JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY **67,** 317–337 (1997)
ARTICLE NO. CH972411

Social and Temperamental Influences on Children's Overestimation of Their Physical Abilities: Links to Accidental Injuries

Jodie M. Plumert and David C. Schwebel

University of Iowa

This study examined social and temperamental influences on children's judgments about their physical abilities and relations between temperamental characteristics, ability overestimation, and accidental injuries. Six- and 8-year-olds first observed a peer succeed or fail on a set of physical tasks and then made judgments about their ability to perform those same physical tasks. At both ages, children who first watched a peer fail on the tasks made more conservative judgments about their own abilities than did children who watched the peer succeed. The relations between temperamental characteristics and judgment ability differed for the two ages. An aggregated temperament measure of Surgency/Undercontrol was related to judgment accuracy for 6-year-olds and to decision times for 8-year-olds. Likewise, the relations between temperament, ability overestimation, and accidental injuries differed for the two age groups. Ability overestimation was related to accidental injuries for 6-year-old boys whereas temperamental characteristics were related to accidental injuries for 8-year-olds. These findings suggest that both the factors that put children at risk for accidental injuries and the relations between temperamental characteristics and cognitive abilities change with development.

Understanding the causes of childhood accidents plays an integral role in promoting children's health, as accidents are the leading cause of death in children under age 18 (Rodriguez & Brown, 1990; Singh & Yu, 1996). Despite the growing national concern over promoting children's safety, however, the causes of childhood accidents remain poorly understood. In particular, little is known about how cognitive abilities, peer influences, and temperamental

This research was supported by a University of Iowa Spelman Rockefeller Child Research Seed Grant. The authors thank the following individuals for their assistance in data collection and coding: Stephanie Dawes, Dana Dunisch, Jason Fine, Andrea Harmon, Damien Ihrig, Jennifer Kreps, Penney Nichols-Whitehead, Melissa Sandvick, Susan Totten, and Azure Welborn. We also thank Grazyna Kochanska for her helpful comments on earlier versions of the manuscript.

Address reprint requests to Jodie M. Plumert, Department of Psychology, 11 SSH East, University of Iowa, Iowa City, IA 52242.

0022-0965/97 $25.00

characteristics contribute to unsafe behavior. For example, immature cognitive skills such as errors in judging danger or the inability to foresee consequences may put children at risk for accidents. Recent research shows that not only do 6- and 8-year-olds frequently overestimate their physical abilities, but that ability overestimation is significantly correlated with accident proneness in 6-year-olds (Plumert, 1995). Although little is known about the role of peer influence in accidental injuries, children may engage in unsafe activities after observing peers perform those same activities. Research also shows that temperamental characteristics such as impulsivity and sensation-seeking are significantly correlated with accident proneness (Bijur, Golding, Haslum, & Kurzon, 1988; Jaquess & Finney, 1994; Langley, McGee, Silva, & Williams, 1983; Matheny, 1986; Nyman, 1988; Pless, Taylor, & Arsenault, 1995; Potts, Martinez, & Dedmon, 1995; Pulkkinen, 1995). Thus, ability overestimation, peer influences, and temperamental characteristics may all play a role in childhood accidents. At present, however, it is not known how these factors work together to put children at risk for accidents.

One important but poorly understood factor that may contribute to accident risk is how children assess danger. Although parents often warn children about potentially dangerous situations (e.g., playing with matches), safe behavior ultimately depends on the child's own ability to assess and avoid dangerous situations. According to J. J. Gibson (1979), adaptive behavior within the environment depends upon perceiving the fit between one's own physical characteristics and the properties of the environment in which actions take place. When assessing danger, children must evaluate their own level of skill in performing an activity in relation to the demands of the specific situation. 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 walk or run safely through traffic gaps (see also Lee, Young, & McLaughlin, 1984).

Several studies have shown that children tend to overestimate their physical abilities (Adolph, 1995; Adolph, Eppler, & Gibson, 1993; Lee et al., 1984; McKenzie & Forbes, 1992; McKenzie, Skouteris, Day, Hartman, & Yonas, 1993; Plumert, 1995). For example, McKenzie et al. (1993) found that even when objects were well out of reach, infants attempted to grasp them. They noted, in fact, that many infants would have fallen while trying to grasp objects that were well out of reach had they not been restrained in their infant seats. Likewise, Adolph and her colleagues (Adolph, 1995; Adolph et al., 1993) found that toddlers overestimated their ability to ascend and descend slopes. Older children also make errors in judging their physical abilities. McKenzie and Forbes (1992) found that 9- and 12-year-old boys consistently overestimated the height of steps that they could climb. Lee et al. (1984)

found that although 5- to 9-year-olds were generally cautious in a simulated road-crossing task, the majority of children accepted a traffic gap that was too short on at least one trial. Finally, Plumert (1995) found that 6- and 8-year-olds overestimated their ability to perform reaching and stepping tasks that were well beyond their ability, and that 6-year-olds who were less accurate in judging their ability to perform tasks beyond their ability had also experienced a greater number of lifetime accidents requiring medical attention. Together, these findings support the notion that children overestimate their abilities and that overestimation of ability is related to accident risk.

One question these studies on children's judgments about their physical abilities raise is *why* do children overestimate their abilities? One possibility is that children have very liberal response criteria for making decisions about their ability to perform physical tasks. That is, only a relatively small chance of success is necessary in order for children to judge that they are capable of performing an activity. This may be particularly true when the consequences of making an error in judgment are minimal. When the consequences are more severe, children may adjust their response criteria in a more conservative direction. For example, Adolph et al. (1993) found that toddlers were more likely to ascend than descend slopes that were beyond their ability, presumably because the consequences of falling are more aversive when descending than when ascending slopes. Likewise, infants are more likely to traverse the shallow side than the deep side of the visual cliff because the consequences of falling are more severe for large dropoffs than for small dropoffs.

Temperamental characteristics may also influence how children adjust their response criteria. Highly active and impulsive children are quick to respond and often rush into activities without thinking about them (Rothbart, Ahadi, & Hershey, 1994). Such children may have difficulty anticipating the aversive consequences of making an error in judgment and therefore fail to adjust their response criteria in potentially dangerous situations. As a result, these children may be at greater risk for accidents. No studies to date, however, have examined whether children who are highly active and undercontrolled are more likely to overestimate their physical abilities than are children who are less active and more controlled.

Although the relations between temperamental characteristics and judgment skills are poorly understood, several studies have established a link between temperament and accident proneness. For example, a longitudinal study revealed that the best predictor of accidents by 27-year-old males was low emotional and behavioral regulation in childhood (Pulkkinen, 1995). Men who had experienced the most accidents at age 27 were also the ones who were rated highest on noncompliance and aggression at ages 8 and 14 by their teachers and peers. Other prospective and retrospective studies have also found links between personality characteristics and accident proneness. In one of the first major studies of personality and accident proneness, Man-

heimer and Mellinger (1967) found that children with a history of accidental injuries were more extraverted, impulsive, daring, and careless. More recent studies have revealed links between accident proneness and characteristics such as impulsivity, hyperactivity, impatience, inattention, and sensation-seeking (Bijur et al., 1988; Jaquess & Finney, 1994; Langley et al., 1983; Matheny, 1986; Nyman, 1988; Pless et al., 1995; Potts et al., 1995). Together, these findings suggest that temperamental characteristics play a role in childhood accidents.

Another factor that may play a role in children's response criteria for making judgments about their physical abilities is peer influence. That is, children may use the performance of other children to inform them about their own abilities. For example, a child who sees another child successfully ride her bike down a steep bumpy hill may think that she also can successfully ride her bike down the hill. Festinger (1954) suggested that such comparisons are useful for making accurate assessments about what one is capable of doing, particularly under conditions of uncertainty. Thus, when children are confronted with situations in which they are not sure whether they are capable of performing an activity, they may look to the performance of peers to inform them about their own ability.

Recent reformulations of Festinger's (1954) social comparison theory emphasize that social comparison processes operate differently depending on one's level of knowledge or skill in a domain. For relatively unskilled and inexperienced individuals such as children, social comparison serves the function of providing information about the requirements of the task (Ruble & Frey, 1991). One critical piece of information that social comparison can provide to children is how difficult the task is to perform. If children observe another child failing, they may perceive the task as difficult. Conversely, if they observe another child succeeding, they may perceive the task as easy to perform. Social comparisons only lead to accurate assessments, however, when both individuals involved in the comparison are highly similar on critical dimensions (Wheeler, Martin, & Suls, 1997). Errors in judgment are likely to result when children fail to take into account age or skill differences between themselves and the other child. Hence, skill in accurately using someone else's performance to make judgments about one's own ability may be a risk factor for childhood injuries.

The purpose of this study was to explore the social, cognitive, and temperamental underpinnings of accident proneness. More specifically, this study examined: (a) social and temperamental influences on children's judgments about their physical abilities; and (b) the links between ability overestimation, temperamental characteristics, and accidental injuries. As in Plumert (1995), the study involved a laboratory assessment in which subjects were asked to judge whether or not they could perform four physical tasks of varying levels of difficulty before they attempted to perform each task. For example, children 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 tip toes. By comparing children's judgments of their ability to perform the tasks with their actual ability to perform the tasks, the accuracy of their judgments was determined. Underestimation of ability was inferred when subjects judged that they could not perform tasks that they actually were able to perform; overestimation of ability was inferred when subjects 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. As in Plumert (1995), it was expected that 6-year-olds would overestimate their abilities more than would 8-year-olds, and that both age groups would perceive tasks that were just beyond their ability as most ambiguous and therefore would exhibit the longest decision times for such tasks.

We assessed the influence of social comparison information on children's judgments about their physical abilities by having children view a videotape of a same-age and same-sex peer either succeeding or failing on the four tasks. Children watched the videotape just before they played the game so that they could use the other child's performance to assess the difficulty of the tasks. We expected that children who watched the peer fail on all the tasks would be more conservative in their judgments about their own abilities than would children who watched the peer succeed on all the tasks. We also examined the relations between children's temperamental characteristics and both the accuracy of their judgments and the length of time they took to make their judgments. We expected that children whose parents characterized them as highly active, impulsive, and impatient would be more likely to overestimate their physical abilities and would exhibit shorter latencies to make judgments. Finally, we examined whether ability overestimation and temperamental characteristics were related to real-world accidental injuries. The measure of accidents employed was the average severity of the day-to-day injuries children experienced over a two-week period. Based on the results of Plumert (1995) we expected that 6-year-olds who overestimated their abilities in the laboratory setting would also experience the most severe day-to-day injuries. Likewise, we expected that children who were highly active, impulsive, and impatient would incur the most severe day-to-day injuries.

METHOD

Participants

Sixty-four children participated, 32 six-year-olds (M = 6 years and 5 months, range = 6 years to 7 years and 1 month) and 32 eight-year-olds (M = 8 years and 7 months, range = 8 years and 1 month to 9 years). There were equal numbers of boys and girls in each age group and experimental condition. Children were recruited from an existing subject registry comprised

of children whose parents had previously expressed interest in child development research. Parents received a letter describing the study followed by a phone call inviting them to participate. Most of the children were Caucasian and from middle or upper-middle class families.

Design and Procedure

Four tasks developed by Plumert (1995) were used to compare children's perceptions of their ability to perform particular physical activities with their ability to actually perform those activities. A figure depicting the apparatus can be found in Plumert (1995). The *vertical reach task* involved removing a toy from a shelf while 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 one-inch increments. The *horizontal reach task* involved reaching out from a squatting position for a toy on a wooden block without touching hands or knees on the floor. Subjects performed the reach while squatting on a flat board with an edge on the front end to keep them from moving their feet closer to the toy. The *stepping task* involved stepping across two sticks on the floor placed parallel to each other. Subjects began the step by putting both toes up against the edge of one of the sticks. The *clearance task* involved moving under a wooden bar attached to two posts without knocking the bar off or putting hands or knees on the floor. Subjects began the clearance task by squatting next to the bar and placing both feet in a box marked by tape on the floor.

There were three variations of each task used for the test trials: (1) the *just-within* version was at the child's maximum level of ability; (2) the *just-beyond* version was 8% above the child's maximum level; and (3) the *well-beyond* version was 13% above the child's maximum level. For all tasks, levels of difficulty were scaled to the abilities of the individual children. For example, if a child's maximum step was 31 inches, the just-within version would be set at 31 inches, the just-beyond version would be set at 33 inches, and the well-beyond version would be set at 35 inches.

Estimates for children's 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 actions used to assess children's maximum levels of ability were similar to, but not identical to the actions used to perform the four tasks described above. The estimate for the vertical reach task was obtained by having children stand facing a wall and reach as high as they could on tip toes. The height of their reach was marked at the second joint of the index finger. The estimate for the horizontal reach task was obtained by having children squat down behind a piece of tape marked on the floor and reach forward as far as they could. The experimenter repeatedly placed his/her pencil just out of reach and asked children to try to touch the top of the pencil without falling over. The extent of the reach was marked at the last place at

which children could touch the pencil. This procedure was adopted because previous research revealed that children often reached further when a goal was present. The estimate for the stepping task was obtained by having children 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. Children then were asked to go back to the starting line and try to step further. This procedure was repeated until the experimenter judged that children had stepped as far as they could. Again, this procedure was adopted because previous research revealed that children often stepped further when a goal was present than when absent. The estimate for the clearance task was obtained by having children squat parallel to a wall and duck their heads down as low as possible. The height of the crouch was marked one inch higher than the highest point on the children's back. This measurement was adopted because previous research revealed that children required an extra inch of height in order to move in a squatting position underneath a bar. All measurements were rounded to the nearest inch.

After threshold estimates were taken, one experimenter prepared the apparatus in the testing room while the second experimenter occupied the child with a filler task. The filler task involved listening to several short vignettes and answering related questions. The vignettes were unrelated to the hypotheses of this study and the results are presented elsewhere (Martin, Plumert, Schwebel, & Nichols-Whitehead, 1996). Immediately after entering the testing room, children watched a short video depicting another child performing the four tasks. The purpose of the video was to provide children with social comparison information. Before starting the video, the experimenter explained to children that they would be seeing another child their same age playing the game. They were told that the video would help them understand how to play the game better. The video began with an experimenter explaining the rules of the game to a child actor of the same sex as the child. The child actors were told that before they attempted an activity, they had to decide whether they could do it correctly. Children in the *peer success* condition watched a video in which the peer always judged that he or she could perform the activities and then also succeeded on each activity. Children in the *peer failure* condition watched a video in which the peer always judged that he or she could perform the activities but then failed on each activity. The child in the video performed each of the four tasks once. The videos were approximately $3\frac{1}{2}$ minutes long.

After watching the video, the experimenter told children that they would be playing the same game as the child on the video. The experimenter first explained how to perform each activity. The experimenter then gave children $8 of play money and explained how they could win and lose money in the game. Children were told that they would be asked to decide whether or not they could perform each one of the four activities. Children 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 another dollar, but if they failed they would have to pay the experimenter a dollar. If they decided they could not perform a given activity, the experimenter explained that they would not be asked to try it nor would they lose or win any money. There was no limit on the amount of time children were given to make a decision on each trial. Finally, children were informed that at the end of the game, the experimenter would add up their dollars and award prizes. At the end of all of the trials, the tasks for which children judged they were unable to perform were set up again and children 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 children's judgments with their actual ability to perform the tasks.

There was a total of 12 test trials. During the first block of trials, all four tasks were set at the well-beyond level of difficulty because we reasoned that children would be more cautious about their ability to perform tasks at this level than at the other levels of difficulty. Hence, it seemed likely that children would be most sensitive to social comparison information when making judgments about their ability to perform tasks well beyond their ability. At the end of the first block of trials, children were taken out of the room to resume the filler task while the second experimenter adjusted each apparatus for the next block of trials. This ensured that children were not given information about whether each task was adjusted to be easier or harder. For the second and third blocks of trials, children received tasks that were just within their ability and tasks that were just beyond their ability. The two levels of difficulty (just-within and just-beyond) were randomized across these last 8 trials with the constraint that children received each level of difficulty only once for each type of task and received two trials of each difficulty level in each block of trials. The entire session was videotaped via a Panasonic camcorder through a one-way mirror. The camcorder was adjusted on a tripod so that the apparatus for each task was visible from a single vantage point. Children received two gifts for participating in the laboratory session.

While children were playing the game, parents filled out two temperament questionnaires. These questionnaires were filled out by mothers, with the exception of three children (5% of the sample) whose fathers completed the questionnaires. Parents completed seven scales from the Child Behavior Questionnaire (CBQ; Rothbart et al., 1994), a widely used measure of children's temperament. Parents rated their children on each item using a 7-point Likert scale. Previous studies suggest good internal reliability (Cronbach's alphas range from .67 to .94; Rothbart et al., 1994). Four of the CBQ scales used in this study were taken from the general factor labeled Surgency: approach, high intensity pleasure, activity level, and impulsivity. Approach measured the amount of excitement and positive anticipation a child exhibits for expected pleasurable activities. High intensity pleasure refers to the

amount of enjoyment a child exhibits in response to situations involving high stimulus intensity, rate, complexity, novelty, and incongruity. Activity level describes the child's overall level of gross motor activity. Impulsivity refers to how quickly a child responds to situations and how likely he or she is to rush into activities without thinking about them. One scale, inhibitory control, was taken from the general factor labeled Effortful Control. Inhibitory control refers to the child's capacity to plan and to suppress inappropriate approach responses under instructions or in novel or uncertain situations. Finally, two scales were taken from the general factor of Negative Affectivity: discomfort and fear. Because preliminary analyses indicated that these two scales were unrelated to the variables of interest in this study, they are not discussed further.

Parents also completed the Matthews Youth Test for Health (MYTH; Matthews & Angulo, 1980). As with the CBQ, parents rated their children on each item using a 7-point Likert scale. Although originally developed to measure Type A personality in children, the MYTH currently is used as a general measure of children's temperament. It has good test-retest reliability (r = .79–.82 across three months) and good internal reliability (Cronbach's alphas range from .88 to .90; Matthews & Angulo, 1980; Matthews & Avis, 1983). The MYTH is comprised of two scales: impatience and competitiveness. Impatience refers to the child's eagerness to begin activities and willingness to wait for others. Because competitiveness is not directly relevant to the variables of interest in this investigation, it is not discussed further.

After children finished the laboratory session, the experimenter explained the Daily Injury Diary to the children and their parents. Each evening for 14 days, children and their parents discussed the day's injuries and recorded them in the diary. The type, circumstances, and treatment of each injury was recorded. For example, one entry may have noted that the child inflicted a scraped knee while playing on the swingset with her older brother and a friend; the scraped knee was treated by washing it in warm water and using a Band-Aid. The return rate for the diaries was 92%. Diaries were returned for all 32 of the 6-year-olds and for 27 of the 32 eight-year-olds. Families who sent back the completed diaries received $10 in compensation.

Measures

Children received three 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 the 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 child was *able* to perform a task that was at the just-within level of difficulty, or was *unable* to perform a task that was at the just-beyond or well-beyond levels of difficulty.

TABLE 1
Means and Standard Deviations for Temperament Measures

Measure	Total	6-year-olds		8-year-olds	
		Boys	Girls	Boys	Girls
Approach	5.11 (.56)	5.24 (.54)	5.06 (.36)	5.15 (.70)	4.99 (.61)
High Intensity Pleasure	4.66 (.88)	4.86 (.75)	4.51 (.90)	4.82 (1.01)	4.45 (.84)
Activity Level	4.54 (.95)	4.89 (.85)	4.36 (.75)	4.90 (1.07)	4.00 (.88)
Impulsivity	4.36 (.79)	4.49 (.82)	4.27 (.82)	4.40 (.85)	4.30 (.71)
Inhibitory Control	4.89 (.95)	4.81 (.98)	5.00 (.54)	4.33 (1.00)	5.47 (.93)
Impatience	4.01 (.87)	4.01 (.93)	4.08 (.71)	4.32 (.93)	3.62 (.81)
Surgency/Undercontrol	4.30 (.68)	4.46 (.70)	4.21 (.57)	4.54 (.76)	3.98 (.59)

Notes. Standard deviations are in parentheses. The only significant difference was between 8-year-old boys and girls' scores for Inhibitory Control.

Incorrectly scaled tasks were excluded from the analyses. For example, if a child answered ''yes'' to a task set at the just-beyond level of difficulty and then actually succeeded on that task, that particular trial would be excluded from analysis. Note that when a task was incorrectly scaled at one level of difficulty, it was likely that the other two levels of difficulty also were off to some degree. For example, success on a task set at the just-beyond level of difficulty made it likely that the just-within and the well-beyond levels of difficulty for that task did not correspond exactly to the child's maximum level and 13% above the child's maximum level, respectively. In such cases, however, only the just-beyond level would be dropped from the analyses.

Children also received three 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 children to make a judgment to the point at which children answered in the affirmative or negative. Two coders recorded the latencies from videotapes on a Macintosh IIfx computer for the entire set of children. Pearson correlations were used to calculate reliabilities on decision times for 16 randomly selected children (25% of the sample). Reliability was high, $r = .99$, with an average difference between coders of 276 ms.

Children also received scores for five scales from the CBQ (approach, high intensity pleasure, activity level, impulsivity, and inhibitory control) and one scale from the MYTH (impatience). These scores were calculated by averaging the ratings parents gave for the items from each scale. Means and standard deviations for the six temperament variables are shown in Table 1. Cronbach's alphas for the six scales were good, ranging from .71 to .90. Temperament data for one 8-year-old girl were not available because the parent took the temperament questionnaires home from the laboratory and did not return

them. Scores for the temperament variables were entered into separate Age × Sex Analyses of Variance (ANOVA) to examine potential moderating effects of age and gender. These analyses revealed that boys ($M = 4.89$, $SD = .95$) had higher ratings for Activity Level than did girls ($M = 4.18$, $SD = .82$), $F(1,59) = 10.05$, $p < .01$. There was also a significant Age × Sex interaction for inhibitory control. As shown in Table 1, there was no difference between the 6-year-old boys ($M = 4.81$, $SD = .98$) and girls ($M = 5.00$, $SD = .54$), but 8-year-old girls ($M = 5.47$, $SD = .93$) had significantly higher ratings for Inhibitory Control than did 8-year-old boys ($M = 4.33$, $SD = 1.00$). No other effects were significant.

To provide a more general picture of children's temperament, we created an aggregated temperament measure (Cronbach's alpha = .85) called Surgency/Undercontrol. This score was calculated by averaging scores on the six temperament scales (inhibitory control was reversed). To assess the potential moderating effects of age and sex, this score was entered into an Age × Sex ANOVA. This analysis revealed that boys ($M = 4.49$, $SD = .72$) were rated higher on Surgency/Undercontrol than were girls ($M = 4.10$, $SD = .58$), $F(1,59) = 5.70$, $p < .05$. The Surgency/Undercontrol score was used as the temperament measure in all analyses that follow.

Finally, children received a score representing the average severity of the injuries reported in the Daily Injury Diary. Injury scores were computed by summing severity ratings and dividing by the number of accidents parents reported in the injury diary. The severity of injuries was used as the measure of accident proneness to partially circumvent biases in the reporting of injuries. For example, one parent might report 10 minor injuries that required no treatment while another parent might report 10 minor injuries that required some home first aid, but fail to report the minor injuries that required no treatment. In this example, the two children would have the same number of reported injuries even though one child incurred more severe injuries than the other child. Therefore, we decided to focus on the severity of injuries rather than on the number of injuries. Each injury reported was given a score of either 1 (minor injury requiring no treatment), 2 (minor injury requiring home first aid), 3 (major injury requiring home first aid), or 4 (major injury requiring medical treatment). Pearson correlations were used to calculate reliabilities on 16 (25% of the sample) randomly selected children. Reliability was high, $r = .98$.

RESULTS

We examined three primary issues: (a) whether children used the performance of the peer to inform them about their own abilities; (b) whether children characterized as highly active, impulsive, and impatient were more likely to overestimate their abilities; and (c) whether ability overestimation and temperamental characteristics were related to accident proneness. Preliminary analyses that were 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.

Validity of Difficulty Levels

An initial set of analyses was conducted to evaluate the validity of the threshold measures used to scale the tasks. An Age (6 years vs 8 years) × Difficulty Level (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 main effect of difficulty level, $F(2,120) = 11.34$, $p < .0001$, but no effect of age and no interaction between age and difficulty level. The mean number of correctly scaled tasks included for the just-within, just-beyond, and well-beyond levels of task difficulty was 3.17 ($SD = .81$), 3.27 ($SD = .67$), and 3.72 ($SD = .45$) out of 4, respectively. The number of tasks included at each level ranged from 2 to 4. Follow-up tests revealed that significantly more tasks at the well-beyond than either the just-within or just-beyond level of difficulty were scaled correctly. There was no significant difference between the just-within and just-beyond levels. As in Plumert (1995), it was easier to scale tasks at the well-beyond level than at either the just-beyond or the just-within level.

Accuracy of Judgments

Two sets of analyses were conducted on the accuracy of children's judgments about their ability to perform the tasks. The first evaluated children's judgments about tasks well beyond their ability and the second evaluated children's judgments about tasks just beyond and just within their ability. Analyses of children's judgments about the well-beyond level were conducted separately because all children received this set of tasks immediately after viewing the peer in the video. Accuracy scores for the well-beyond level of difficulty were entered into an Age (6 years vs 8 years) × Condition (peer success vs peer failure) ANOVA. This analysis yielded significant effects of age, $F(1,60) = 13.34$, $p < .001$, and condition, $F(1,60) = 4.12$, $p < .05$. Although both age groups overestimated their ability to perform tasks well beyond their ability, 8-year-olds ($M = 41\%$, $SD = 22$) made a greater percentage of accurate judgments than did 6-year-olds ($M = 23\%$, $SD = 20$). Moreover, children who watched the peer fail on the tasks ($M = 37\%$, $SD = 25$) made a greater percentage of accurate judgments than did children who watched the peer succeed on the tasks ($M = 27\%$, $SD = 20$). In other words, children who watched the peer fail were more conservative about their own ability to perform the tasks than were children who watched the peer succeed.

The second set of analyses evaluated children's judgments about the just-beyond and just-within levels of difficulty. Accuracy scores were entered into

TABLE 2
Stability of Judgments across Tasks within Each Level of Difficulty

Response category	Difficulty level		
	Well-beyond level	Just-beyond level	Just-within level
Always same judgment	20%	31%	88%
Same judgment on three-fourths of tasks	44%	16%	4%
Same judgment on two-thirds of tasks	19%	31%	5%
Same judgment on half of tasks	17%	22%	3%

Note. Percentages refer to the percentage of children in each response category.

an Age (6 years vs 8 years) × Condition (success vs failure) × Difficulty Level (just-beyond vs just-within) 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, $F(1,60) = 6.27$, $p < .05$, and difficulty level, $F(1,60) = 123.51$, $p < .001$, but no effect of condition, $F(1,60) = 1.96$, ns. The 8-year-olds ($M = 72\%$, $SD = 32$) made a greater percentage of accurate judgments than did 6-year-olds ($M = 63\%$, $SD = 40$), and children of both ages made a greater percentage of accurate judgments about tasks just within their ability ($M = 94\%$, $SD = 18$) than about tasks just beyond their ability ($M = 41\%$, $SD = 30$). There was also a significant Age × Difficulty Level interaction, $F(1,60) = 9.59$, $p < .01$. Simple effects tests revealed that the percentage of correct judgments that 6- and 8-year-olds made about tasks just within their ability did not differ significantly ($M = 96\%$, $SD = 12$ and $M = 91\%$, $SD = 23$, respectively), $F(1,62) = 1.57$, ns, but that 8-year-olds made a greater percentage of accurate judgments about tasks just beyond their ability ($M = 53\%$, $SD = 28$) than did 6-year-olds ($M = 29\%$, $SD = 27$), $F(1,62) = 11.35$, $p < .01$.

Stability of Judgments

We also examined individual patterns of responding to determine whether children's judgments were stable across the different tasks within each level of difficulty. Table 2 shows the percentage of children who gave the same judgment all of the time, three-fourths of the time, two-thirds of the time, or half of the time for each level of task difficulty. The two-thirds response category was included because responses for only three out of the four tasks were included for some children due to inaccurate scaling of tasks. It appears that the majority of children made reasonably consistent responses across the different tasks, although a sizeable proportion of children responded inconsistently.

Decision Times

Two sets of analyses were conducted on the length of time children took to make judgments about their ability to perform the tasks. The first evaluated children's decision times about tasks well beyond their ability and the second evaluated children's decision times about tasks just beyond and just within their ability. Again, analyses of children's decision times about the well-beyond level were conducted separately because all children received this set of tasks first following the peer manipulation. Prior to statistical analysis, trials for which decision times 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. Extremely long decision times usually reflected children being off task and hence added unwanted noise to the data. The number of individual trials removed for 6-year-olds and 8-year-olds was 13 and 11, 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.

To analyze decision times about tasks well beyond children's ability, latency scores were entered into an Age (6 years vs 8 years) $\times$ Condition (success vs failure) ANOVA. This analysis yielded no significant effects. The overall mean latency for the well-beyond level of difficulty was 2459 ms (SD = 2226). In the analysis of decision times about tasks just within and just beyond children's ability, latency scores were entered into an Age (6 years vs 8 years) $\times$ Condition (success vs failure) $\times$ Difficulty Level (just-within vs just-beyond) repeated measures ANOVA with the first two factors as between-subjects variables and the third as a within-subjects variable. This analysis yielded a significant effect of difficulty level, $F(1,60) = 20.67$, $p < .001$, indicating that the mean decision time for the just-beyond level (M = 2209 ms, SD = 1824) was significantly longer than that for the just-within level (M = 1155 ms, SD = 820). Thus, children made quicker decisions about tasks that were just within their ability than about tasks that were just beyond their ability.

Temperamental Characteristics and Judgments about Physical Abilities

To address how individual differences in temperamental characteristics were related to judgments about physical ability, correlations were conducted for each age group between accuracy scores for the well-beyond level of difficulty and the aggregated Surgency/Undercontrol temperament measure. As shown in Table 3, the correlations differed widely for the two age groups. Surgency/Undercontrol was related to accuracy of judgments for 6-year-olds, but not for 8-year-olds. Thus, 6-year-olds characterized by high levels of activity, impatience and strong approach tendencies and low levels of inhibitory control made less accurate judgments about their ability to perform tasks well beyond their ability. Correlations also were conducted separately for

TABLE 3
Correlates of Surgency/Undercontrol: Judgment Accuracy, Decision Times, and Accident Severity

	Surgency/Undercontrol	
	6-year-olds	8-year-olds
Judgments accuracy		
Well-beyond level	−.47**	−.03
Just-beyond level	−.31	−.10
Decision times		
Well-beyond level	−.22	−.14
Just-beyond level	−.11	−.46*
Accident severity	−.05	.54**

* $p < .05$.
** $p < .01$.

each age group between accuracy scores for the just-beyond level of difficulty and the aggregated Surgency/Undercontrol temperament measure. Neither of these two correlations reached significance, although there was a trend for 6-year-olds to overestimate their ability to perform tasks at this level of difficulty as well (see Table 3). Correlations between Surgency/Undercontrol and accuracy of judgments on the just-within level of difficulty were not conducted because accuracy scores for this difficulty level were near ceiling.

To assess whether temperamental characteristics were related to how quickly children made judgments about their ability to perform the tasks, correlations between the aggregated temperament variable and latency scores were conducted for each age group. The first set of correlations examined the relations between temperamental characteristics and decision times about tasks well beyond children's ability. These correlations failed to reach significance for either age group (see Table 3). The second set of correlations examined the relations between temperamental characteristics and decision times about tasks just beyond children's ability. As shown in Table 3, 8-year-olds characterized as high on Surgency/Undercontrol made quicker decisions about their ability to perform tasks that were just beyond their ability. The correlation between decision times and Surgency/Undercontrol did not reach significance for the 6-year-olds.

Accidental Injuries

The findings reported above provide insight into cognitive and temperamental factors that may contribute to childhood accidents. These results do not, however, provide information about whether these skills are related to individual differences in children's *actual* accident histories. The measure of accidental injuries we used was the average severity of children's day-to-day injuries

reported over a two-week period. Average severity ratings ranged between 1.00 and 1.83 ($M = 1.23$, $SD = .21$). As expected, most of the everyday accidents children experienced were very minor and did not require treatment. The lowest score children received for a particular injury was a 1 (i.e., minor injury requiring no treatment) and the highest score was a 3 (major injury requiring home first aid). None of the children received the highest severity score for any injury (i.e., major injury requiring medical treatment). Accident severity did not differ as a function of age, $F(1,55) = .21$, ns, or gender, $F(1,55) = .57$, ns.

Accuracy of judgments and accident severity. Correlations were conducted separately for each age group between accuracy scores for the well-beyond level of difficulty and average injury severity ratings. These analyses revealed no significant effects. The correlation between accuracy scores and severity ratings for 6- and 8-year-olds were $r = -.22$ and $r = -.03$, respectively. Closer inspection of the data, however, revealed that the relation between judgment accuracy and injury severity differed for males and females across the two age groups. Specifically, there was a significant negative correlation between the accuracy of 6-year-old boys' judgments of their physical abilities and average injury severity, $r = -.51$, $p < .05$. Thus, 6-year-old boys who were less accurate in judging their physical abilities in the laboratory setting experienced more severe day-to-day injuries. There was no relation between 6-year-old girls' judgments of their physical abilities and average injury severity, $r = .05$, ns. Likewise, the relation between judgment accuracy and injury severity was not significant for 8-year-old boys ($r = -.01$, ns) or for 8-year-old girls ($r = -.10$, ns).

Temperamental characteristics and accident severity. To examine the relations between temperamental characteristics and accidental injuries, we conducted separate correlations for each age between the Surgency/Undercontrol temperament measure and the average injury severity measure. As shown in Table 3, the correlations again differed for the two age groups. Eight-year-olds characterized by high levels of gross motor activity, intense excitement for pleasurable activities, enjoyment of high intensity situations, and fast speeds of response initiation incurred more severe day-to-day injuries ($r = .54$). The correlation between temperament and injury severity was nonsignificant for 6-year-olds ($r = -.05$), however.

Multivariate analysis of accident severity. A stepwise multiple regression analysis was conducted to examine the unique contributions of age, temperament, and ability overestimation to accidental injuries. The accuracy score for the well-beyond level of difficulty was used as the measure of ability overestimation, and the Surgency/Undercontrol score was the measure of temperament. Scores for temperament, ability overestimation, and accident severity were standardized prior to analysis. The following predictors were entered into the regression equation: Age, Ability Overestimation, Temperament, Age $\times$ Temperament, and Age $\times$ Ability Overestimation. The two

interaction terms were entered because the correlational analyses reported above showed that the relations between accident severity, ability overestimation, and temperament differed for the two ages. Only one predictor, Age $\times$ Temperament ($\beta = .58$), emerged as a significant predictor of accident severity, $R^2_{\text{adjusted}} = .16$, $F(1,57) = 12.13$, $p < .01$. The results of this analysis offer further evidence that temperament operated differently at the two ages.

DISCUSSION

The results of this investigation suggest that both temperamental characteristics and social comparison information play a role in children's judgments about their physical abilities. When children observed the peer in the video fail on all of the activities, they were more conservative about their own abilities than when they observed the peer succeed. Thus, both 6- and 8-year-olds used the peer's performance to inform them about their own abilities. The relations between temperamental characteristics and children's judgments about their physical abilities differed for the two age groups, however. For 6-year-olds, temperamental characteristics were correlated with the accuracy of their judgments but not with the amount of time they took to make their judgments. Specifically, children who were more active and undercontrolled made less accurate judgments about their abilities. For 8-year-olds, however, the same temperamental characteristics were correlated with the amount of time they took to make their judgments, but not with the accuracy of their judgments. Eight-year-olds who were more active and undercontrolled made quicker judgments about their ability to perform the tasks. The relations between temperament, ability overestimation, and accident severity also differed for the two age groups. Specifically, ability overestimation was related to accident severity for 6-year-olds males whereas temperamental characteristics were related to accident severity for 8-year-olds.

The finding that children were more cautious about their own abilities after observing a peer fail than succeed on the activities demonstrates that social comparison information plays a role in children's judgments of their physical abilities. Interestingly, social comparison information influenced children's judgments about tasks well beyond their ability, but not their judgments about tasks just beyond or just within their ability. There are at least two possible explanations for this finding. One is that children are more sensitive to social comparison information in situations where they are particularly unsure about their abilities. In the present study, children may have been less sure about their ability to perform the tasks well beyond their ability than the tasks that were just within or just beyond their ability. Another possible explanation is that the initial information children gained through watching someone else perform the tasks was supplanted by direct experience with performing the tasks. That is, social comparison information only influenced children's judgments about the well beyond tasks because they were performed first. As Ruble and Frey (1991) suggest, the function of social comparison information

may change as children gain more experience with performing a task. In this study, the peer's performance provided useful information for making an initial task assessment because children had no way of knowing how difficult the tasks were before they began. After experiencing the tasks first-hand, children no longer needed the peer's performance to inform them about the difficulty of the tasks because more objective information was available. One possible implication of this finding is that faulty use of social comparison information may be a risk factor for accidental injuries only when children have little or no experience with a particular activity. Further research is needed, however, to disentangle these alternative interpretations.

The results of this investigation also suggests that the relation between children's temperamental characteristics and their judgments about their physical abilities varies with development. Specifically, temperamental characteristics were correlated with the accuracy of 6-year-olds' judgments but not with the amount of time they took to make their judgments. Thus, 6-year-olds who were characterized as highly active and undercontrolled made more errors in judging their physical abilities. For 8-year-olds, these same temperamental characteristics were not related the accuracy of their judgments. These differences may be the result of developmental changes in how children assess their physical capabilities. In both this investigation and in Plumert (1995), 8-year-olds made significantly more accurate judgments about their abilities than did 6-year-olds. Older children may have more conservative response criteria than do 6-year-olds because they are better able to anticipate the consequences of making an incorrect judgment. By age 8, children may have developed some of the cognitive skills necessary to make accurate judgments about their physical abilities. At age 6, however, these cognitive skills are more fragile and easily disrupted. As a result, temperamental characteristics may influence the accuracy of younger, but not older, children's judgments about their physical abilities.

Although temperament was not significantly related to the accuracy of older children's judgments about their physical abilities, temperament played a role in other aspects of their judgments. Specifically, 8-year-olds who were characterized as highly active and undercontrolled made faster judgments about their ability to perform tasks just beyond their ability. Thus, the same temperamental characteristics that were related to the accuracy of 6-year-olds' judgments were also related to subtler aspects of older children's judgments about their physical abilities. Clearly, the amount of time children take to make judgments taps a more refined aspect of their decision-making than does the content (yes-no) of their judgments. This is particularly true for highly ambiguous tasks such as those just beyond children's ability. Thus, it appears that as children's cognitive skills advance, temperamental influences shift to subtler aspects of cognitive processing.

Finally, this investigation revealed that the links between temperamental characteristics, ability overestimation, and accidental injuries differed for 6-

and 8-year-olds. For 6-year-old males, overestimation of ability in the laboratory setting was related to the severity of the day-to-day injuries that they experienced. This relation held even though the consequences for making an incorrect judgment are much less severe in the laboratory than in real life. Together, the findings of this study and Plumert (1995) suggest that the relation between ability overestimation and accidental injuries holds for both major injuries and more minor day-to-day injuries. We found no relation between overestimation of ability and accidental injuries for 8-year-olds. Instead, we found relations between temperament and accidental injuries for 8-year-olds. Eight-year-olds who were characterized as highly active and undercontrolled had more severe day-to-day injuries. The correlation between Surgency/Undercontrol and accident proneness did not even approach significance for 6-year-olds.

These findings suggest that the factors that put younger and older children at risk for accidents may differ. For younger children, ability overestimation may mediate the relation between temperamental characteristics and accident proneness. That is, temperamental characteristics such as high impulsivity and low inhibitory control put younger children at risk for accidents because these characteristics lead to errors in children's judgments about their abilities. When judging whether it is safe to cross a street, for example, children need to stop and think before running across. Impulsive children may be more prone to errors in judgment than children who make decisions more carefully. The significant correlations between temperamental characteristics and ability overestimation and between ability overestimation and accident proneness supports such a model. Further research is warranted, however, before any definitive conclusions can be drawn about potential mediating effects of ability overestimation on the relation between temperament and accident proneness in younger children.

The fact that temperamental characteristics were correlated with accident severity but not overestimation of ability in 8-year-olds suggests that temperament may exert a more direct influence on accident proneness in older children. That is, older children who are highly active and undercontrolled may have more opportunities to injure themselves because they seek out dangerous situations. Unlike younger children, older children may engage in risky behavior not because they cannot assess danger accurately, but because they enjoy being in dangerous situations. Furthermore, older children experience less direct supervision from adults which may compound accident risk for children prone to ''sensation-seeking'' behavior.

More generally, the results of this investigation have implications for understanding the relations between temperament and cognitive development. In particular, these results suggest that the contribution of temperament to cognitive skills may change with development. Although temperamental characteristics may remain stable (e.g., Kochanska, Murray, & Coy, 1997), the contribution of temperament to cognitive skills may change as those skills undergo

development. In other words, temperamental influences on cognitive processing may be most apparent when children are in the process of acquiring a skill. During times of transition, performance is variable and may be more susceptible to other influences such as temperament. As particular cognitive skills mature and become less variable, however, there is less room for other factors such as temperament to influence performance. In the present investigation, the same temperamental characteristics were related to two different aspects of children's judgments, judgment accuracy and decision times, at the two ages studied. For 6-year-olds, temperament was related to the accuracy of judgments, whereas for 8-year-olds, temperament was related to how quickly children made their judgments. It seems plausible that because 6-year-olds' ability to make accurate judgments about their physical capabilities is somewhat fragile and easily disrupted, temperamental characteristics such as impulsivity and approach played a role in their judgments. Eight-year-olds, in contrast, have developed a better ability to make accurate judgments about their physical abilities. Therefore, temperamental characteristics manifested themselves in the amount of time they took to make their judgments. Impulsive children, for instance, decided quickly without careful thought, whereas more inhibited children considered their options carefully before making a judgment.

A note of caution is warranted about our conclusions concerning the relations between temperament, ability overestimation, and accidental injuries. Accidental injuries are complex phenomena that are mediated by many factors besides temperamental characteristics and ability overestimation. These include factors such as motor coordination, parental supervision, anticipation of consequences, and knowledge about safety rules. Although further research is needed to determine how multiple factors work together to put children at risk for accidents, these findings represent a step forward in our understanding of the complex interplay between temperament and ability overestimation and its links to accident proneness in children.

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.

Bijur, P., Golding, J., Haslum, M., & Kurzon, M. (1988). Behavioral predictors of injury in school-age children. *American Journal of Diseases of Children,* **142,** 1307–1312.

Festinger, L. (1954). A theory of social comparison. *Human Relations,* **7,** 117–140.

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

Jaquess, D. L., & Finney, J. W. (1994). Previous injuries and behavior problems predict children's injuries. *Journal of Pediatric Psychology,* **19,** 79–89.

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development,* **68,** 263–277.

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

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

Manheimer, D., & Mellinger, G. (1967). Personality characteristics of the child accident repeater. *Child Development,* **38,** 491–514.

Martin, R., Plumert, J. M., Schwebel, D., & Nichols-Whitehead, P. (1996, May). *Common sense models of stress and illness in children.* Poster presented at the Annual Meeting of the Midwestern Psychological Association, Chicago, IL.

Matheny, A. P. (1986). Injuries among toddlers: Contributions from child, mother, and family. *Journal of Pediatric Psychology,* **11,** 163–176.

Matthews, K. A., & Angulo, J. (1980). Measurement of the Type A behavior pattern in children: Assessment of children's competitiveness, impatience-anger, and aggression. *Child Development,* **51,** 466–475.

Matthews, K. A., & Avis, N. E. (1983). Stability of overt Type A behaviors in children: Results from a one-year longitudinal study. *Child Development,* **54,** 1507–1512.

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.

Nyman, G. (1988). Infant temperament, childhood accidents, and hospitalization. In S. Chess, A. Thomas, & M. E. Hertzig (Eds.), *Annual progress in child psychiatry and child development.* New York: Brunner/Mazel.

Pless, I. B., Taylor, H. G., & Arsenault, L. (1995). The relationship between vigilance deficits and traffic injuries involving children. *Pediatrics,* **95,** 219–224.

Plumert, J. M. (1995). Relations between children's overestimation of their physical abilities and accident proneness. *Developmental Psychology,* **31,** 866–876.

Potts, R., Martinez, I. G., & Dedmon, A. (1995). Childhood risk taking and injury: Self-report and informant measures. *Journal of Pediatric Psychology,* **20,** 5–12.

Pulkkinen, L. (1995). Behavioral precursors to accidents and resulting physical impairment. *Child Development,* **66,** 1660–1679.

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

Rothbart, M. K., Ahadi, S., & Hershey, K. L. (1994). Temperament and social behavior in children. *Merrill-Palmer Quarterly,* **40,** 21–39.

Ruble, D. N., & Frey, K. S. (1991). Changing patterns of comparative behavior as skills are acquired: A functional model of self-evaluation. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 79–113). Hillsdale, NJ: Lawrence Erlbaum.

Singh, G. K., & Yu, S. M. (1996). U.S. childhood mortality, 1950–1993: Trends and socioeconomic differentials. *American Journal of Public Health,* **86,** 505–512.

Wheeler, L., Martin, R., & Suls, J. (1997). The proxy model of social comparison for self-assessment of ability. *Personality and Social Psychology Review,* **1,** 54–61.

Received: October 14, 1996; revised: July 10, 1997