3 months (range = 6 years to 6 years, 10 months) and 8 years, 6 months (range = 8 years, 2 months to 8 years, 11 months). There were equal numbers of male and female participants in each age group. The adults were college students who received course credit for their participation. Children were recruited from an existing participant registry composed of children whose parents had previously expressed interest in child development research. The majority of children were Caucasian and from middle- to upper-middle-class families.

Design and Procedure

Four tasks were used to compare children's perceptions of their ability to perform particular physical activities with their ability to actually perform those activities. The apparatus used for each task is shown in Figure 1. The vertical reach task involved removing a toy from a shelf standing on tip toes. The shelf was mounted on brackets that were attached to shelving strips on a wall. The height of the shelf was adjustable in 1-in. increments. The horizontal reach task involved reaching out from a squatting position for a toy duck on a wooden block without touching hands or knees on the floor. Participants performed the reach task while squatting on a flat board with an edge on the front end to keep them from moving their feet closer to the duck. The stepping task involved stepping across two sticks placed parallel to each other. Children began the step task by putting both toes up against the edge of one of the sticks. The clearance task involved sidling under a wooden bar attached to two posts without knocking the bar off or putting hands or knees on the floor. Participants began the clearance task by squatting parallel to the bar and placing both feet in a box marked by tape on the floor. This set of heterogeneous tasks was chosen to increase the generalizability of the findings and because pilot testing showed that they produced relatively reliable baseline performance from children.

There were four variations of each task used for the test trials: (a) the well within version was 13% below the participant's maximum level of

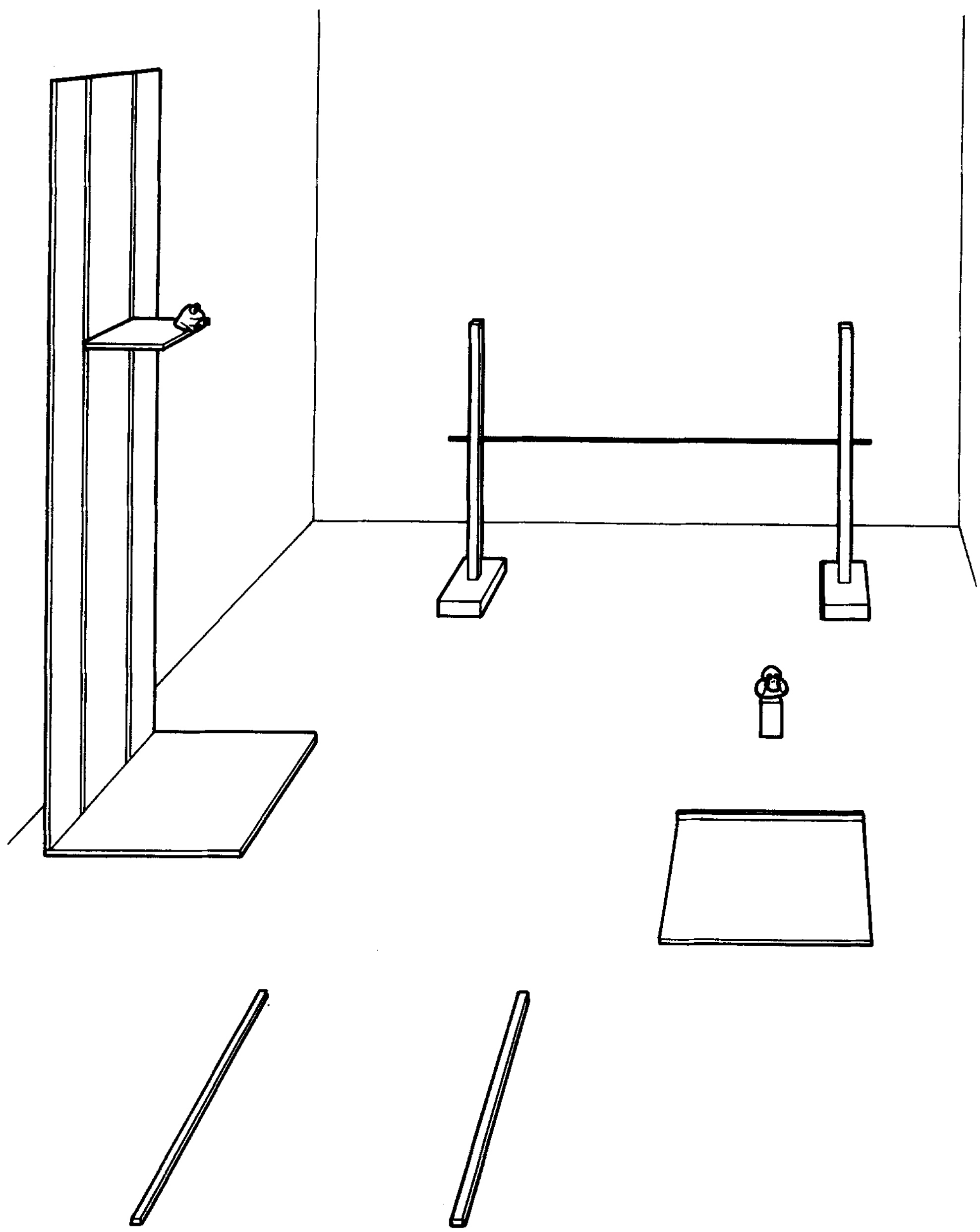

Figure 1. Schematic representation of vertical reach, horizontal reach, stepping, and clearance tasks used in Experiments 1 and 2.

ability; (b) the just within version was at the participant's maximum level of ability; (c) the just beyond version was 8% above the participant's maximum level; and (d) the well beyond version was 13% above the participant's maximum level. In all tasks, levels of difficulty were scaled to the abilities of the individual participants. Baseline measures for participants' maximum levels of ability were obtained before the test trials by having them perform reaching, crouching, and stepping movements in an area outside of the testing room.

The baseline for the vertical reach task was obtained by having children stand facing a wall and reach as high as they could on their tip toes. The height of their reach was marked at the second joint of the index finger. The baseline for the horizontal reach task was obtained by having participants squat down behind a piece of tape marked on the floor and reach forward as far as they could. The experimenter then placed his or her pencil just out of reach and asked the participants to try to touch the pencil. The extent of the reach was marked at the last place at which the participants could touch the pencil. This procedure was adopted because extensive pilot testing revealed that individuals often reached much farther when a goal was present than when it was absent. The baseline for the stepping task was obtained by having participants stand with both toes at a line marked by tape on the floor and take as big a step as they could. A line was marked where the heel of the trailing foot landed. The participants then were asked to go back to the line and try to step farther. This procedure was repeated three times or until the experimenter judged that the participants had stepped as far as they could. Again, this procedure was adopted because pilot testing revealed that participants often stepped farther when a goal was present than when absent. The clearance baseline was obtained by having the participants squat parallel to a wall and duck their heads down as low as possible. The height of the crouch was marked 1 in. higher than the highest point on the participant's back. This measurement was adopted because pilot testing revealed that participants required an extra inch of height to move in a squatting position underneath a bar. All baseline measurements were rounded to the nearest half inch.

Participants received 16 test trials grouped in blocks of four. Within each block, participants received one of each type of task (vertical reach, horizontal reach, stepping, and clearance) in a random order. The four levels of difficulty (well within, just within, just beyond, and well beyond) were randomized across all 16 trials with the constraint that participants received each level of difficulty only once for each type of task. Randomization of difficulty levels ensured that task difficulty was not confounded with trial order.

At the beginning of the first block of test trials, the experimenter explained that they would be playing a game involving the four activities in the room. The purpose of the game scenario was to make the potential payoffs and risks of trying the activities analogous to real-life decision making. The experimenter explained to the participants that they would be asked to decide whether or not they could perform each one of the four activities. The participants were told that if they thought that they could perform a given activity, they would be asked to try it. If they succeeded, they would be awarded a point, but if they failed, a point would be taken away. If they decided they could not perform a given activity, they would not be asked to try it nor would they lose or be given a point. However, at the end of the four blocks of trials, the tasks for which the participants judged they were unable to perform were set up again and they were asked to try to perform each one. This manipulation maintained the integrity of the game during the test trials but also made it possible to compare participants' judgments with their actual ability to perform the tasks. Children also were informed that at the end of the game, the experimenter would add up their points and award prizes. Points were marked on a chalkboard in the testing room.

After explaining the game and the point system, the experimenter explained how each activity was performed. Before making each judgment, the participants were instructed to put themselves in the starting position for each activity. To make their judgment about the vertical reach, participants first stepped up to a line on the floor positioned directly below the shelf. While keeping their hands at their sides and their feet flat on the floor, they were asked whether or not they thought they could grab the frog off the shelf standing on their tip toes. For the horizontal reach, participants were asked to place their toes up to the front edge of the board and squat down. Without reaching their hands forward, the participants were asked whether they could pick the duck up off the block without knocking it over or touching their hands or knees on the floor. For the stepping task, participants were asked to place their toes up against the edge of the near stick and decide whether they could step all the way across the other stick without touching it or putting their hands or feet on the floor between the two sticks. For the clearance task, participants were asked to stand parallel with the bar and put their feet in the box marked with tape on the floor. They then were asked to squat down and look at the bar and decide whether they could get under the bar without knocking it off or touching their hands or knees on the floor. At the end of each block of four trials, participants were taken out of the room while the second experimenter adjusted each apparatus for the next block of trials. This ensured that the participants were not given information about whether each task was adjusted to be easier or harder. Participants were videotaped with a Panasonic camcorder through a one-way mirror. The camcorder was adjusted so that the apparatus for each task was visible from one vantage point.

While children were playing the game, parents filled out an Accidental Injury Questionnaire requesting the ages and types of treatment for any accidental injuries their child had received that required medical attention. The injury categories on the questionnaire included: animal bites, broken bones, bruises, burns, concussions, cuts/scrapes, dental injuries, heat exhaustion/dehydration, choking, near drowning, and joint injuries. There was also a blank category in which parents could specify an injury type not listed on the questionnaire.

Measures

Participants received four accuracy scores based on the percentage of tasks at each level of difficulty they judged correctly. These scores were calculated by dividing the number of correct judgments by the number of correctly scaled tasks at each level of difficulty. In the case of well within and just within level tasks, the correct response was to answer yes to the experimenter's query, and for the just-beyond and well-beyond level tasks, the correct response was to answer no. A task was deemed correctly scaled if the participant was able to perform a task that was at the well-within or just-within levels of difficulty or was unable to perform a task that was at the just-beyond or well-beyond levels of difficulty.

Participants also received four latency scores based on the average amount of time taken to make a judgment for each level of task difficulty. Again, incorrectly scaled tasks were excluded from these scores. Latencies represented the time interval from the point at which the experimenter finished asking participants to make a judgment to the point at which participants answered in the affirmative or negative. Two coders recorded the latencies from videotapes on a Macintosh IIfx computer for all participants. Pearson correlations were used to calculate reliabilities on five randomly selected participants for the four levels of task difficulty. Reliabilities for the well-within, just-within, just-beyond, and well-beyond levels of difficulty were .997, .997, .994, and .999, respectively.

Results

The primary issue addressed in the analyses was whether the three age groups differed in the accuracy of their judgments about the four levels of task difficulty. Also of interest was whether participants took longer to make judgments about levels of task difficulty that were more ambiguous. Preliminary analyses conducted to evaluate the potential moderating effect of gender revealed no significant results. Therefore, the two genders were pooled in all subsequent analyses. Tukey's honestly significant difference (HSD) test was used in all follow-up analyses.

Scaling of Tasks

An initial set of analyses was conducted to evaluate the scaling of tasks. One issue concerns how reliable baseline measures of maximum ability were for scaling the tasks at each age. An Age (6 years vs. 8 years vs. adult) × Difficulty Level (well within vs. just within vs. just beyond vs. well beyond) repeated measures analysis of variance (ANOVA) with the first factor as a between-subjects variable and the second as a within-subjects variable was conducted on the mean number of correctly scaled tasks at each level of difficulty. This analysis yielded a significant effect of difficulty level, $F(3, 171) = 33.94$, $p < .001$, but no effect of age and no interaction between age and difficulty level. The mean number of correctly scaled tasks included for the well-within, just-within, just-beyond, and well-beyond levels of task difficulty was 3.87, 2.85, 3.42, and 3.68, respectively. Follow-up tests revealed that significantly more tasks were included in the well-within, just-beyond, and well-beyond levels than in the just-within level. In addition, more tasks were included in the well-within and well-beyond levels than in the just-beyond level. There was no difference between the well-within and the well-beyond levels.

A second issue concerned whether proportional rather than absolute scaling of the four levels of task difficulty was appropriate. Because the four levels of difficulty were body-scaled, the absolute distances from maximum ability were greater for the larger than for the smaller participants. For example, if a smaller person's maximum reach was 23 in. and a larger person's maximum reach was 34 in., then the just-beyond level of difficulty would be 25 in. and 37 in. for the smaller and larger person, respectively. Although absolute distance from maximum ability differs for the two people in this example, the body-scaled distance remains the same because the object is about half a hand length away for both people. Thus, proportionally scaling difficulty levels to a participant's size would seem to be more appropriate for keeping levels of task difficulty constant across participants. Nonetheless, it was important to test this hypothesis statistically by inspecting correlations between judgment accuracy and participant size. Separate correlations for each age group were computed between participant size and the mean percentage of correct judgments about tasks just beyond and well beyond participants' ability. (Only judgments about tasks beyond participants' ability were used in these analyses because as is shown below, all three age groups almost always gave correct judgments about tasks within their ability.) The measure of a participant's size was the average of the baseline measurements for the four tasks. The correlations between size and percentage of correct judgments about tasks just beyond and well beyond participants' ability were −.28 and .24 for 6-year-olds, .12 and .03 for 8-year-olds, and −.21 and .16 for adults. All were nonsignificant, suggesting that proportional scaling of task difficulty was appropriate.

Accuracy of Judgments

To evaluate whether the three age groups' judgments about the four levels of task difficulty differed, accuracy scores for each level of task difficulty were entered into an Age (6 years vs. 8 years vs. adult) × Difficulty Level (well within vs. just within vs. just beyond vs. well beyond) repeated measures ANOVA with the first factor as a between-subjects variable and the second as a within-subjects variable. This analysis yielded significant effects of age, $F(2, 57) = 10.80$, $p < .001$, and difficulty level, $F(3, 171) = 76.92$, $p < .001$.

Both of these effects, however, were subsumed under an Age × Difficulty Level interaction, $F(6, 171) = 3.82$, $p < .01$ (see Figure 2). Follow-up tests revealed that the pattern of responding to the four levels of task difficulty differed among the three age groups. Both the 6- and the 8-year-olds were significantly more accurate in judging their ability to perform the well-within and just-within levels than the just-beyond and well-beyond levels. Thus, children were nearly always accurate about tasks that were within their ability but overestimated their ability to perform tasks that were beyond their ability about half of the time. In contrast, adults were significantly more accurate in their judgments about the well-within, just-within, and well-beyond levels than about the just-beyond level. Moreover, adults' judgments about the well-within, just-within, and well-beyond levels of task difficulty did not differ. Thus, adults were adept at judging their ability to perform tasks that were within and well beyond their level of ability but sometimes overestimated their ability to perform tasks that were just beyond their ability.

Decision Times

A second question of interest was whether the length of time participants took to make decisions about their ability to per-

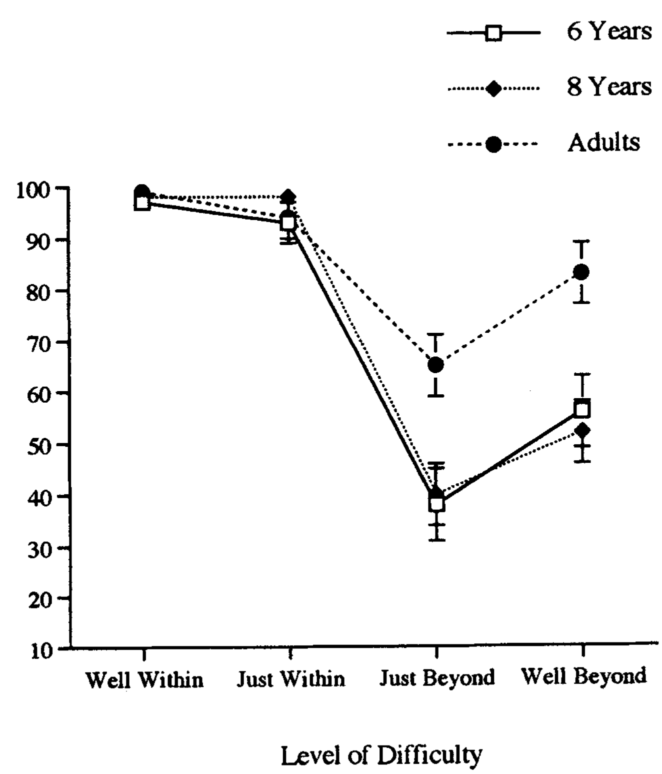

Figure 2. Mean percentage of correct judgments as a function of age and difficulty level in Experiment 1.

form the tasks varied as a function of task difficulty. That is, did decision times increase as the perceived ambiguity of judgments increased? To address this question, mean decision times for each level of task difficulty were entered into an Age × Difficulty Level repeated measures ANOVA with the first factor as a between-subjects factor and the second as a within-subject factor. Before the statistical analysis, decision times that were three or more standard deviations greater than the mean for each level of task difficulty within each age group were classified as outliers and removed. The number of outliers removed for 6-year-olds, 8-year-olds, and adults was 6, 9, and 5, respectively. The analyses were carried out on the remaining decision times regardless of whether participants judged their ability to perform a particular task correctly or incorrectly. This analysis yielded a significant effect of difficulty level, $F(3, 171) = 14.47, p < .001$. Follow-up tests revealed that mean decision times for the just-within, just-beyond, and well-beyond levels of difficulty were significantly longer than that for the well-within level (see Figure 3). Furthermore, the mean decision time for the just-beyond level was significantly longer than that for the just-within level. There were no other significant differences. Thus, participants made the fastest decisions about tasks that were easy for them to perform, followed by tasks that were just within their ability, followed finally by tasks that were just beyond and well beyond their ability.

Relations Between Judgment Accuracy and Accident History

The findings reported above provide insight into the development of perceptual–cognitive skills that may contribute to childhood accidents. These results do not, however, reveal whether these skills are related to individual differences in children's actual accident histories. One way to address this issue is to examine whether individual differences in accuracy of judgments are related to accident proneness. In the present investigation, the number of accidents requiring medical attention that children had experienced was used as an indicator of accident proneness.

Number of accidental injuries. Accidental injury scores were computed by summing the number of accidents requiring medical attention that parents reported on the Accidental Injury Questionnaire. The number of accidents children had experienced ranged between 0 and 4 ($M = 1.33$). Preliminary analyses revealed that the number of accidents did not differ as a function of age, $F(1, 36) = 0.02$, *ns*, or gender, $F(1, 36) = 0.83$, *ns*.

Accuracy of judgment and accident proneness. A composite score representing how accurately children judged tasks that were beyond their level of ability was calculated by averaging scores on the just-beyond and well-beyond levels of task difficulty. Correlational analyses revealed that the relation between judgment accuracy and accident proneness differed for the two age groups. Specifically, there was a significant negative correlation between the accuracy of 6-year-olds' judgments of their physical abilities and the number of accidents they had experienced ($r = -.44, p = .05$). Thus, 6-year-olds who were less accurate in judging their physical abilities had experienced more accidents. In contrast, there was no relationship between 8-year-olds' judgments of their physical abilities and the number of accidents they had experienced ($r = .10$, *ns*).

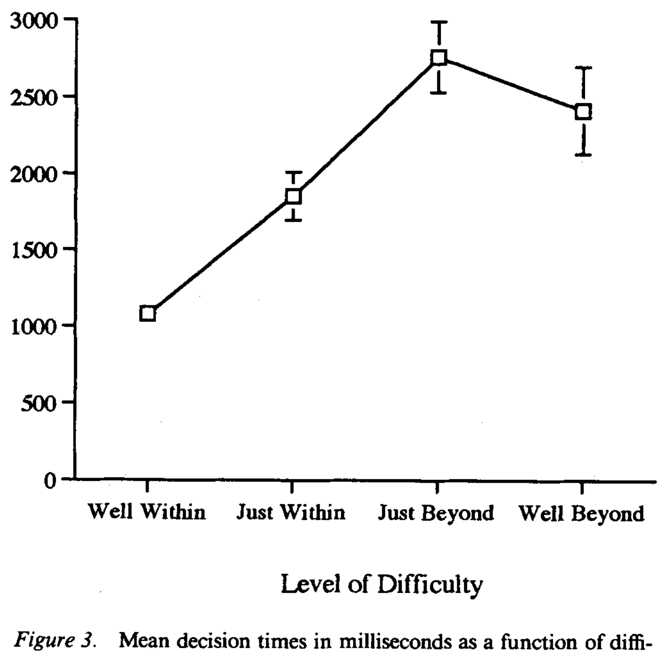

Figure 3. Mean decision times in milliseconds as a function of difficulty level in Experiment 1.

Discussion

The results of this experiment demonstrate that 6- and 8-year-olds often overestimated their ability to perform physical tasks that were beyond their ability. This was true of tasks that not only were just beyond children's ability but also of those that were well beyond their ability. Adults also had difficulty making judgments about tasks that were just beyond their ability but had less difficulty than children in making judgments about tasks that were well beyond their ability. Adults and children did not differ, however, in the length of time they took to make judgments about their ability to perform tasks at the four levels of difficulty. Participants' decision times suggest that they became increasingly unsure of their judgments as tasks increased in difficulty from well within their ability to just beyond their ability.

The inaccuracy of children's assessments of their physical skills raises the issue of what factors influence their judgments. One factor that may play an important role is children's sensitivity to feedback about their physical skills. Presumably, one way children learn about the relation between their abilities and the demands of the task is by attempting to perform various physical activities. Success and failure, therefore, provide children with useful sources of information about their abilities. The dilemma of success and failure, however, is that although it informs children about whether or not they are capable of doing a particular activity, it does not tell them how much more or less they are capable of doing. Appropriate decision-making, therefore, always involves weighing previous successes and failures against new challenges that present themselves. Failure to take into account both past experience and the present situation when making a decision is likely to result in errors of judgment, and this in turn may lead to an accident. Younger children and accident-prone children may have a particularly difficult time integrating both sources of information and hence may make more errors in judging their physical abilities.

Experiment 2 investigated the question of whether there are developmental differences in children's ability to use prior experience to inform their judgments about their physical abilities. This issue was addressed by manipulating 6- and 8-year-old's initial experiences of success and failure in the four tasks from the laboratory assessment used in Experiment 1. Children received four practice trials before the test trials that were either possible or impossible to perform. All children then received 16 additional tasks of the four levels of difficulty. Again, parents filled out the Accidental Injury Questionnaire.

Experiment 2

Method

Participants

Twenty-four 6-year-olds and 24 8-year-olds participated. The mean ages of the children were 6 years and 4 months (range = 6 years 1 month to 6 years 11 months) and 8 years and 4 months (range = 8 years 1 month to 8 years 11 months). There were equal numbers of male and female children in each age group and condition. The children were recruited in the same manner as Experiment 1.

Design and Procedure

The same procedure was used as in Experiment 1 except that children received four practice trials before the test trials. The practice trials were used to manipulate children's experiences of initial success or failure. Equal numbers of children at each age were randomly assigned to either the success or the failure condition. In the *success* condition, children received possible versions of each task as practice trials. These tasks were adjusted at 10% below the child's maximum level of ability. In the *failure* condition, children received impossible versions of each task as practice trials. These tasks were adjusted at 10% above the child's maximum level of ability. Children in both conditions then received four blocks of test trials. Within each block, children received one of each type of task (vertical reach, horizontal reach, stepping, and clearance) in a random order. The four levels of difficulty for the test trials were scaled and randomized as in Experiment 1.

At the beginning of the block of practice trials, the experimenter explained to the children that they would be playing a game involving the four activities in the room. The experimenter then explained how to do each activity. Children were informed that they would get a chance to practice each activity before playing the game. For the practice trials, children were asked to try to perform each activity without making any judgments about their ability to perform each activity. After the practice trials, the experimenter took the children out of the testing room while the second experimenter set up the first block of test trials. When the experimenter brought the children back into the room for the first block of test trials, he or she explained the rules of the game. During the test trials, children made judgments about their ability to perform the tasks as in Experiment 1. Children again were videotaped with a Panasonic camcorder through a one-way mirror. While children were playing the game, parents filled out the Accidental Injury Questionnaire requesting the ages and types of treatment for any accidental injuries their child had received that required medical attention.

Measures

Accuracy scores, decision time scores, and accidental injury scores were calculated as in Experiment 1. Decision time reliabilities for the well-within, just-within, just-beyond, and well-beyond levels of difficulty were .934, .999, .999, and .996, respectively.

Results

The primary issue addressed in the analyses was whether the prior experiences of success or failure influenced the accuracy of children's judgments. The amount of time children took to make their judgments again was analyzed to provide information about whether the more ambiguous difficulty levels resulted in longer decision times. Preliminary analyses evaluating the potential moderating effect of gender revealed no significant results. Therefore, the two genders were pooled in all subsequent analyses. Tukey's HSD test again was used in all follow-up analyses.

Scaling of Tasks

An initial set of analyses was conducted to evaluate the scaling of tasks. An Age (6 years vs. 8 years) × Difficulty Level (well within vs. just within vs. just beyond vs. well beyond) repeated measures ANOVA with the first factor as a between-subjects variable and the second as a within-subjects variable was conducted on the mean number of correctly scaled tasks at each level of difficulty. This analysis yielded a significant effect of difficulty level, $F(3, 171) = 14.87, p < .001$, but no effect of age and no interaction between age and difficulty level. The mean number of correctly scaled tasks included for the well-within, just-within, just-beyond, and well-beyond levels of task difficulty were 3.88, 3.27, 3.31, and 3.77, respectively. Follow-up tests revealed that significantly more tasks were included in the well-within and well-beyond levels than in the just-within and just-beyond levels. There was no difference between the well-within and the well-beyond levels nor between the just-within and just-beyond levels.

A second issue concerned whether larger children within each age group had an advantage over smaller children because the absolute distances from maximum ability were greater for the larger than for the smaller children. As in Experiment 1, separate correlations for each age group were computed between child's size and the mean percentage of correct judgments about tasks just beyond and well beyond the childrens' ability. The respective correlations between size and percentage of correct judgments about tasks just beyond and well beyond childrens' ability were −.05 and .31 for 6-year-olds and .05 and .19 for 8-year-olds. All were nonsignificant, suggesting that proportional scaling of task difficulty was appropriate.

Accuracy of Judgments

The major question of interest was whether 6- and 8-year-olds in the success and failure conditions responded differently to the four levels of task difficulty. This question was addressed by entering children's accuracy scores for the four levels of difficulty into an Age (6 years vs. 8 years) × Practice Condition (success vs. failure) × Difficulty Level (well within vs. just within vs. just beyond vs. well beyond) repeated measures ANOVA with the first two factors as between-subjects variables and the third as a within-subjects variable. This analysis yielded significant effects of age, $F(1, 44) = 16.66, p < .001$, and difficulty level, $F(3, 132) = 107.30, p < .001$.

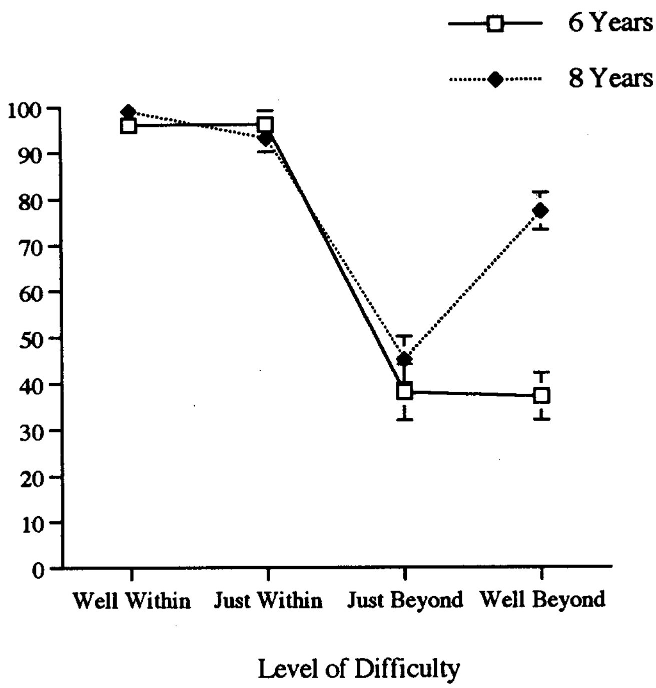

Figure 4. Mean percentage of correct judgments as a function of age and difficulty level in Experiment 2.

These effects were subsumed, however, by a significant Age × Difficulty Level interaction, $F(3, 132) = 12.99, p < .001$ (see Figure 4). Follow-up tests revealed that the pattern of responding to the four levels of task difficulty differed between the two age groups. As in Experiment 1, 6-year-olds were significantly more accurate in judging their ability to perform the well-within and just-within levels than the just-beyond and well-beyond levels (see Figure 4). Thus, 6-year-olds were nearly always accurate about tasks that were within their ability but overestimated their ability to perform tasks that were beyond their ability about half of the time. Eight-year-olds also were more accurate in judging their ability to perform the well-within and just-within levels than the just-beyond and well-beyond levels. Unlike in Experiment 1, however, 8-year-olds' judgments about tasks at the well-beyond level were significantly more accurate than their judgments about tasks at the just-beyond level. Thus, 8-year-olds were just as likely not to respond that they could perform tasks that were just beyond their ability but often responded that they were unable to perform tasks that were well beyond their ability. These findings suggest that 8-year-olds, but not 6-year-olds, benefited from prior experience with performing the tasks.

Decision Times

A second question of interest was whether children took more time to make judgments about difficulty levels that were more ambiguous. To address this question, mean decision times for each level of task difficulty were entered into an Age × Practice condition × Difficulty Level repeated measures ANOVA with the first two factors as between-subjects variables and the third as a within-subjects variable. Before the statistical analysis, decision times that were three or more standard deviations greater than the mean for each level of task difficulty within each age group were classified as outliers and removed. The number of outliers removed for 6- and 8-year-olds was 10 and 5, respectively. The analyses were carried out on the remaining decision times regardless of whether subjects judged their ability to perform a particular task correctly or incorrectly. This analysis yielded a significant effect of difficulty level, $F(3, 132) = 15.20$, $p < .001$, and a significant interaction between age and difficulty level, $F(3, 132) = 3.40, p < .05$. As shown in Figure 5, 6-year-olds made significantly faster decisions about tasks that were just within and well within their ability than about tasks just beyond and well beyond their ability. Eight-year-olds made significantly faster decisions about tasks that were well within their ability than about tasks just within, just beyond, and well beyond their ability. Thus, children again made the fastest decisions about tasks that were well within their ability to perform and the slowest decisions about tasks that were just beyond and well beyond their ability. Compared with 6-year-olds, however, 8-year-olds exercised more caution in their decisions about tasks that were just within their ability.

Relations Between Judgment Accuracy and Accident History

Number of accidental injuries. The number of accidents children had experienced ranged between 0 and 6 ($M = 1.23$). Preliminary analyses revealed that the number of accidents did not differ as a function of age, $F(1, 44) = 0.10$, *ns*, or gender, $F(1, 44) = 0.28$, *ns*.

Accuracy of judgment and accident proneness. A composite score representing how accurately children judged tasks that were beyond their level of ability was calculated as in Experiment 1. Correlational analyses again revealed that the relation between judgment accuracy and accident proneness differed for

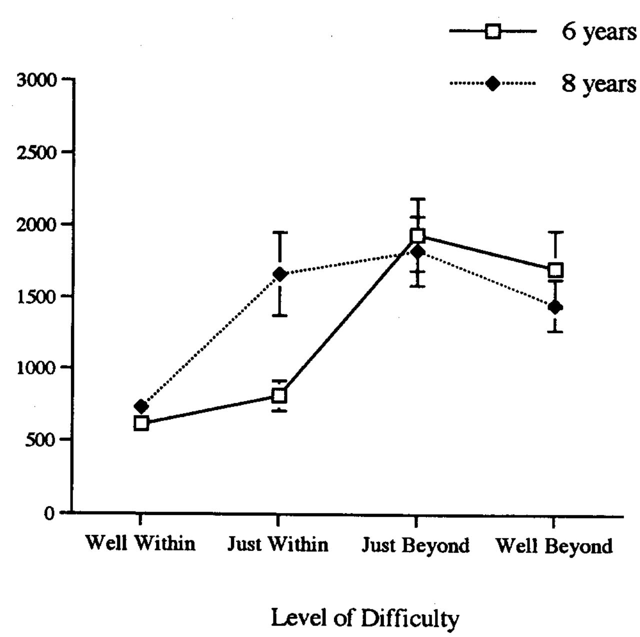

Figure 5. Mean decision times in milliseconds as a function of age and difficulty level in Experiment 2.

the two age groups. There was a significant negative correlation between the accuracy of 6-year-olds' judgments of their physical abilities and the number of accidents they had experienced ($r = -.48$, $p < .05$) but no relationship between 8-year-olds' judgments of their physical abilities and the number of accidents they had experienced ($r = -.02$, *ns*).

Discussion

The results of this experiment again demonstrate that 6-year-olds were less accurate about tasks that were just beyond and well beyond their ability than about tasks that were within their ability. In contrast to the previous experiment, however, 8-year-olds' judgments were more accurate for tasks that were well beyond their ability than for those that were just beyond their ability. This result was obtained for 8-year-olds who experienced success on the practice trials and for those who experienced failure on the practice trials. This suggests that experiences with success and failure may operate similarly to inform individuals about their physical skills. Alternatively, the nonthreatening consequences of failing in the present investigation may have lessened the likelihood that more cautious judgments would result from failing than from succeeding. In either case, the results of the two experiments suggest that 8-year-olds are more likely than 6-year-olds to incorporate prior experience into their subsequent decision making. Without additional experience, 8-year-olds are just as likely as 6-year-olds to make errors about their ability to perform physical tasks that are well beyond their ability.

General Discussion

The results of these studies show that children consistently overestimated their physical abilities and that individual differences in younger children's overestimation was related to accident proneness. In both experiments, 6-year-olds overestimated their ability to perform even tasks that were well beyond their ability. Moreover, 6-year-olds who were less accurate in judging their ability to perform tasks that were beyond their ability had also experienced a greater number of serious accidents. Older children also overestimated their physical abilities, but the tendency to do so for tasks well beyond their ability was attenuated by more experience with performing the activities. Interestingly, adults also made errors in judging their physical abilities but only when tasks were just beyond their ability and apparently highly ambiguous. When tasks were well beyond their ability, adults were much less likely to make errors in judgment.

Although some have suggested that adults' perception of whether activities are within or beyond their ability is categorical (e.g., Warren, 1984), the results of these experiments suggest that this perception may have a graded structure. In other words, individuals seem to perceive the boundary between activities that are within and beyond their ability as ambiguous. Why might this be the case? One possible reason is that motor performance is variable (Carron, 1970; Carron & Bailey, 1969). That is, how far one can reach, step, or jump varies somewhat from one attempt to the next. For example, in the present investigation, variability in performance resulted in fewer trials being included in the just-within and just-beyond levels of difficulty than in the well-within and well-beyond levels. One consequence of even small variations in motor performance is that individuals may be reasonably good at making rough approximations about their ability but have difficulty making precise judgments about what they can and cannot do. Further research is needed to examine whether the degree of variability in motor performance maps onto the degree of ambiguity individuals perceive concerning the boundary between actions that are within and beyond their ability.

The finding that adults overestimated less than did children and that 8-year-olds benefited more from experience than did 6-year-olds suggests that there are developmental changes beyond infancy in the ability to perceive the relation between one's own physical abilities and the demands of the situation. These findings directly parallel findings from the metacognitive literature showing that 5- and 6-year-olds often overestimate their memory abilities (Flavell, Friedrichs, & Hoyt, 1970) and that young children often continue to have unrealistically high expectations about their memory abilities even after experiencing failure (Yussen & Levy, 1975). Thus, it appears that there may be domain-general developmental change in children's ability to evaluate their level of skill in relation to the demands of the task.

What might account for developmental changes in children's judgments about their physical abilities? First, motor performance becomes less variable with age (e.g., Parker, Larkin, & Ackland, 1993; Smoll & DenOtter, 1976). For example, Smoll and DenOtter (1976) found that intraindividual variability in rolling balls toward a target steadily declined between the ages of 5 and 11. Likewise, Parker et al. (1993) found that children's hopping performance increased and within-trial variability decreased with age. One consequence of this reduced variability in motor performance is that older children and adults may have more consistent information available to them about their capabilities and limitations than may younger children. This also suggests that the range of ambiguity may be larger for younger than for older children and adults. In the present investigation, errors dropped significantly for 8-year-olds in Experiment 2 and for adults in Experiment 1 when tasks were scaled at 13% beyond their ability. In contrast, 6-year-olds continued to overestimate at these distances. Although errors at greater distances were not examined in the present investigation, one would expect that 6-year-olds' errors would drop at some point beyond the 13% mark.

Second, as children grow older, they undoubtedly accumulate more experience with succeeding and failing on physical tasks, which may facilitate their ability to make judgments about the relation between their physical skills and the demands of the task. This explanation is consistent with findings showing that infants with more crawling experience show more avoidance of the visual cliff than do infants with less crawling experience (Bertenthal & Campos, 1984) and that 12-year-olds given experience with climbing steps make more accurate judgments about whether stairs are climbable than do 12-year-olds without such experience (McKenzie & Forbes, 1992). The results of the present investigation also suggest that there are developmental differences in children's ability to take advantage of experience. Quite likely, younger children require greater amounts of experience to make accurate judgments about their abilities than do older children.

The present investigation also revealed that individual differences in 6-year-olds' ability to make judgments about their physical abilities was related to the number of accidents they had experienced requiring medical attention. Thus, 6-year-olds who made more errors in judging their abilities had also experienced a greater number of serious accidents than had those who made fewer errors. The fact that this finding was replicated across both experiments is noteworthy and suggests that the relation is stable. Although other factors may mediate the relation between overestimation of physical abilities and accident proneness, these findings suggest that deficits in children's ability to make judgments about the relation between their physical abilities and the demands of the situation may put them at risk for accidents during the early years of childhood.

A cautionary note also is warranted, however, because the correlation for 6-year-olds was moderate, and others have not found a correlation between falls at home and avoidance of steep slopes or the visual cliff (e.g., Adolph, 1995; Bertenthal, Campos, & Barrett, 1984). One reason for this discrepancy, however, may be the ages of the children involved. Because the range of serious falls during infancy is likely to be quite restricted, the relation between falls and behavior in the lab may be difficult to detect. Errors in parental reporting of accidents may also contribute to discrepancies among studies. When asked only to report "serious" falls (e.g., Bertenthal et al., 1984; Walk, 1966), for example, parents may differ widely in what they classify as a serious fall or in their ability to remember their children's mishaps. Parental reports of accidents requiring medical attention such as those gathered in the present investigation may be less likely to be influenced by such problems. Moreover, accidents requiring medical attention may better differentiate children who are accident prone from those who are not.

Both studies in the present investigation revealed a correlation between overestimation of ability and number of accidents for 6-year-olds but not for 8-year-olds. Why might this be the case? One explanation for this discrepancy is that the factors that contribute to errors in the laboratory differ for the two age groups. More specifically, although the factors that cause 6-year-olds to make more errors in judging their physical abilities in the laboratory may be similar to the ones that resulted in accidents in early childhood, the same may not be true for 8-year-olds. Further research, however, is necessary before any definitive conclusions can be drawn about developmental differences concerning the correlation between overestimation and accidents.

A final issue raised by this investigation and others like it is why children and adults exhibit a bias to overestimate their abilities. One factor that may play a role is the aversiveness of the consequences that result from making errors. Although children and adults in the present investigation lost points when they made errors, there was no bodily penalty for overestimating one's ability. Therefore, even adults may have decided it was better to risk failing than to avoid trying. When errors have more aversive consequences, however, individuals may exercise more caution. For example, Adolph (1995) recently found that infants are less likely to attempt going down slopes than up slopes presumably because falling is more aversive when going down than up a hill.

A second factor that may play a role in overestimation of ability is the attractiveness of the goal. In the present investigation, children wanted to win points and try the activities. The only way to achieve these goals was by making affirmative judgments about their ability. As a result, children's desire to achieve these goals may have overridden doubts they had about their ability to perform activities that were beyond their ability. Likewise, infants' desire to grasp an attractive object or reach their mother may explain why they often overestimate their abilities. In studies of infants' judgments about the traversability of slopes and surfaces, for example, mothers stand at the end of the runway and coax infants with toys or food (Adolph, 1993, 1995; Gibson et al., 1987). In ambiguous situations, infants may rely heavily on the mother's behavior to make decisions about whether to attempt an activity. In fact, Sorce, Emde, Campos, and Klinnert (1985) found that when the visual cliff was adjusted to an ambiguous height, infants crossed the cliff when their mother posed expressions of joy or interest but not when she posed expressions of fear or anger.

The extent to which judgments err in a conservative or nonconservative direction therefore, may be jointly constrained by perceptual and nonperceptual factors. Furthermore, developmental changes in nonperceptual abilities may play an important role in determining the extent to which children of different ages overestimate their abilities. For example, research on achievement motivation has shown that young children have difficulty separating their desire to perform an activity from their ability to perform the activity (e.g., Stipek & Mac Iver, 1989). If so, younger children may be highly influenced by the attractiveness of the goal and hence be more likely than older children to overestimate their abilities. Further research investigating the role of these nonperceptual factors may enrich our understanding of how adults and children make judgments about the relation between their own abilities and the demands of the situation.

From a more general standpoint, the finding that children have a tendency to overestimate their physical abilities raises an interesting developmental issue. Namely, by its very nature, development involves aspiring to do things that are beyond one's current level of ability. Without such motivation, it seems unlikely that development would move forward. As others have pointed out (e.g., Bjorklund & Green, 1992), there may be a beneficial component to a general bias to overestimate one's ability that offsets some of the potential hazards in doing so. The developmental dilemma, therefore, is to continually aspire to trying new and difficult things but not to try things that might have disastrous consequences. Children who have difficulty making this distinction are likely to be at risk for serious accidental injuries. The present experiments represent a step forward in understanding the developmental changes in judgments about physical abilities during middle childhood and in unravelling possible factors contributing to childhood accidents.

References

Adolph, K. E. (1995). Psychophysical assessment of toddlers' ability to cope with slopes. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 734–750.

Adolph, K. E., Eppler, M. A., & Gibson, E. J. (1993). Crawling versus walking infants' perception of affordances for locomotion over sloping surfaces. *Child Development, 64,* 1158–1174.

Bertenthal, B. I., & Campos, J. J. (1984). A reexamination of fear and its determinants on the visual cliff. *Psychophysics, 21,* 413–417.

Bertenthal, B. I., Campos, J. J., & Barrett, K. C. (1984). Self-produced locomotion: An organizer of emotional, cognitive, and social development in infancy. In R. N. Emde & R. J. Harmon (Eds.), *Continuities and discontinuities in development.* Plenum Press.

Bjorklund, D., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist, 47,* 46–54.

Brooks, P. H., & Roberts, M. C. (1990). Social science and the prevention of children's injuries. *Social Policy Report of the Society for Research in Child Development, 4,* 1–11.

Carello, C., Grosofsky, A., Reichel, F. D., Solomon, H. Y., & Turvey, M. T. (1989). Visually perceiving what is reachable. *Ecological Psychology, 1,* 27–54.

Carron, A. V. (1970). Intra-task reliability and specificity of individual consistency. *Perceptual and Motor Skills, 30,* 583–587.

Carron, A. V., & Bailey, D. A. (1969). Evidence for reliable individual differences in intra-individual variability. *Perceptual and Motor Skills, 28,* 843–846.

Christoffel, K. K., Schofer, J. L., Jovanis, P. P., Brandt, B., White, B., & Tanz, R. (1986). Childhood pedestrian injury: A pilot study concerning etiology. *Accident Analysis and Prevention, 18,* 25–35.

Coppens, N. M. (1986). Cognitive characteristics as predictors of children's understanding of safety and prevention. *Journal of Pediatric Psychology, 11,* 189–202.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology, 1,* 324–340.

Gibson, E. J., Riccio, G., Schmuckler, M. A., Stoffregen, T. A., Rosenberg, D., & Taormina, J. (1987). Detection of the traversability of surfaces by crawling and walking infants. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 533–544.

Gibson, E. J., & Walk, R. D. (1960). The "visual cliff." *Scientific American, 202,* 64–71.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Langley, J., McGee, R., Silva, P., & Williams, S. (1983). Child behavior and accidents. *Journal of Pediatric Psychology, 8,* 181–189.

Lee, D., Young, D., & McLaughlin, C. (1984). A roadside simulation of road crossing for children. *Ergonomics, 27,* 1271–1281.

McKenzie, B. E., & Forbes, C. (1992). Does vision guide stair climbing? A developmental study. *Australian Journal of Psychology, 44,* 177–183.

McKenzie, B. E., Skouteris, H., Day, R. H., Hartman, B., & Yonas, A. (1993). Effective action by infants to contact objects by reaching and leaning. *Child Development, 64,* 415–429.

Nixon, J., Pearn, J., Wilkey, I., & Corcoran, A. (1986). Fifteen years of child drowning—A 1967–1981 analysis of all fatal cases from the brisbane drowning study and an 11 year study of consecutive near-drowning cases. *Accident Analysis and Prevention, 18,* 199–203.

Parker, H. E., Larkin, D., & Ackland, T. R. (1993). Stability and change in children's skill. *Psychological Research, 55,* 182–189.

Peterson, L., & Mori, L. (1985). Prevention of child injury: An overview of targets, methods, and tactics for psychologists. *Journal of Consulting and Clinical Psychology, 53,* 586–595.

Pick, H. J., Jr., Plumert, J. M., & Arterberry, M. (1987). *ATVs and children: Perceptual-motor, cognitive, and social risk factors.* Report prepared for the Consumer Product Safety Commission (Contract No. CPSC-C-87-1289).

Pufall, P. B., & Dunbar, C. (1992). Perceiving whether or not the world affords stepping onto and over: A developmental study. *Ecological Psychology, 4,* 17–38.

Roberts, M. C. (1986). Health promotion and problem prevention in pediatric psychology: An overview. *Journal of Pediatric Psychology, 11,* 147–161.

Rochat, P. (1993, March). *Perceived reachability for self and for others in children.* Poster presented at the Biennial Meeting of the Society for Research in Child Development, New Orleans, LA.

Rodriguez, J. G., & Brown, S. T. (1990). Childhood injuries in the United States. *American Journal of Disease Control, 144,* 627–646.

Smoll, F. L., & DenOtter, P. (1976). Intraindividual variability in development of accuracy of motor performance. *Journal of Motor Behavior, 8,* 195–201.

Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology, 21,* 195–200.

Stipek, D., & Mac Iver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 60,* 521–538.

Walk, R. D. (1966). The development of depth perception in animals and human infants. *Monographs of the Society for Research in Child Development, 31* (5, Serial No. 107).

Warren, W. H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 683–703.

Yussen, S. R., & Levy, V. (1975). Developmental changes in predicting one's own span of short-term memory. *Journal of Experimental Child Psychology, 19,* 502–508.

Received May 16, 1994
Revision received October 12, 1994
Accepted October 12, 1994 ■