Psicológica (2010), 31, 677- 687

Skill-Related Uncertainty and Expected Value in 5- to 7-Year-Olds

Sarah Bayless[1] & Anne Schlottmann[2]

University of Winchester, UK[1]; University College London[2], UK

Studies using an Information Integration approach have shown that children from four years have a good intuitive understanding of probability and expected value. Experience of skill-related uncertainty may provide one naturalistic opportunity to develop this intuitive understanding. To test the viability of this view, 16 5- and 16 7-year-olds played a marble rolling game in which size of the target and distance from it varied factorially. Task difficulty judgements (prior to practical experience with the game) reflected both objective task structure and subsequent performance for both age groups. Children then judged how happy they would be playing games of variable difficulty for different prizes. These judgements had the multiplicative structure predicted by the normative expected value model, again for both age groups. Thus children can use task difficulties as estimates of personal success probability in skill-related tasks. Our findings therefore extend previous work on early probability understanding from games of chance to games of skill.

The ability to evaluate personal success probability is important for efficient behaviour in situations of uncertainty. The present study investigates how young children evaluate success probabilities and utilities of outcomes in situations that depend on skill, in order to learn more about the natural sources of early intuitive probability understanding.

Contrary to traditional theory (Hoemann & Ross, 1982; Piaget & Inhelder, 1958, 1975), recent work using an Information Integration approach (Anderson, 1981, 1982, 1991, 1996) has shown that children from 4 years of age have good intuitive understanding of probability and expected value (EV) (see review in Schlottmann & Wilkening, 2010, in press). For instance, children's judgements (on a continuous graphic scale) of how easy it is to randomly draw a blue winner marble from a plate with blue and black marbles vary appropriately with the number of winners and losers. More importantly, the barrel-shaped pattern of judgements corresponds to the predictions of the normative probability ratio model (Anderson & Schlottmann, 1991; also see Acredolo, O'Connor, Banks, & Horobin, 1989; Wilkening & Anderson, 1991). Similarly, children's judgements of how good it is to play a game of chance for a prize vary appropriately with the

[1]Department of Psychology, University of Winchester, Sparkford Road, Winchester, SO22 4NR, Email: sarah.bayless@anapoda.co.uk; [2]Division of Psychology & Language Sciences, University College London, Gower Street, WC1E 6BT, Email: a.schlottmann@ucl.ac.uk.

likelihood of winning and size of the prize. Crucially, the judgements follow a fan-shaped pattern, which corresponds to the predictions of the normative model in which EV is defined as the product of probability and value (Schlottmann, 2001; Schlottmann & Anderson, 1994; Schlottmann & Tring, 2005). Because the structure of children's judgements is so close to appropriate formal models, they cannot easily be discounted as non-probabilistic. Instead these findings highlight a genuine intuitive probability competence.

The studies demonstrating this intuitive competence have typically used tasks which involve simple games of chance, like the random draw of a marble or spinning of a roulette wheel. Young children have little experience with such devices, so it is remarkable that they can reason about probability in these unfamiliar contexts. Furthermore, it is quite unlikely that that they learned about probability in these contexts. Instead, everyday experiences of uncertainty at achieving desired outcomes may play an important part in developing this intuitive understanding.

Uncertainty may be particularly salient for children in skill-related tasks. Young children are beginning learners of most skills who will experience over and over again that task achievement is not guaranteed. They may fall off their scooter or stay on, may read the correct word or a substitute, their ball may hit or miss the target. The contexts in which children learn about their own performance and its determinants (effort, skill and objective difficulty factors) also provide opportunities to learn about success probability and its implication for outcome. Children's intuitive probability competence is easier to understand if we consider that they may learn in such everyday scenarios which bear little resemblance to the lottery-style tasks typically considered in the judgment-decision literature. The goal of the present study is to investigate whether and how young children assess subjective success probabilities in a skill-dependent task.

The issue of how young children assess and predict performance/ability has been studied before in the areas of memory monitoring (Schneider & Pressley, 1997) and achievement motivation (Stipek, 1984; Stipek & MacIver, 1989). Both lines of work found that young children typically overestimate their performance, gradually becoming more realistic over the school years. This has been attributed to a number of factors: deficient monitoring (Flavell, 1979), overweighting of effort (Wellmann, 1985), non-differentiation of effort and ability (Nicholls, 1978) or non-differentiation of wishes and expectations (Schneider, 1998; Stipek, 1984). Regardless of the precise explanation, unrealistic optimism may boost children's motivation to practice and improve their skills (Schneider, 1998).

Only one study (Schneider, Hanne, & Lehmann, 1989) has related children's performance expectations to their experiences of uncertainty. From 3 years, children discriminated between difficulty levels in a box-lifting task, with no overestimation of success. Subjective feelings of uncertainty were indicated by children taking longer to make verbal predictions for more uncertain task levels, and approaching these faster than task levels in which success/failure was relatively certain. Similar findings appeared for a marble-rolling task, but children discriminated less between difficulty levels and overestimated their success. Schneider et al. (1989) argue that more experience with lifting objects than rolling marbles at younger ages might account for more realistic expectations in the box-lifting task. Here we aim to extend these findings.

The present study addresses two issues. First, we investigated children's judgements of task difficulty in a more complex situation in which difficulty varied along two dimensions simultaneously. Second, we wanted to determine how children's judgements of difficulty correspond to objective success probability, and whether these judgements directly reflect estimates of personal success probability. If so, then children should be able to incorporate task difficulty into EV judgements.

Children were invited to help a puppet play a "shoot the marble through the gate game" in which gate size and distance from the start line varied (see Figure 1). Children first judged the difficulty of each game combination, then played all games, and finally judged how happy the puppet would be playing some games, with difficulty level and size of the prize varied. The latter task is an adaptation of a standard EV task for children (e.g., Schlottmann, 2001).

METHOD

Participants. Thirty-five children took part in the experiment; three were eliminated due to not understanding the task or not paying attention. There were 16 children in the younger age group (range =5,4-6,2, mean age = 5,8); and 16 children in the older age group (range = 6,6–8,2, mean age = 7,5). Children were volunteers from two London primary schools, attended by predominantly white children from middle-class homes.

Materials. The marble game was played on a 60x60cm mat with a start line and three distances (20cm, 40cm, 60cm) marked. Three gates (internal width 2.5cm, 4.5cm, and 6.5cm, marked by different symbols to facilitate discrimination) could be placed at these distances (see Figure 1). Two further gates of 1cm and 7.5cm width were used as anchors. Children were given 3 identical marbles of 1.5cm diameter to play. The prizes were small, medium or large bags of M&M candies, laid out next to a gate during the EV judgements.

Children's judgements were made on a graphic response scale, consisting of 17 wooden dowels increasing in height from 2.5 to 18.5cm, with each stick 1cm taller than the previous one. Children pointed to a stick to indicate how difficult a game would be, or how happy they would be. Bigger sticks corresponded to greater difficulty, or for EV judgements, to better games. Even 4-year-olds can successfully use this scale (Anderson & Schlottmann, 1991; Schlottmann, 2001; Schlottmann & Anderson, 1994). Scale usage was elicited in the standard way by instruction with end anchors (Anderson, 1982, chapter 1).

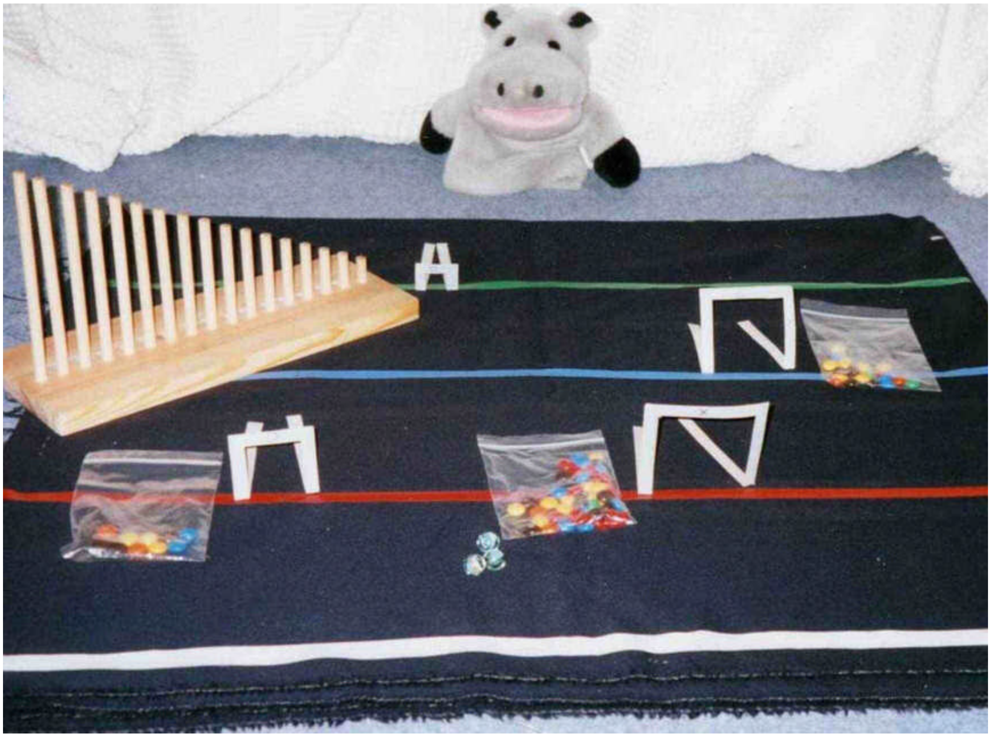

Figure 1. Materials for the marble-rolling task. (Children helped "Hilda Hippo" roll the marble through one of three gates (an anchor gate is also shown). Gates were set up in the centre of the mat at one of the three marked distances and the marble was rolled from the white start line. Bags of M&M candies served as prizes for the EV game. Children made both difficulty and EV judgements on the stick scale on the left.)

Design. The design for the task difficulty judgements and for the subsequent performance task was a 3 gate size x 3 distance within subjects factorial. Children first judged two individually randomized replications of the 9 game combinations, then had three attempts with the marble for each, also in a random sequence. The design for the final EV task was a 3 game (large gate/20cm, medium gate/40cm, small gate/60cm) x 3 prize within subjects factorial. Children again judged two individually randomized replications. Age was a between subjects factor.

Procedure. Children were tested individually in a single session at their school. First the puppet showed the child the marble game and asked the child to help her play. The large anchor gate was placed at the 40cm line and children were encouraged to roll the marble through this gate from the starting line. This was repeated with the small anchor gate. Children then sorted all gates according to difficulty. Following this two identical large anchor gates were placed at the closest and furthest line and children were asked to indicate the easier game. Children had no difficulties with this.

The stick scale was now introduced, with long sticks for good (easy) games, short sticks for bad (hard) games, and medium sticks for ok (not too hard, not too easy) games. The largest stick was associated with the easiest game (large anchor at 20 cm). The smallest stick was associated with the hardest game (small anchor at 60 cm). Pictures of the easy and hard games were placed beside the corresponding ends of the scale throughout the session. Children were shown an easy, medium and hard game and asked to point to a corresponding stick to ensure understanding. The 18 experimental trials followed.

Following the judgements, children were asked which of three gate-distance combinations (Large/20, Medium/40, Small/60) they would like to play first. Responses on 3 rolls for the chosen game were recorded, followed by three rolls for each of the other eight games, presented in a random order.

The expected value task was presented subsequently. Children were told that the puppet might now win a prize if the marble rolled through the gate, and the M&M prizes were shown. The easy anchor game was paired with the largest prize and the difficult anchor game with no prize. Children indicated the better and worse game. They were then instructed to point to a stick to show how good each game was. Pictures of the anchor games were placed by the corresponding scale end and children were reminded to use all sticks so they could evaluate in-between games. The 18 experimental trials followed. The session was concluded by asking children to choose one game combination to play for a sticker prize that they would keep.

RESULTS

Mean judgements of difficulty, task performance and judgements of expected value made by 5- and 7-year olds are presented in Figure 2. These data were submitted to mixed model ANOVAs. Greenhouse-Geisser adjusted degrees of freedom are reported as appropriate.

Task Difficulty Judgements. Children's judgements of task difficulty prior to experience with the task are shown in the top panel of Figure 2. Results of the 3 (Gate Size) x 3(Gate Distance) x 2(Age) ANOVA indicate a main effect of gate size $F(1.41, 42.34)=100.83$, $p<0.001$ and distance $F(2,60)=30.18$, $p<0.001$, but no interaction, $F(4,120)=4.23$, ns. The only significant effect involving age was the gate size x age interaction, $F(1.41, 42.34)=5.80$, $p<0.05$, with a larger effect of gate size for 7- than 5-year-olds, as seen in steeper slopes in the right panel.

Individual subject analyses confirmed that children considered both factors in their judgements. In single subject Anovas ($p<0.1$), 11 of 16 5-year-olds showed a main effect of gate size, nine of distance, six showed both main effects and four showed interactions. For 7-year-olds, all showed a main effect of gate size, nine of distance, nine showed both, and four showed interactions. That few 5-year-olds showed statistical main effects of both variables is due to low power. If either statistically reliable or sizable main effects (means difference 3 points or more) are considered, then 9 5 and 9 7-year-olds show effects of both gate size and distance.

Performance. Children's success in rolling the marble was similarly affected by both gate size and distance, $F(2,60)=9.59$, $p<0.001$ and

$F(2,60)=30.18$, $p<0.001$, respectively, (see middle panel of Figure 2). Thus the structure of the subjective difficulty ratings considered above reflects the objective success probabilities. Despite the irregularity in the bottom line of the performance data (presumably due to the limited number of attempts with each combination), the interaction was not significant, $F(4,120)=1.14$, nor were there significant effects involving age, $F(2,60)<1$.

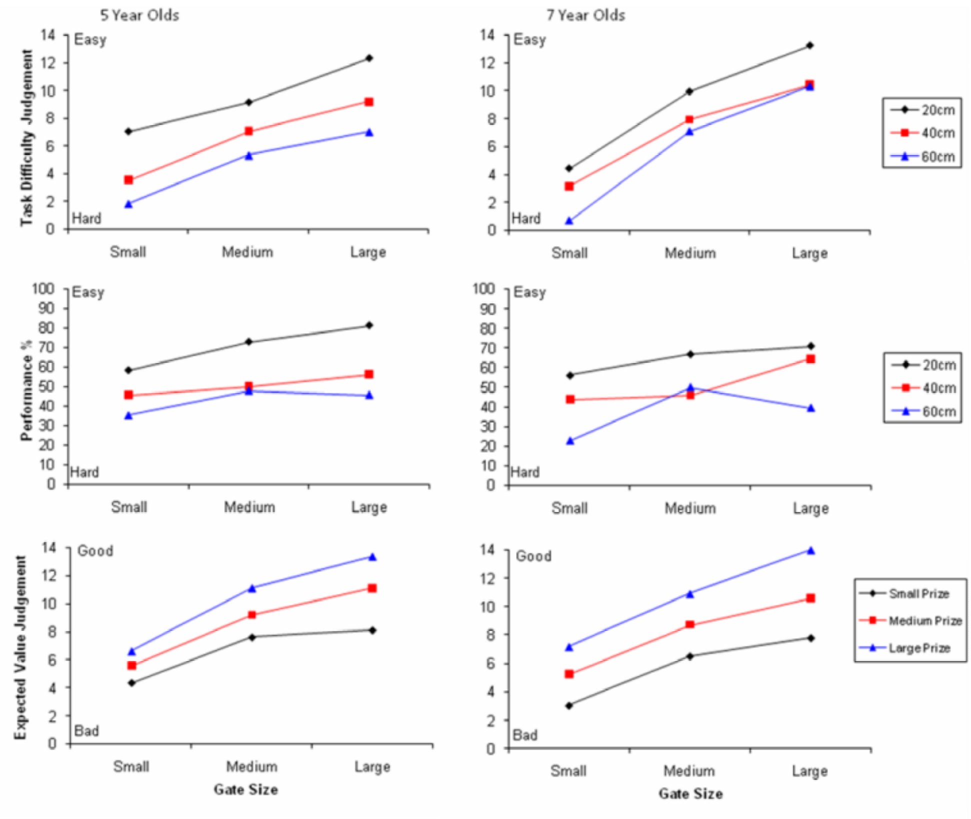

Figure 2. Mean task difficulty judgements (top panels), performance (middle) and EV judgements (bottom) for two age groups. (Difficulty judgements and performance are shown as a function of gate size (horizontal) and distance (curve factor); EV judgements are shown as a function of gate size/distance combination (horizontal) and prize (curve factor). Higher scores indicate easier games, better performance and better games. Effects of both factors appear clearly in all panels, but while the pattern for difficulty judgements and actual performance is near-parallel, EV judgements show the expected normative fan-shape pattern.)

To evaluate quantitative accuracy of the subjective difficulty ratings, we compared these ratings with task performance using a 3(gate size) x 3(distance) x 2(task) x 2(age) mixed model Anova. There was no main effect of Task $F<1$, i.e., there was no overall overestimation of success. The only significant effect involving task was the gate size x task interaction, $F(2,60)=23.99$, $p<0.001$. All other effects were non-significant $F<3.12$, $p>0.62$, except the expected main effects of gate size and distance, $F>64$.

The mean judgements and performance data for gate size, distance and task (collapsed across ages, and with both judgements and performance re-scaled to %) are presented in Table 1. Inspection of these means indicated that regardless of distance, children underestimated the difficulty for the largest gate (greater mean judgements compared to performance), gave reasonably accurate estimations of difficulty for the medium sized gate, and overestimated difficulty of the smallest gate (lower mean judgements compared to performance).

Table 1: Mean judgement and performance scores for each gate size and distance combination. N.B. judgement scores are rescaled from 1-14 to 1-100 to align with the performance scores.

		Judgement	Performance
	20cm	91.26	76.04
Large	40cm	70.28	60.42
	60cm	61.92	42.72
Mean		74.49	59.73
	20cm	68.17	69.80
Medium	40cm	53.55	47.92
	60cm	44.40	48.96
Mean		55.37	55.56
	20cm	41.06	57.29
Small	40cm	23.87	44.79
	60cm	9.04	29.16
Mean		24.66	43.75
Grand Mean		51.51	53.01

Expected Value Judgements. Mean expected value judgements were submitted to a 3(Game Difficulty) x 3(Size of Prize) x 2(Age) mixed model ANOVA. There were significant main effects of difficulty $F(1.53, 45.92)=54.36$, $p<0.001$ and prize $F(1.34, 40.27)=35.16$, $p<0.001$. Importantly, the interaction between gate and prize was also significant $F(3.65, 109.50)=3.37$, $p<0.05$, with a linear x linear component, $F(1,30)=12.77$, $p<0.01$, reflecting the fan-shaped pattern in the lower panels of Figure 1; there were no other significant effects, all $F<1$. The shape of the difficulty x prize interaction indicates that children's judgement follows the multiplicative pattern predicted by the formal EV model.

Individual Anovas revealed that 11 of the 5-year-olds showed a main effect of difficulty, seven of prize, two showed both main effects and one a significant interaction. Eleven of the 7-year-olds showed a main effect of difficulty, 12 of prize, six showed both main effects and three an interaction.

Despite few children showing a significant interaction, 10 5-year-olds and eight 7-year-olds showed the predicted fan shape in the data (with the game effect more than two points larger for the most desirable prize than the least desirable prize). Thus the individual analyses agreed with the impression from the group data that the structure of children's judgements corresponds to the predictions of the EV model.

Choices. When choosing which game to try first at the beginning of the performance task, 11 5- and 10 7-year-olds opted for the easiest game (large gate at 20 cm distance), the remainder were split between the two more challenging options. When choosing which game to play for a real sticker prize at the very end of the session, even more children, 12 5- and 14 7-year-olds opted for the easiest game, and none of the remainder chose the most difficult option (small gate at 60 cm). This shift towards the easiest game was significant (z=-2.06, p=.40, Wilcoxon, collapsed over age).

DISCUSSION

In this study, young children's difficulty judgements in a skill-dependent game reflected both objective task structure, and corresponded reasonably well to success probability. Moreover, the pattern of children's happiness judgements when games were paired with prizes agreed well with the formal EV model.

Judgements of Task Difficulty and Performance. In the first part of the study, 5- and 7-year-olds made realistic judgements of the difficulty of games with different gate-size/distance combinations prior to practical experience with the games. These findings extend those of Schneider et al. (1989) who demonstrated that 3- to 6-year-olds can make systematic predictions of success on physical tasks when given objective cues to difficulty.

Comparison of judgements with performance showed that children were realistic about the effects of distance, but not gate size, on task difficulty: They overestimated their success for the largest gate, and underestimated it for the smallest gate. Unrealistic optimism in children's predictions of performance has been reported when stimuli varied uni-dimensionally (e.g. Schneider, 1998; Stipek, 1984; Wellmann, 1985). In the present two-dimensional task the picture was more complicated, with one dimension assessed realistically, while the other was assigned too much weight.

It is not entirely clear how to best model task difficulty, which depends on the physical structure and on the child's action systems. A major factor is the precision of the aim needed, a function of the ratio of gate size to distance. This predicts a multiplicative pattern, but performance was additive. This could reflect the small number of trials, or distance might contribute significantly in a second way. For example, deviations from the initial trajectory due to surface irregularities increase with distance, or the more forceful push needed may make the aim harder to control.

Children's judgements, at any rate, may simply rely on the multi-purpose addition rule proposed by Anderson and Cuneo (1978), reflecting their recognition that both dimensions are relevant, without clear understanding of physical structure and situation-specific parameters. Further work on developmental changes in judgment, performance and understanding at older ages is desirable.

Why children focus too much on the role of gate size is unclear. Anecdotal evidence suggested that children believed they could push the marble harder to make up for it having to roll further, but they did not report compensation beliefs for small gates. Alternatively gate size may be more salient for children than distance, as the explicit goal was to roll the marble through the gate. Further work is necessary on this issue as well. But despite these open questions, the main point is that children's difficulty judgements were qualitatively consistent with physical task structure, and on the whole well calibrated with actual performance.

Task Difficulty and Expected Value. In the second part of the study, children evaluated games for M&M prizes. The resulting data pattern corresponded closely to the multiplicative predictions of the formal EV model. This extends previous work on children's EV concepts from games of chance to games of skill.

The multiplicative nature of children's EV judgements in probabilistic games is remarkable because in intuitive physics children make additive judgements for multiplicative concepts until around age 8. Multiplication may be more difficult when it involves a conjunction of two dimensions to form a third (e.g., length x width = area) than when one dimension merely weights another (e.g., probability x value = expected value) (see Schlottmann & Wilkening, 2010, in press). Such weighting effects appear to extend beyond the domain of formal probability.

Traditional theory saw probability understanding as late emerging. However, while children's computational accuracy increases as they grow older, good structural understanding at an intuitive level appears from pre-school age, prior to formal instruction (see review in Schlottmann & Wilkening, 2010, in press). This agrees with Fischbein's (1975) view that probability intuitions are adaptive in this uncertain world and thus likely to emerge early. In fact, Teglas, Girotto, Gonzalez and Bonatti (2007; also see Xu & Garcia, 2008) argue for probability understanding from infancy.

The present results provide a first demonstration that unorthodox, but natural probabilities may be co-opted into children's developing probability concepts. Differentiated judgements of task difficulty show that information about skill-related success probability is available to children. EV judgements incorporating these task difficulties go further to show that they use this information in probabilistic reasoning. Thus our results suggest one mechanism through which probability understanding might develop without experience with formal probability: It might be a corollary of skill learning. There may be yet other sources of a natural probability understanding.

Finally, children's choices also fit with EV and achievement motivation theory. At two points in the study, children chose which game to play: At the very end, most chose the highest EV option. On the initial performance trial, when no prize was at stake, children significantly more often chose a more difficult game, in line with the view that they, like adults, are often motivated to try tasks when they are uncertain whether they can succeed (Schneider et al., 1989). These choices underscore children's sound understanding: They were motivated to play games high in intrinsic motivation in the absence of a prize, but high in EV when playing for a prize.

Experience of personal success probability could be a precursor to probability understanding in formal, lottery-style situations. However, understanding of lottery-style probabilities appears from at least 4 years of age (Anderson & Schlottmann, 1991; Schlottmann & Christoforou, in prep.). It remains to be seen whether skill-related probability intuitions appears even earlier. Regardless, studies of skill-dependent and other unorthodox probabilities provide a promising new approach to the study of children's emerging probability understanding.

REFERENCES

Acredolo, C., O'Connor, J., Banks, L., & Horobin, K. (1989). Children's ability to make probability estimates: Skills revealed through application of Anderson's functional measurement methodology. *Child Development, 60*, 933-945.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anderson, N. H. (1982). *Methods of information integration theory*. New York: Academic Press.

Anderson, N. H. (1991). *Contributions to information integration theory* (Vol. I-III). Hillside, NJ: Erlbaum.

Anderson, N. H. (1996). *A functional theory of cognition*. Mahwah, NJ: Erlbaum.

Anderson, N. H., & Schlottmann, A. (1991). Developmental study of personal probability. In N. H. Anderson (Ed.), *Contributions to information integration theory: Vol III Developmental* (pp. 111-134). Hillsdale, NJ: Erlbaum.

Anderson, N. H., & Cuneo, D. O. (1978). The height + width rule in children's judgements of quantity. *Journal of Experimental Psychology: General, 107*, 335-378.

Fischbein, H. (1975). *The intuitive sources of probabilistic thinking in children*. Dordrecht, The Netherlands: Reidel.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906-911.

Hoemann, H. W., & Ross, B. M. (1982). Children's concepts of chance and probability. In C. J. BRainerd (Ed.), *Children's logical and mathematical cognition* (pp. 93-121). New York: Springer.

Nicholls, J. G. (1978). The development of concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development, 49*, 800-814.

Piaget, J., & Inhelder, B. (1958). *The growth of logical thinking in children and adolescents* (A. Parsons & S. Milgram, Trans.). New York: Basic Books.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children* (L. Leake, Jr., P. Burrell & H. Fishbein, Trans.). New York: Norton.

Schlottmann, A., & Wilkening, F. (2010, in press). Early developments in judgment and decision: Probability and expected value, serial processing, heuristics and biases.

In M. Dhami, A. Schlottman & M. Waldmann (Eds.), *Judgment and decision-making as a skill: Learning, Development, Evolution.* Cambridge: University Press.

Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *Child Development, 72*(1), 103-122.

Schlottmann, A., & Anderson, N. H. (1994). Children's judgements of expected value. *Developmental Psychology, 30*(1), 56-66.

Schlottmann, A., & Christoforou, M. (in prep.). Why are young children so good at expected value judgement?

Schlottmann, A., & Tring, J. (2005). How children reason about gains and losses: Framing effects in judgement and choice. *Swiss Journal of Psychology, 64*(3), 153-171.

Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science, 1*(2), 291-297.

Schneider, W., Hanne, K., & Lehmann, B. (1989). The development of children's achievement-related expectancies and subjective uncertainty. *Journal of Experimental Child Psychology, 47*, 160-174.

Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty*. Psychology Press.

Stipek, D. J. (1984). Young children's performance expectations: Logical analysis or wishful thinking? In J. Nicholls (Ed.), *Advances in motivation and achievement: The development of achievement motivation* (Vol. 3, pp. 33-56). Greenwich, CT: JAI.

Stipek, D. J., & MacIver, D. (1989). Developmental changes in children's assessment of intellectual competence. *Child Development, 60*(521-538).

Teglas, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences of the United States of America, 104*(48), 19156-19159.

Wellmann, H. M. (1985). A child's theory of mind: the development of Conceptions of cognition. In S. R. Yussen (Ed.), *The growth of reflection in children* (pp. 169-206). New York: Academic Press.

Wilkening, F., & Anderson, N. H. (1991). Representation and diagnosis of knowledge structures in developmental psychology. In N. H. Anderson (Ed.), *Contributions to information integration theory: Vol III Developmental* (pp. 45-80). Hillsdale, NJ: Erlbaum.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America, 105*(13), 5012-5015.

(Manuscript received: XXX; accepted: XXX)