

preprint of: Judgment and decision making in young children. Schlottmann A, Wilkening F (2011)

In Judgment and Decision Making as a Skill. Editors: Dhani M, Schlottmann A, Waldmann M. 55-83. Cambridge University Press 31 Dec 2011

Judgement and Decision Making in Young Children:

Probability, Expected Value, Belief Updating, Heuristics and Biases.

Anne Schlottmann

University College London

Friedrich Wilkening

University of Zürich

(2011, in press). In M. K. Dhani, A. Schlottmann, & M. Waldmann (Eds.), *Judgement and decision making as a skill: Learning, development, and evolution*. Cambridge, UK: Cambridge University Press.

Understanding probability and utility is core to Judgement and Decision Making (JDM). Under a Piagetian view these are formal operational concepts, achieved in adolescence. Most developmental work on JDM indeed involves adolescents (Bruine de Bruin, this volume; Jacobs and Klaczynski, 2005). The Piagetian approach also fits with a view of JDM as a skill: If we take adults as somewhat under-skilled, then one reason might be lack of practice, prevented during childhood by lack of basic cognitive abilities (Furby and Beth-Marom, 1992). Thus, when a child contemplates a run for the forbidden sweetie jar, on the standard view, rational considerations of the probabilities of success versus discovery and the utilities of a sweet snack versus parental dismay are not possible.

A standard Piagetian approach has, however, been largely abandoned in cognitive development research, with substantially earlier abilities than predicted by Piaget apparent in practically all cognitive domains, and we will not defend it here. Instead, we review a growing body of research, mostly carried out since earlier reviews (Hoemann and Ross, 1982; Reyna and Brainerd, 1994), showing quite the opposite: that children even before entering school have functional probability understanding. Thus, adults' lack of skill cannot be blamed on delayed conceptual development. Instead, reasons might lie in the complexity, structure, and meta-cognitive demands of prototypical JDM tasks used with adults and adolescents.

Our review begins with the issue of whether children distinguish deterministic (certain) from probabilistic (uncertain) events at all. We find that they can, then show that they can also make reasonable judgments of probabilities/utilities in these uncertain events – at an intuitive level. This early intuitive competence contrasts with the well-known difficulties even adults experience in situations requiring explicit application of probability concepts or computation. We then address how children might learn about specific probabilities, through event sampling or causal-logical analysis of the situation. This research largely arose within a cognitive-developmental framework, but we go on to discuss work derived from a standard JDM perspective, on heuristics and biases, which operate from early childhood as well. All in

all, it will become obvious that there is no simple developmental trajectory from less normative to more normative processing (or the reverse), so the question looming at the end is how to integrate evidence on simultaneous strengths and weaknesses from childhood.

Although the details are far from clear, some form of multiple systems/process theory seems inevitable to accommodate the widely differing levels of understanding found in children (and adults).

A crucial concept for our review is that of intuition. In cognitive contexts, intuition is usually defined negatively, by contrast with more explicit, conscious, or computational forms of reasoning (Hammond, 1996). Intuition is typically considered not only inarticulate, but also undifferentiated, which would severely limit its use. Thus Piaget and Inhelder (1951/1975) dismissed children's intuitions as immature and dysfunctional, to be replaced by late emerging, verbalisable concepts. Fischbein (1975), in contrast, argued that even young children have functional probabilistic intuitions that are adaptive in this uncertain world. Both Piaget and Fischbein's methods, however, were too qualitative to provide convincing evidence on the role of intuition. Both largely relied on isolated nonverbal responses (see below) that were interpreted based on how children talked about them. To some extent, Fischbein simply re-interpreted Piaget's data -- highlighting their ambiguity.

The more recent work we review largely disagrees with Piaget, but agrees with Fischbein. In contrast to both approaches, however, it uses tight experimental methodologies that do not rely on verbalisations – this is fortunate because children tend to know much more than they can say, again a lesson quite clear from other domains of cognitive development. These new data demonstrate that probability intuitions can be highly structured. Moreover, this structure is often close to that of formal probability models, making it difficult, if not unwarranted, to discount children's intuitive knowledge as “non-probabilistic”.

Can Children Distinguish Deterministic and Probabilistic Events?

Without the ability to distinguish deterministic situations from chance, understanding of probability is impossible. The first part of Piaget and Inhelder's seminal book (1951/1975), accordingly, was devoted to this issue. Children predicted outcomes in devices producing random mixtures of elements or different forms of distributions. The conclusion was that preschoolers – because of lack of causal understanding – fail to distinguish the necessary outcomes of deterministic devices from the possible outcomes of chance devices, and that school children – because of limitations in logical thought – still have trouble recognizing how the actual outcome distribution also depends on how often observations are repeated.

Few studies have re-considered this basic question. One might surmise, however, that Piaget's tasks, as in many other domains, drew on more than just the basic deterministic-probabilistic distinction. High linguistic, memory, and meta-cognitive demands in his tasks may add to their difficulty. Indeed, Kuzmak and Gelman (1986) found that when asked “do you know which colour will come out, yes or no?”, most four-year-olds appropriately distinguished a lottery-like spinning-marble cage and a clear tube containing a row of marbles ejected in a visibly pre-determined order. Appropriate explanations appeared from five years.

In tasks with higher meta-cognitive demands, children may not appreciate the implications of the deterministic-probabilistic distinction until later. Rapp and Wilkening (2005) asked children to find out if a deterministic ball-dropping and a probabilistic spinner device worked properly; children could keep a record of outcomes for one or the other device. From age ten, children understood that a record is more useful for the probabilistic device because it takes more observations to check its operation. Children's appreciation of the distinction between necessity and uncertainty thus depends on task complexity, but is recognized at a basic level from four years, in contrast to Piaget's assertions.

Can Children Understand Probability?

Having seen that children understand that probabilistic events differ from deterministic events, one might ask what they know about the structure of probabilistic events. Piaget and Inhelder addressed this in the influential second part of their book, which mostly deals with children's understanding of random draws from mixtures with variable proportions of desirable and undesirable elements. Two tasks were developed: In one, children predicted the outcome of a draw from a set of marbles with different colours. In the other, children saw two sets of marbles, each containing some with a winning and some with a non-winning colour, and chose the better set for winning in a blind draw. Both tasks have been used extensively since, in somewhat different literatures. The two-sample choice task has been reviewed before (Hoemann and Ross, 1982; Reyna and Brainerd, 1994), while the prediction task appears mainly in the literature on probability matching (Fischbein, 1975).

Children do well on Piaget's two-sample choice task, but only if both samples contain the same total number of elements. It has been argued that children answer correctly because they use a non-probabilistic magnitude estimation strategy: They compare the number of targets, without relating targets to non-targets. Only if the two sets differ in total number, children must use a probability strategy to generate the correct choice, that is, they must compare the targets to the total number in each set to find the set with the higher proportion (Figure 1). Under this condition performance drops substantially. Half of Siegler's (1981) eight-year-olds, for example, still attended to the targets only, while the other half already employed the appropriate probability strategy. Compared to choice tasks involving proportional concepts from other domains, performance on the probability task lagged behind initially, but improved to a more advanced level.

Figure 1

Performance factors can play a confounding role, but are not always crucial. Non-verbal tasks, for instance, are typically somewhat easier. However, giving children feedback

on the actual outcome of a draw, or a reward for correct guesses, does not help much (Hoemann and Ross, 1982; Reyna and Brainerd, 1994). Over the years, the theoretical conceptualization of children's difficulties shifted implicitly from Piaget's emphasis on their logical limitations to an information processing view stressing computational deficiencies. Either way, from the literature on probability choice tasks one would conclude that understanding probability is beyond young children's capabilities (for an exception, see Huber and Huber, 1987).

When tasks involve graded judgements of probabilities, in contrast, children do rather well (e.g., Acredolo, O'Connor, Banks and Horobin, 1989; Anderson and Schlottmann, 1991, Wilkening and Anderson, 1991). Because young children lack the number skills to make numerical probability estimates, graphical rating scales are employed in these tasks. Children indicate how easy it is to win a game by pointing to one of a series of sticks, with longer sticks for easier games, or show how happy they would be to play a game by pointing to one of a series of faces, ranging in expression from sad to happy. These studies use the framework of information integration theory, also known as functional measurement (Anderson, 1981, 1982, 1991, 1996). Children encounter a large number of games, say, randomly drawing a blue winner marble from sets of blue and black marbles, with factorial variations of numbers of winners and losers. With such designs, the underlying strategy can be directly diagnosed from the data pattern, without need for verbalizations. The response demands of rating tasks might seem higher than in choice tasks, but children's performance is reliably more advanced. To illustrate, Figure 2 shows results from Anderson and Schlottmann (1991).

Figure 2

In Figure 2, in contrast to expectations from the binary choice literature, probability judgements at all ages systematically increase with the number of winners (horizontal) and losers (curve parameter). This pattern appeared at the group and individual level. Most strikingly, the data for all ages show a slanted barrel shape (i.e., curves converge at both

extremes), as predicted by the normative probability model. However, the model predicts less barrelling than found empirically. The decrease in barrelling with age in Figure 2 thus indicates development toward the normative law. Exaggerated barrelling, as found for the younger children, occurs when winners are weighted too heavily; this is the only feature of the results reminiscent of the traditional Piagetian view. Instead of focusing only on targets, however, children in this study clearly integrated both targets and non-targets.

The data leave open whether children used the probability ratio rule (number of winners divided by total, i.e., winners plus losers), or a simpler serial rule for this integration. Under the latter, they first respond based on whether winners or losers are more frequent, then make an (insufficient) adjustment in the other direction for the less frequent cue. The serial strategy may be formalized as an averaging rule (also suggested for expected value judgement later). Regardless of this issue, Figure 2 demonstrates a far higher level of competence from pre-school age than expected traditionally.

A similar conclusion arises from Acredolo et al.'s study (1989) with seven- to eleven-year-olds that varied the number of winners and total number of outcomes. In this case, the probability ratio rule predicts a diverging fan instead of barrelling, with a Winner x Total interaction, while the serial subtractive-averaging rule (similar to a winners minus losers rule) predicts a parallel pattern, without interaction. The group data showed clear fanning, but data for individual children often lacked a significant interaction, so both rules may have been used.

Acredolo's study also showed that the strong overweighting of winners, apparent in one of his tasks and in Figure 2 above, is not an inevitable developmental outcome. A second task managed to produce more equal cue weighting, downplaying the role of the targets by making children judge how likely it was that a jumping bug would land on a flower rather than a spider. Thus, target dominance may simply be a function of the salience of a win-lose interpretation of the task.

These studies leave no doubt that children from age four understand the structure of the probability concept. Accuracy of individual judgements, in contrast, or the question of how children might learn about specific probabilities, is not of major interest in this line of work. Information integration theory emphasizes the subjective representation of the external world. Judgements are seen as subjective estimates, and following longstanding psychophysical tradition, it is taken as read that subjective estimates and objective values do not normally coincide. The approach could be extended to investigate how objective values and subjective estimates are related, that is, to derive psychophysical probability functions for young children and study the accuracy of their judgements, but this has not been done yet (though there is work with adolescents, see Bruine de Bruin, this volume). The question of how children learn about specific probabilities is taken up in a later section.

Can Children Understand Expected Value?

In everyday life, probabilities typically appear in the context of goal attainment. Children's goals – like adults' – vary in desirability and attainability, as reflected in the concept of Expected Value (EV). It seems natural, therefore, to extend the inquiry in this direction. In fact, probability judgement is typically made meaningful to children by giving outcomes the values of winning or losing a game (i.e., 1 and -1). Such tasks can be converted into explicit EV tasks by associating the outcomes with variable prizes (Figure 3) and asking children how happy they (or a puppet) would be with each game.

Figure 3

Normatively, EV of an outcome is defined as its value weighted by its probability, or as the product of probability and value, while overall EV of an event with two or more outcomes is the sum of the component EVs. Thus, two operations are required: multiplication for the within-outcome integration and addition for the across-outcome integration.

Within-outcome EV: Precocious multiplication?

If children multiply probability and value, the data pattern should be a divergent fan, illustrated in Figure 4. The diverging fans for the top and bottom pair of curves show the multiplication of probability and prize on the blue outcome, with two different yellow outcomes in the 3 probability x 2 blue value x 2 yellow value factorial design. One fan slopes upwards, the other downwards, because the blue outcome is either the larger or smaller of the alternatives, and EV increases or decreases with the probability of winning it.

Figure 4

The early emergence of multiplication, apparent – with some initial ambiguity – since Anderson's (1980) study, was surprising. In other multiplicative tasks, children younger than eight years show additive patterns (parallelism in the factorial graph), interpreted as the application of a multipurpose rule that children use when they know that two variables are relevant, but not exactly how they are related. This appears reliably for area judgement (e.g., Anderson and Cuneo, 1978; Wilkening, 1979), time and velocity judgement (Wilkening, 1981), and numerosity judgement (Cuneo, 1982). Fanning, however, appears equally reliably in EV tasks (Anderson, 1980; Schlottmann and Anderson, 1994; Schlottmann, 2001; Schlottmann and Tring, 2005; Bayless and Schlottmann, in press), including in a study with four-year-olds (Schlottmann and Christoforou, in preparation), well before instruction with probabilities in school.

The likely interpretation is that there are two kinds of multiplication concepts. In conjunctive concepts, as typical in intuitive physics, two dimensions combine to form a qualitatively new third dimension. Thus, length multiplied by width of a rectangle yields its area, and distance divided by time yields velocity. In weighting concepts, in contrast, one dimension merely modifies the other, without a qualitatively new resultant. Thus, the value of a goal is modified by its probability, but the outcome is still a value. Less need for conceptual

abstraction in weighting tasks may allow for earlier application of the multiplicative rule. Schlottmann and Christoferou (in preparation) tested this possibility by giving four- to ten-year-olds a conjunctive area task and two weighting tasks, an EV task and a novel source-message task. As predicted, the younger children multiplied in both weighting tasks, but added in the area task, supporting the view that cognitive complexity may differ for different kinds of multiplication concepts in formally isomorphic situations. Thus, the multiplicative nature of children's EV judgements appears clearly established.

A second point is that children's probability understanding is remarkably abstract. Even the youngest children's judgements co-vary with probability, not the physical quantity used to represent it, in contrast to the magnitude estimation hypothesis (Hoemann and Ross, 1982). Children clearly understand that probability is a relative, not an absolute quantity. They also interpret the physical probability cues selectively, adapting their meaning to the value structure of the situation. In Figure 4, for instance, judgements of the same physical game reverse, depending on which outcome carries the larger prize. Thus, children do not just associate physical targets and responses, but clearly know that the meaning of a physical cue can differ, depending on how desirable and probable the outcome is.

Across-outcome EV: Subadditivity caused by a serial strategy?

So far we have highlighted children's remarkable intuitive competence. However, while it seems clear that the within-outcome integration is multiplicative, as normatively required, there is doubt about the additivity of the across-outcome integration, even though addition seems mathematically simpler than multiplication.

Adults (Shanteau, 1975) and primary-school-aged children as well (Schlottmann, 2001) often show subadditivities in games with two probabilistic outcomes, as in the task of Figure 3, in which both blue and yellow outcomes could carry prizes. Normatively, overall EV is the sum of the component EVs, implying that the effect of one prize is independent of

the other. Only a few individuals in Schlottmann (2001) showed this normative independence, although the individual differences averaged out in the group data: For many children, the same prize on blue had substantially greater effect when it was the larger of the two prizes. Another, smaller group showed the opposite pattern, with the blue prize having greater effect when it was smaller.

One interpretation of these subadditivities is that children use a serial strategy, similar to how adults' judge single outcome events (Anderson, 1980; Lopes and Ekberg, 1980). In this view, children first locate the two outcome values on the response dimension, then fractionate the distance between them in proportion to the probabilities. "Risk-seeking" children may attend to the risky but larger value outcome first, with insufficient adjustment for the smaller value. "Risk-averse" children, in contrast, may start with the smaller value, sure outcome, with insufficient adjustment for the larger value. This interpretation is supported by reaction time data in adults, but such data are still lacking in children.

Nevertheless, this explanation of subadditivity seems more plausible than a more technical one: that it might be an artefact of the rating method, resulting from non-linear use of the response scale. This is logically possible, but whenever it has been tested, this was ruled out empirically (Anderson and Wilkening, 1991; Schlottmann, 2000; Schlottmann and Anderson, 1994; Surber, 1984).

Across-outcome EV: Averaging versus adding

The normative across-outcome EV model is additive, but with alternative, dependent probabilities this is structurally equivalent to averaging (i.e., the two differ only in division by a constant). The two rules become distinguishable, if outcomes have independent probabilities. If, for instance, a lottery with high probability of a prize is compared to a compound lottery with the same probability to win this prize, plus an additional chance to win

a second prize, then under the normative addition model, the compound lottery has higher worth. Children up to eleven years at least, nevertheless average, as shown in Figure 5.

Figure 5

The solid lines in Figure 5 show children's EV judgements for pairs of games with independently varied probability of winning a skipping rope or marble prize. The near-parallel curve pattern agrees with both adding and averaging. The dashed line is for judgements of individual games to win only the rope. Under the adding model, EV for these one-prize games lies below that for two-prize games, as for the adults on the right. For the children, however, the dashed curve crosses over, indicative of averaging. Consider the rightmost points: When a high probability for the rope combines with a medium probability for the marbles, this lower component worth is averaged in, pulling the judgement down – even though the same medium probability pulls the judgement up when combined with a low probability for the rope. This crossover rules out adding and supports averaging. Similar opposite direction effects have appeared in non-probabilistic situations for adults (Gaeth, Levin, Chakraborty and Levin, 1990; Hsee, 1998).

Children (even infants) have no problem with understanding addition as an increase in quantity (Hughes, 1986; Wynn, 1992). However, averaging is wide-spread in everyday value judgement in adults (Anderson, 1981) and children (Butzin and Anderson, 1974) and may appear in EV judgement as a non-normative overgeneralization from everyday life. In the compound spinner task, which involved the explicit instruction that “two is better than one”, adults replaced averaging by adding, but averaging processes might still contribute to EV judgement in less explicit situations. This finding illustrates how biases in adult judgements might appear as vestiges of developmentally earlier processes.

Origins of the probability-value link

The claim that EV is a basic intuition operative prior to formal instruction with probabilities and possibly linked developmentally to other forms of value judgement seems well supported by the above data. However, this raises the question of how children know that probability modifies value in the first place. This insight might not be obvious from preschoolers' limited experience with games of chance. We believe, however, that children have other, more transparent opportunities to learn about the role of uncertainty in their everyday lives, in guises that do not involve random draws. Uncertainty about an event may occur, for instance, because communication is unreliable. Uncertainty may also occur in skill-dependent tasks, where success depends on ability, effort, objective task difficulty, and so forth. Skill-related uncertainty is highly salient for beginning learners experiencing repeatedly that goal achievement is not guaranteed, and children from three years can reliably relate their chances of success to the difficulty level of skill tasks (Schneider, Hanne and Lehmann, 1989). Thus, children may learn about the implications of uncertainty in scenarios with little resemblance to lottery-type JDM tasks.

If skill-dependent uncertainty is linked to children's understanding of probability from early on, then they should be able to easily make EV judgements in skill-related tasks. To explore this, we showed children a shoot-the-marble-through-the-gate game, varying gate size and distance from the start line to create different difficulty levels (Bayless and Schlottmann, in press). Even five-year-olds' success judgements varied with both task difficulty factors, prior to practical experience with the game. Importantly, children gave multiplicative EV judgements when prizes were attached to a selection of games of variable difficulty. This supports the argument that children evaluated unorthodox, but natural probabilities here. Thus, EV is a basic concept in an uncertain world, that children may learn about the uncertainty of their goals from early on, and that rarely considered, but frequently encountered sources of uncertainty information may be co-opted into children's reasoning.

Judgement versus choice

Weaker performance in choice than judgement appears not only in developmental studies of probability, but also of EV (Levin, Weller, Pederson and Harshman, 2007b; Harbaugh, Krause and Vesterlund, 2002). Choice-judgement discrepancies are equally common in JDM research with adults (Lichtenstein and Slovic, 1971; Mellers, Chang, Birnbaum and Ordonez, 1992; Payne, 1982), and in other domains with children (e.g., Wilkening and Anderson, 1991).

That children's judgements typically reveal more advanced concepts than their choices may reflect that choice requires a more complex, comparative stimulus field. Siegler (1981) suggested that choice tasks are solved through sequential, dimension-by-dimension comparison of the two options, a highly analytic approach. Judgement tasks, generally presenting one stimulus at a time, seem to elicit more intuitive, approximate estimation strategies. The difference lies in stimulus structure, not response mode per se. This follows from studies using continuous adjustment responses in which children equalize two choice options (Falk and Wilkening, 1998; Levin, Wilkening and Dembo, 1984; Wilkening and Anderson, 1991). This ability to equalize options not only lags behind performance in single stimulus judgement tasks, but also behind binary choice tasks (and behind verbal understanding), presumably because in addition to the comparison elicited by the choice structure, the adjustment format suggests a very fine-graded, deliberate approach (Falk and Wilkening, 1998). Different probability task formats thus can elicit widely different levels of performance.

Figner and Schaub (2009) report an apparent exception. Children played a game in which a cat chased a mouse that advanced towards safety if the winning side of a dice was rolled. Trials varied in how many steps the mouse was away from safety ("aspiration level") and probability of winning. On each trial, children chose between rolling a dice of known

probability and drawing randomly from a bag containing dice with 0 to 6 winning sides. Five- to ten-year-olds chose advantageously, i.e., when the known probability was low and the mouse far from safety, it might be better to draw from the bag. Children also judged how happy the mouse would be with each option, taking both factors into account for the single dice, with unreliable ratings for the more complex bag option. On the whole, choice and judgement were at a similar level here, but note that the choice options were dissimilar in structure, so did not suggest a simple, short-cut comparison. Instead, judgement and choice might correspond, because children could focus on the better understood, single dice in both formats, accepting this option when the mouse was happy with it, rejecting it otherwise. This study highlights once more the importance of stimulus structure over response format.

Performance has many other determinants, of course, such as whether the situation involves gains or losses. For adults, reasoning about losses seems more demanding, involving anticipated regret (Connolly and Zeelenberg, 2002) and induced conflict (Lopes, 1987). In line with this, children make more advanced EV choices (Levin et al., 2007b) and judgements (Schlottmann and Tring, 2005) for gains.

Overall, it makes little sense to elevate any probability task over another on principle. Nevertheless, when assessing children's competence is a primary goal, intuitive judgement tasks using single-stimulus presentations and continuous scales may be most sensitive. Again, these intuitive judgements cannot be discounted as non-probabilistic or unstructured. They may well serve to support explicit understandings achieved later in development, in agreement with Fischbein's claims. This perspective, emphasizing the range of task requirements and the multiplicity of children's learning opportunities, makes it less controversial to grant even young children considerable probability understanding.

The research reviewed so far largely focused on how children understand the structure of basic JDM concepts, showing both substantial strengths and some limitations from early on. However, at least two other questions of much interest in JDM have been addressed in

work with young children: One is how children's intuitive competence fits with the literature on heuristics and biases. The other is how children might learn about the probabilities of particular events. These questions are addressed below.

Learning about Probability

The philosophical literature distinguishes multiple senses of probability, which seem to entail different ways of learning. Empirical or frequentist probability is based on repeated experience with events. In children, this should show up as sensitivity to the role of event frequencies. Classical or theoretical probability, in contrast, is based on specification of the set of possible outcomes. In a spinner game, for instance, the outcome has to be one of several spinner segments. Children's appreciation of how these physical features relate to outcome probability might either be based on analysis of the causal-logical structure of the situation or on event frequencies. The debate between extensional (Johnson-Laird, Legrenzi, Girotto, Sonino-Legrenzi and Caverni, 1999) and frequentist (Cosmides and Tooby, 1995; Gigerenzer and Hoffrage, 1995) accounts of probability understanding partly reflects this distinction. Furthermore, there is subjective probability, seen as degree of belief in or expectation about an outcome. Children might achieve graded beliefs from experienced event frequencies, causal-logical analysis, and/or a generalisation from experience with similar situations. This subjective approach to probability would seem most encompassing: it applies to single-case and non-lottery scenarios and allows for multiple ways of learning.

Most studies reviewed above involved classical probability, with spatial or numerical summary representations of a set of possibilities, such that the temporal frequency of outcomes is proportional to, say, size of winning space or number of winning marbles. This approach makes sense as it reduces memory demands. There is the implicit assumption, however, that children's experience with such devices plays a major role in their understanding, that is, theoretical and empirical probability are closely aligned. Even if

children have little direct experience with lotteries, understanding might be supported by the more solid empirical knowledge they have achieved in other probabilistic situations, as discussed earlier. In this view, event sampling and frequency processing is a main route to learning about probability, even if there are other ways as well (a divergent opinion appears below).

Undeniably, even infants are sensitive to event-frequency probability. This appears in the literature on socio-emotional development, as sensitivity to contingency in interaction patterns (Gergely and Watson, 1999). There are also striking examples of infants' capacities in statistical learning of linguistic (Saffran, Aslin and Newport, 1996) and visual patterns (Fiser and Aslin, 2002) and in matching sample and population proportions (Xu and Garcia, 2008). While this probability sensitivity is entirely implicit in visual or motor behaviour, it may become more explicit over the years.

Probability matching and event sampling

Slightly older children's intuitions of relative frequency appear in probability matching behaviour (see Fischbein, 1975; Stevenson, 1970), which emerges from three years. It is, however, inexact at first and improves considerably with age. The standard task involves repeated predictions about which of two randomly drawn events will occur next. Correct predictions are rewarded, so one should always predict the more frequent event. Even adults, however, typically behave sub-optimally, with predictions proportional to experienced event frequencies (but see Shanks, Tunney and McCarthy, 2002). In Boyer (2007), even five- and six-year-olds performed above chance level in identifying the more frequent event, but they considerably lagged behind older children in task awareness.

This gap between awareness and performance also appears in an event-sampling task from the neuropsychology literature (Crone and Van der Molen, 2004; Garron and Moore, 2004): In the Iowa Gambling task, children repeatedly choose a card from two or more decks.

Some are profitable in the short, but disadvantageous in the long run, with large wins on each card offset by occasional larger losses. Other decks are less profitable initially, with smaller gains on each card, but advantageous in the long run, with only small losses occasionally. With a simplified task, four-year-olds eventually chose the advantageous set, while 3-year-olds chose the disadvantageous one (Kerr and Zelazo, 2004). The complex nature of this task, however, leaves open whether this developmental change involves probability understanding or predominantly memory and executive functions.

Another unresolved issue is that probability matching appears in the mean responses over many trials. Any particular response, however, depends on the pattern of preceding trials, with strong order effects. Children, for instance, tend to predict the more frequent colour initially, then often alternate responses, which may reflect retrieval failures and working memory limitations (Brainerd, 1981). To understand the sequential process of belief revision, a model would seem necessary to decompose the overt response into the immediate reactions to recent events, and the stable-state reaction to the long-term probability.

Belief revision in event sampling

Although such a model-based approach has not been applied to children's choices, it has been used to study directly how children estimate event frequencies in a sampling-and-updating task (Schlottmann and Anderson, 1995, 2007). Because children seem to live largely in the here-and-now, it was unclear initially whether past information would influence their judgements at all. The findings, however, converged with the serial-strategy interpretation of the EV studies in showing that even five-year-olds can form integrated beliefs with short- and long-term components.

Children heard about turtles catching gold and red starfish for a living and judged (on a scale from all gold to all red) their proportion in the catch, based on up to five starfish revealed sequentially (one sample visible at a time). Children updated their belief about the

overall proportion after each sample, judging several sequences with different sample composition in this and a structurally identical task with more social content. In both, even five-year-olds made more extreme judgements for more extreme sample proportions (see Figure 6), consistent with Jacobs and Narloch's (2001) finding that elementary school children can consider both sample size and variability when estimating population size based on simultaneously presented samples.

Figure 6

At the same time, children's judgements showed strong non-normative recency, that is, larger effects of more recent samples: In the left panel of Figure 6, for instance, judgements at Position 2 are higher when identical sample proportions were presented in --+ than in +- order. A serial averaging model, under which each informer is represented by an information value and an importance weight, previously established for adults (Anderson, 1981; Hogarth and Einhorn, 1992) fit the data. This permitted decomposition of the responses into the contributions of early and late informers, showing that recency effects were largely short-term and that early informers had stable, long-term effects on the response from at least six years. Thus, children can meaningfully estimate relative frequencies, although this is masked by strong order effects in judgement, not memory (Hastie and Parks, 1986; Anderson, 1981).

In the above study, continuous updating was built into the procedure, because children responded by moving a marker that also provided a visible reminder of the previous response. In a follow-up study (Schlottmann and Anderson, 2007) this marker was removed, yet children still employed running adjustment. Previously established as an effective strategy in adults' serial judgement, running adjustment operates from childhood. It should therefore be possible to study children's judgements in more standard JDM belief revision tasks.

So far, only two studies have attempted this. Girotto and Gonzalez (2008) asked children which colour would be drawn at random from a bag containing, say, four black

circles, one black square, and three white squares. Children saw the bag contents, then had the initial hypothesis (dis)confirmed verbally, e.g., when learning that the target is square this changes the odds from 5:3 for black to 3:1 for white. Performance was above chance from five years, improving further until nine years. In line with the work above, therefore, children clearly responded to the most recent informer. This task, however, does not require any integration of earlier and later information, because the later information simply overrules the earlier one. In contrast, Zhu and Gigerenzer (2005) required children to integrate base rate and test data in Bayesian problems. From nine years, children could solve problems couched in natural frequency, but not probability form. The problems used written summaries, so event sampling was not involved. Clearly, further investigations of how children update probabilistic beliefs are needed.

In summary, children can sensibly adjust their beliefs. They neither fixate on initial information, nor live in the perpetual present. They are not only sensitive to event frequency in mean response rates, but can also make direct estimates of relative frequency, even from a series of observations. Much work is still needed to learn how they apply these skills in JDM, but the sparse existing data show that children can learn about probabilities from various forms of frequency information.

Theoretical probability intuitions in infants?

Teglas, Girotto, Gonzalez and Bonatti (2007) and Girotto and Gonzalez (2008) argue nevertheless that early probability understanding does not involve experienced event frequency, but rather a primary intuition of theoretical probability represented as a set of possibilities. Twelve-month-olds saw an animated lottery with three bouncing balls of one colour and one ball of another colour, followed by a ball of either the unique or more frequent colour appearing below the opening. Infants looked longer at the improbable than probable result, without habituation or experience with this device.

Alternatively, however, infants might simply track the more salient unique ball for longer. To test this, a control device had a divider across the middle, such that the three balls bouncing above could not reach the opening, while the uniquely-coloured ball below could. This time infants looked longer when, impossibly, one of the top balls appeared below. Infants know that objects cannot go through other objects (Spelke, 1992), so it is unsurprising that they understood the divider's implications, and because control infants did not simply track the unique ball, the authors argue that experimental infants did not do this either. However, that infants refrain from a salience approach with a deeper understanding of the situation, as in the control, does not imply they do the same in the experimental condition, where the existence of a deeper understanding is unclear. Accordingly, the conclusion that infants understood the outcomes as probable and improbable would seem somewhat premature.

Teglas et al. (2007) also studied pre-school children's reactions when the set of possibilities conflicted with the experienced event frequencies. A ball bounced in a box with three exit holes on one side and one hole on the other. Children had to press a button as soon as it exited on either side. Three- and five-year-olds initially reacted faster to three-hole exits. Outcomes, however, were rigged such that the ball mostly exited on the one-hole side. Five-year-olds, but not three-year-olds eventually became faster at detecting this. When asked where the ball would exit, five-year-olds, but not three-year-olds, stated the probable three-hole side initially; neither age learned to say that the ball had mostly exited elsewhere. This lack of effect for the frequency manipulation was taken to imply that early probability intuitions are independent of observed frequencies. However, this conclusion may also go too far: Children's expectation of where the ball should exit may have been a generalisation from links between observed frequencies and causal structure in related situations, with the experimental evidence simply too weak to override over-learned prior knowledge.

While this research is not conclusive, it asks provocative and timely questions about where children's early appreciation of classical probability might come from. It may not seem overly plausible to endow children with rational intuitions about lottery machines, which are not a feature of the natural world, but a feat of engineering. However, foraging animals are also sensitive to the spatial probability distribution of resources, even though most work on animals' sensitivity to resource distributions employs event-sampling paradigms, with frequentist assumptions (Real, 1991).¹

Heuristics and Biases in Childhood

The final issue we address is how children's early competence with probability, expected value, and serial processing relates to development of the heuristics and biases so prominent in adult JDM. This question is difficult to answer, however, in part because experimental designs that elicit heuristics and biases do not generally allow for investigation of the underlying processes. In any event, for most heuristics/biases only one or two developmental studies are published. Only the framing phenomenon has attracted more attention.

Framing studies largely agree that children are more risk-seeking for losses than gains, just like adults (Harbaugh et al., 2002; Levin and Hart, 2003; Levin, Hart, Weller and Harshman, 2007a; Schlottmann and Tring, 2005; but see Reyna and Ellis, 1994; Levin et al., 2007b). While susceptibility to frame differences does not change much developmentally, children seem somewhat more risk-seeking than adults in equivalent tasks (Harbaugh et al., 2002; Levin and Hart, 2003; Levin et al. 2007a; Reyna and Ellis, 1994; but see Schlottmann and Tring 2005), consistent with popular opinion. Children's risk attitudes – like adults' – depend on the level of risk, but the direction of this effect varies with task specifics

¹ Whether or not children, like animals, are initially sensitive to the probabilities of resources, children soon develop probability intuitions at a more abstract level, as shown earlier.

(Harbaugh et al., 2002; Schlottmann and Tring, 2005). Schlottmann and Tring (2005) also showed that frame differences were strong in standard choice tasks, but minimal (albeit in the same direction) in EV judgements for each choice option, consistent with findings, discussed earlier, of more advanced judgement than choice strategies.

Additionally, children's risk seeking is predicted by personality factors such as their impulsivity, as well as by parental risk attitude (Levin and Hart, 2003; Levin et al., 2007a). When children and parents were re-tested after three years, risk attitudes for both groups were correlated across time, and children's risk attitude was still predicted by their initial personality measure, but not anymore by parental attitude. These data show essential continuity between child and adult risk attitudes, with intriguing suggestions about their ontogenetic formation.

Looking more widely across developmental research on all biases/heuristics, two main findings appear: First, with three exceptions (Arkes and Ayton, 1999; Reyna and Ellis, 1994; Levin et al, 2007b), the target bias/heuristic operated from the youngest age studied. While most studies were limited to elementary school age, the conjunction fallacy (a tendency to consider a more specific event more likely than a more general event), notably, has now been reported from age four (Fisk, Bury and Holden, 2006). Non-normative strategies thus emerge as early as the normative strategies reviewed earlier.

The second general finding is that the direction of development varies. Some studies found increases with age, in framing effects (Reyna and Ellis, 1994, Levin et al., 2007b), use of the representativeness heuristic (Jacobs and Potenza, 1991; Davidson, 1995), the conjunction fallacy (Davidson, 1995), and sunk cost decisions (Arkes and Ayton, 1999). Other studies found no change, in framing effects (Schlottmann and Tring, 2005; Levin and Hart, 2003), use of the representativeness heuristic (Agnoli, 1991), the conjunction fallacy (Fisk et al., 2006), and anchoring effects (Smith, 1999). Yet others found a decrease, in use of the availability heuristic (Davies and White, 1994), sunk cost decisions (Klaczynski and

Cottrell, 2004), and averaging errors (Schlottmann, 2000). These findings do not support the traditional view of a general decrease in non-normative heuristics/biases as children's logical and computational skills grow. These findings also disagree with a fuzzy-trace view of a general increase in heuristics/biases, as children move from quantitative-verbatim towards gist-based processing (Reyna, 2006; Reyna and Ellis, 1994). Developmental trajectories clearly differ depending on the bias/heuristic and the specific task at issue.

At this point, it is difficult to predict which developmental trend should appear under which conditions (Jacobs and Klaczynski, 2002). Little is known about the underlying processes. Another complicating factor is that increases in some bias (or normative behaviour) might reflect nothing more than growth in ancillary skills required for the task. This is particularly problematic in JDM, where child tasks often derive from complex adult tasks. These are difficult to strip of extraneous demands, which might overload children. A high level of "not-on-task" behaviour might indicate unspecific comprehension/processing difficulties; more specific problems can also sometimes be identified. For example, decreasing use of the availability heuristic might be a memory effect (Davies and White, 1994), and increases in anchoring-and-adjustment might reflect computational improvements (Smith, 1999). Such findings are significant because comparable adult tasks may also draw on multiple processes. As Jacobs and Klaczynski (2002) point out, concern with demonstrating the origins of some bias or capacity under highly simplified conditions should not detract from the social, motivational, affective – and cognitive – conditions under which it comes into play.

Summary, Open Questions and Conclusions

The work we reviewed showed early competence in simple probability and expected value tasks. On the whole, there was little evidence of developmental change or fundamental deficits in how children grasp the structure of these concepts. Furthermore, some ability to

dynamically extract probabilities from ongoing events already appears from infancy, although the use of sequentially sampled probabilities in JDM tasks improves substantially with age. Good performance in probability choice, adjustment, and computation also develops relatively late. Biases and heuristics, on the other hand, appear early, with variable developmental trajectories. How can we put these findings together?

It seems hard to avoid some sort of dual/multiple processing framework for integrating these data, contrasting intuitive with analytic forms of cognition (Evans, 2008; Hammond, 2002; Weber and Johnson, 2009). Development is extended for more analytic, explicit tasks, that is, tasks using comparative choice or computational formats. Intuitive processes, in contrast, are available early. Estimation tasks draw on these (e.g., Dixon and Moore, 1996) – but so do tasks eliciting heuristics and biases (e.g., Kahneman and Frederick, 2005). Similar performance contrasts appear for children and adults, who do well in scaled-up versions of intuitive estimation, but not computational tasks, or in tasks designed to induce heuristics/biases. Dual process accounts may be theoretically underspecified (Evans, 2008), but from the present point of view they capture well that probability concepts are not unitary, all-or-none affairs (Anderson, 1991; Falk and Wilkening, 1998), with different knowledge states elicited in different tasks or contexts.

Is there development at all from this multiple process/state perspective, or just task variability? Traditionally, analytical thinking is seen as the more advanced, adaptive form of cognition, with development proceeding from undifferentiated – and therefore maladaptive – intuition to reflective and deliberate analysis. The first claim about the ultimate standard does not hold generally. Intuitive approaches to JDM can come quite close to formal models, as reviewed above, and are within the grasp of children and adults, while equivalent computational approaches are beyond their capabilities. More generally, intuition would not exist if it were not largely adaptive. Where computation provides precise answers, intuition

provides cognitively efficient, approximate shortcuts, becoming faulty mainly when misapplied (Gigerenzer, Todd and the ABC Research Group, 1999).

The second traditional claim, about developmental trajectory, also does not hold. The data reviewed here do not demonstrate that children generally reason with inferior intuitive strategies, which eventually are replaced by superior analytical strategies. Only occasionally biases appear to decrease with age. One dual process theory, Fuzzy Trace theory, takes the opposite view: that development proceeds towards intuition (e.g., Reyna, 2006; Reyna and Ellis, 1994). Very young children's intuitive competence in estimation tasks argues against this view as well, although again there are some examples of adults being better than children at discerning the intuitive gist of a problem. Overall, however, JDM development is not unidirectional in nature (Jacobs and Klaczynski 2002).

For a better understanding of specific developmental effects, we need to work on the interplay of intuition, analysis, and task characteristics (Hammond, 1996). Which tasks trigger intuitive or analytical strategies, or combinations of both? Ideas about particular tasks, reviewed above, were developed after it became evident, for instance, that choice and judgement yielded radically different data. What we need, however, is a better a priori understanding of the conditions triggering intuitive or analytic approaches.

Equally crucial is a better understanding of salience effects. Whether behaviour is successful or not largely depends on whether the strategy triggered is the one required. If relevant task features are not salient, or if irrelevant features are overly salient, inappropriate strategies ensue. Functional measurement methodology is successful at providing task structures that trigger highly appropriate strategies, while heuristics/biases research specializes in designing misleading tasks. In both cases, performance is a function of the person-situation match, not just a reflection of peoples' competence or incompetence. The Achilles heel of intuition is, perhaps, that in the absence of an articulated understanding of why a strategy applies, the reasoner has little control over its application, but is at the mercy

of circumstance. This is why basic intuitions need to be educated and integrated with later developing more analytical understanding.

Another gap in developmental research concerns the role of emotional processes. Recent progress appears in discussions of development in the neuro-cognitive systems underlying JDM, in particular, an evolutionarily older system for socio-emotional reward processing in limbic/paralimbic areas and a more neo-cortical cognitive-control system (e.g., Steinberg, 2007). These systems dissociate in adolescence when the pleasure/reward system matures faster than the cognitive system. This can account for the increased risk-taking traditionally attributed to adolescents. Figner, Mackinlay, Wilkening and Weber (2009) showed this increased risk-taking and simplified information use experimentally, in tasks comparing adolescents' and adults' affective ("hot") and deliberative ("cold") decision-making, although further work is necessary to separate affective/deliberative processing from processing of serially/summarily presented information. More generally, little is known about how these neuro-cognitive systems relate to more traditional intuition-analysis dualities or how emotion affects JDM in younger children.

All in all, we have shown that children first learn about probability and EV intuitively, prior to formal instruction from everyday experience. Some even speculate about a biological disposition. Probability instruction could build on this intuitive knowledge and attempt to introduce basic concepts earlier in the primary school years. Nevertheless, intuitive knowledge can also mislead. Thus our final conclusion is not much different than in other, more intensely studied domains of cognitive development: Basic processes and concepts, even structurally complex ones as considered here, operate from early on. What develops only slowly is the ability to explicitly understand the implications and consciously control the application of these concepts and processes.

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..
- Agnoli, F. (1991). Development of judgemental heuristics and logical reasoning: Training counteracts the representativeness heuristic. *Cognitive Development, 6*, 195-217.
- Anderson, N. H. (1980). Information integration theory in developmental psychology. In F. Wilkening, J. Becker, & T. Trabasso (Eds.), *Information integration by children* (pp. 1-45). Hillsdale, NJ: Erlbaum.
- 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). (Ed.) *Contributions to information integration theory*. Vols.I-III. Hillsdale, NJ: Erlbaum.
- Anderson, N. H. (1996). *A functional theory of cognition*. Mahwah, 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.
- 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.
- Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals? *Psychological Bulletin, 125*, 591-600.
- Bayless, S. & Schlottmann, A. (2009, in press). Skill-related uncertainty and expected value in 5- to 7-year-olds. *Psicologica. Special issue on Functional Measurement*.

Boyer, T. W. (2007). Decision-making processes: Sensitivity to sequentially experienced outcome probabilities. *Journal of Experimental Child Psychology*, *97*, 28-43.

Brainerd, C. J. (1981). Working memory and the development of probability judgement. *Psychological Review*, *88*, 463-502.

Butzin, C. A., & Anderson, N. H. (1973). Functional measurement of children's judgements. *Child Development*, *44*, 529-537.

Connolly, T., & Zeelenberg, M. (2002). Regret in decision making. *Current Directions in Psychological Science*, *11*, 212-216.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgement under uncertainty. *Cognition*, *58*, 1-73.

Crone, E. A., & Van der Molen, M. W. (2004). Developmental changes in real life decision making: Performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Developmental Neuropsychology*, *25*, 251-279.

Cuneo, D. O. (1982). Children's judgements of numerical quantity: A new view of early quantification. *Cognitive Psychology*, *14*, 13-44.

Davidson, D. (1995). The representativeness heuristic and the conjunction fallacy effect in children's decision making. *Merrill-Palmer Quarterly*, *41*, 328-46.

Davies, M., & White, P. A. (1994). Use of the availability heuristic by children. *British Journal of Developmental Psychology*, *12*, 503-505.

Dixon, J.A., & Moore, C. F. (1996). The developmental role of intuitive principles in choosing mathematical strategies. *Developmental Psychology*, *32*, 241-253.

Evans, J. S. B.T. (2008). Dual-processing accounts of reasoning, judgement, and social cognition. *Annual Review of Psychology*, *59*, 255-278.

Falk, R., & Wilkening, F. (1998). Children's construction of fair chances: Adjusting probabilities. *Developmental Psychology, 34*, 1340-1357.

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberate processing in risky choice: Age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 709-730.

Figner, B., & Schaub, S. (2009, submitted). Development of adaptive risky decision making: Risk sensitivity in judgement and choice.

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

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Science, 99*, 15822-15826.

Fisk, J. E., Bury, A. S., & Holden, R. (2006). Reasoning about complex probabilistic concepts in childhood. *Scandinavian Journal of Psychology, 47*, 497-504.

Furby, L., & Beyth-Marom, R. (1992). Risk taking in adolescence: A decision making perspective. *Developmental Review, 12*, 1-44.

Gaeth, G. J., Levin, I. P., Chakraborty, G., & Levin, A. M. (1990). Consumer evaluation of multi-product bundles. *Marketing Letters, 2*, 47-57.

Garron, N., & Moore, C. (2004). Complex decision-making in early childhood. *Brain and Cognition, 55*, 158-170.

Gergely, G., & Watson, J. S. (1999). Early socio-emotional development: Contingency perception and the social-biofeedback model. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp.101-136). Hillsdale, NJ: Erlbaum.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684-704.

Gigerenzer, G., Todd, P. M. & the ABC Research Group (1999). Simple heuristics that make us smart. Oxford: Oxford University Press.

- Girotto, V., & Gonzalez, M. (2008). Children's understanding of posterior probability. *Cognition*, *106*, 325-344.
- Hammond, K. R. (1996). *Human judgement and social policy*. Oxford: Oxford University Press.
- Harbaugh, W. T., Krause, K. & Vesterlund, L. (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics*, *5(1)*, 53-84.
- Hastie, R., & Parks, B. (1986). The relationship between memory and judgement depends on whether the judgement task is memory-based or on-line. *Psychological Review*, *93*, 258-268.
- 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.
- Hogarth, H. J., & Einhorn, R. M. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*, 1-55.
- Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, *11*, 107-121.
- Huber, B. L., & Huber, O. (1987). Development of the concept of comparative subjective probability. *Journal of Experimental Child Psychology*, *44*, 304-316.
- Hughes, M. (1986). *Children and number: Difficulties in learning mathematics*. Oxford: Blackwell.
- Jacobs, J. E., & Klaczynski, P. A. (2002). The development of judgement and decision making during childhood and adolescence. *Current Directions in Psychological Science*, *11*, 145-149.
- Jacobs, J. E., & Klaczynski, P. A. (Eds., 2002). *The development of judgement and decision making in children and adolescents*. Mahwah, NJ : Lawrence Erlbaum.

- Jacobs, J. E., & Narloch, R. H. (2001). Children's use of sample size and variability to make social inferences. *Applied Developmental Psychology, 22*, 311-331.
- Jacobs, J. E., & Potenza, M. (1991). The use of judgement heuristics to make social and object decisions: A developmental perspective. *Child Development, 62*, 166-178.
- Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., Caverni, J.-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106*, 62-88.
- Kahneman, D., & Frederick, S. (2005). A model of heuristic judgement. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267-293). Cambridge: Cambridge University Press.
- Kerr, A., & Zelazo, P. D. (2004). Development of "hot" executive function: The children's gambling task. *Brain and Cognition, 55*, 148-157.
- Klaczynski, P. A., & Cottrell, J. M. (2004). A dual-process approach to cognitive development: The case of children's understanding of sunk cost decisions. *Thinking and Reasoning, 10*, 147-174.
- Kuzmak, S., & Gelman, R. (1986). Young children's understanding of random phenomena. *Child Development, 57*, 559-566.
- Levin, I. P., & Hart, S. (2003). Risk preferences in young children: Early evidence of individual differences in reaction to potential gains and losses. *Journal of Behavioral Decision Making, 16*, 397-413.
- Levin, I. P., Hart, S. S., Weller, J. A., & Harshman, L. A. (2007a). Stability of choices in a risky decision-making task: A 3-year longitudinal study with children and adults. *Journal of Behavioral Decision Making, 20*, 241-252.

- Levin, I. P., Weller, J. A., Pederson, A. A., & Harshman, L. A. (2007b). Age-related differences in adaptive decision-making: Sensitivity to expected value in risky choice. *Judgement and Decision Making*, 2, 225-233.
- Levin, I., Wilkening, F., & Dembo, Y. (1984). Development of time quantification: Integration and non-integration of beginnings and endings in comparing durations. *Child Development*, 55, 2160-2172.
- Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, 89, 46-55.
- Lopes, L. L., & Ekberg, P. S. (1980). Test of an ordering hypothesis in risky decision making. *Acta Psychologica*, 45, 161-167.
- Lopes, L. L., (1987). Between hope and fear: The psychology of risk. *Advances in Experimental Social Psychology*, 20, 255-295.
- Mellers, B. H., Chang, S., Birnbaum, M. H., & Ordonez, L. D. (1992). Preferences, prices, and ratings in risky decision-making. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 347-361.
- Payne, J. W. (1982). Contingent decision behaviour. *Psychological Bulletin*, 92, 382-402.
- Piaget, J., & Inhelder B. (1958). *The growth of logical thinking in children and adolescents*. (A. Parsons & S. Milgram, Trans.). New York: Basic Books. (original publ. 1955)
- Piaget, J., & Inhelder B. (1975). *The origin of the idea of chance in children*. (L. Leake, Jr., P. Burrell, & H. Fischbein, Trans.). New York: Norton. (original publ. 1951)
- Rapp, A., & Wilkening, F. (2005). Children's recognition of the usefulness of a record: Distinguishing deterministic and probabilistic events. *European Journal of Developmental Psychology*, 2, 344-363.
- Real, L. A. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science*, 253, 980-986.

- Reyna, V. F. (2006). How people make decisions that involve risk. *Current Directions in Psychological Science*, *13*, 60-66.
- Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgement: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability*. Wiley: Chichester.
- Reyna, V. F., & Ellis, S. C. (1994). Fuzzy trace theory and framing effects in children's risky decision making. *Psychological Science*, *5*, 275-279.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-months-old infants. *Science*, *13*, 274(5294), 1926-1928.
- Schlottmann, A. (2000). Children's judgements of gambles: A disordinal violation of additive utility. *Journal of Behavioral Decision Making*, *13*, 77-89.
- Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *Child Development*, *72*, 103-122.
- Schlottmann, A., & Anderson, N. H. (1994). Children's judgements of expected value. *Developmental Psychology*, *30*, 56-66.
- Schlottmann, A., & Anderson, N. H. (1995). Belief revision in children: Serial judgement in social cognition and decision making domains. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1349-1364.
- Schlottmann, A., & Anderson, N. H. (2007). Belief learning and revision studied with information integration theory. *Teorie & Modelli*, *12*, 63-76.
- Schlottmann, A., & Christoforou, M. (in preparation). *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*, 153-171.
- Schneider, K., Hanne, K., & Lehmann, B. (1989). The development of children's achievement-related expectancies and subjective uncertainty. *Journal of Experimental Child Psychology*, *47*, 160-174.

Shanks, D. R., Tunney, R. J., & McCarthy, J. D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making, 15*, 233-250.

Shanteau, J. (1975). An information integration analysis of risky decision making. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgement and decision processes* (pp. 109-137). New York: Academic Press.

Siegler, R. S. (1981). *Developmental sequences within and between concepts*. Monographs of the Society of Research in Child Development, 46(2, Serial No. 189).

Smith, H. D. (1999). Use of the anchoring and adjustment heuristic by children. *Current Psychology: Developmental, Learning, Personality, Social, 18*, 294-300.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). The origins of knowledge. *Psychological Review, 99*, 605-632.

Steinberg, L. (2007). Risk taking in adolescence. *Current Directions in Psychological Science, 16*, 55-59.

Stevenson, H. W. (1970). Learning in children. In P. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 1, pp. 849-938). New York: Wiley.

Surber, C. F. (1984). Issues in quantitative rating scales in developmental research. *Psychological Bulletin, 95*, 226-246.

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 Science, 104*, 19156-19159.

Weber, E. H., & Johnson, E. J. (2009). Mindful judgement and decision making. *Annual Review of Psychology, 60*, 53-85.

Wilkening, F. (1981). Integrating velocity, time, and distance information: A developmental study. *Cognitive Psychology, 13*, 231-247.

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.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358, 749-750.

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

Zhu, L., & Gigerenzer, G. (2005). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition*, 98, 287-308.

Figure Captions

Figure 1. (a) In the 2-sample probability choice task children can use magnitude estimation and only compare the number of red targets if both samples have the same total. (b) If the totals vary, then a probability, target/target+non-target strategy produced a different response than the target-only strategy.

Figure 2. Mean probability judgement for four ages ($n=32$ per age group) in the Anderson and Schlottmann (1991) study. Very similar data for 8-year-olds appeared in Wilkening and Anderson (1991).

Figure 3. A marble is shaken in a tube lined with coloured paper. For a probability task, games vary in number of segments of winning and losing colours, and children judge how easy it is to win each game. For an EV task, prizes are placed by one/both segments, e.g., if the marble lands on blue (dark), children win one prize, if it lands on yellow (light), they win another prize. Note that identical physical cues differ in meaning in different games: Yellow represents .2 probability in the top, but .5 probability in the bottom game, while the 6 crayon prize represents the higher value, risky alternative at the top, but the lower value, sure thing at the bottom. If judgements show the normative pattern, we can therefore be sure that children judge EV, not the physical cues. Note also that there is no feedback on actual outcomes, as the goal is to assess pre-existing understanding, not children's ability to learn from feedback (after Schlottmann, 2001).

Figure 4. Mean EV judgements of four to six-year-olds ($N=28$), seven- to eleven-year-olds ($N=28$) and adults ($N=17$) judging how happy a puppet would be with various games differing in size of the crayon prize and outcome probability (after Schlottmann, 2001).

Figure 5. Mean judgements of EV for the worth of single (dashed lines) and pairs of independent spinner games (solid lines) for one or two prizes in three age groups (4-7, 8-11, adult) (after Schlottmann, 2000).

Figure 6. (a) Mean judgements of 16 6-year-olds based on 4 successive samples for 16 different sequences (sample composition is listed on the right, + = gold; - =red). Successive bifurcation shows the effects of sample frequency. Recency appears in different judgements for identical sample proportions presented in different order. (b) Serial weight curves derived from these judgements show recency as indicated by the upswing of the curves. The bottom curve shows the influence of samples at Position 1 to 4 on judgement at Position 4; the other weight curves are for judgement at Positions 2 and 3. The final sample always has the strongest influence on the judgement, but its weight is much reduced as soon as the next informer is presented (after Schlottmann & Anderson, 1995).

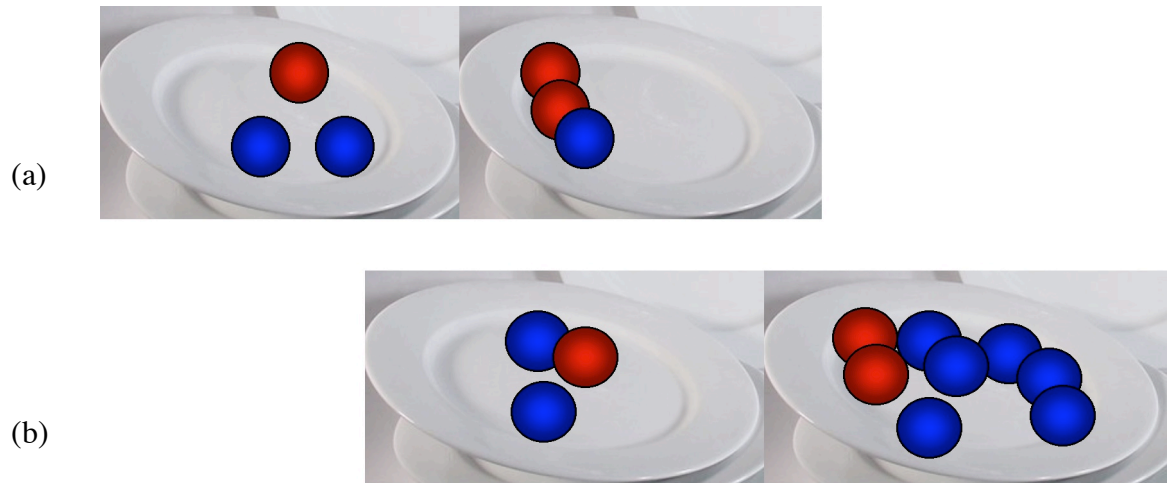


Figure 1

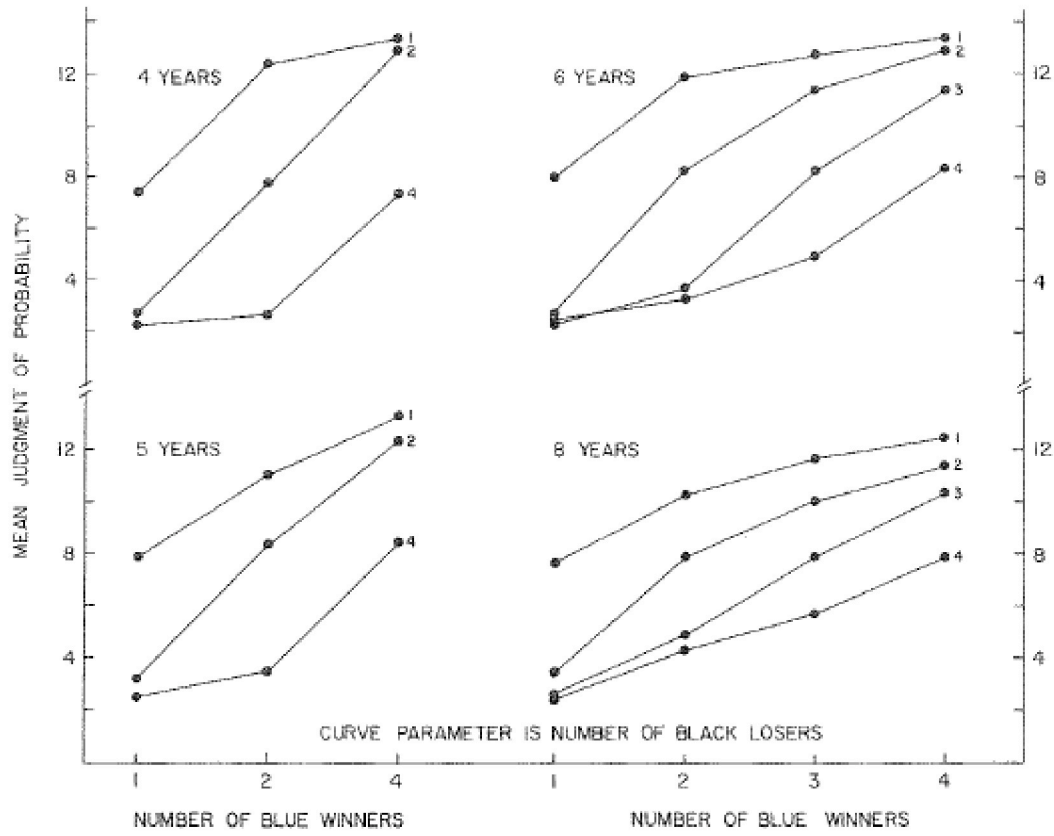


Figure 2

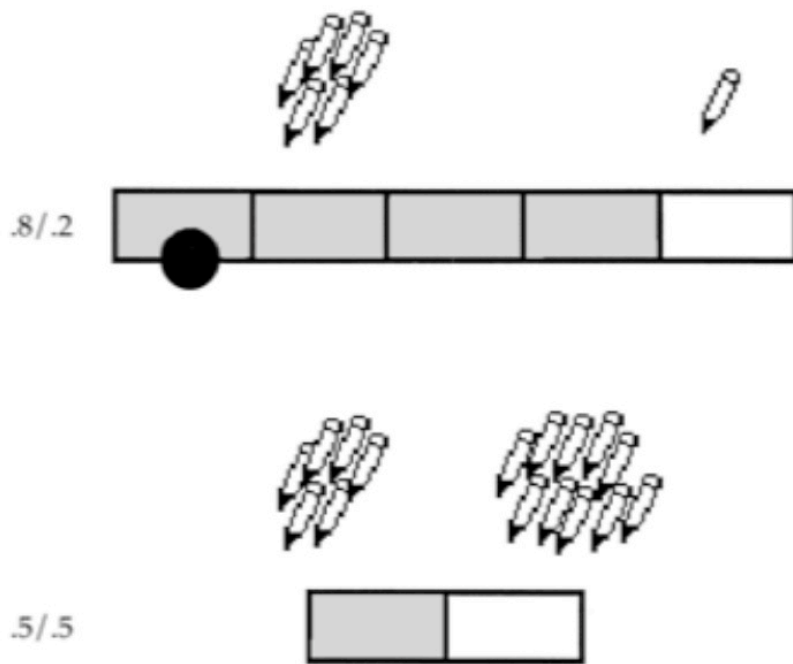


Figure 3

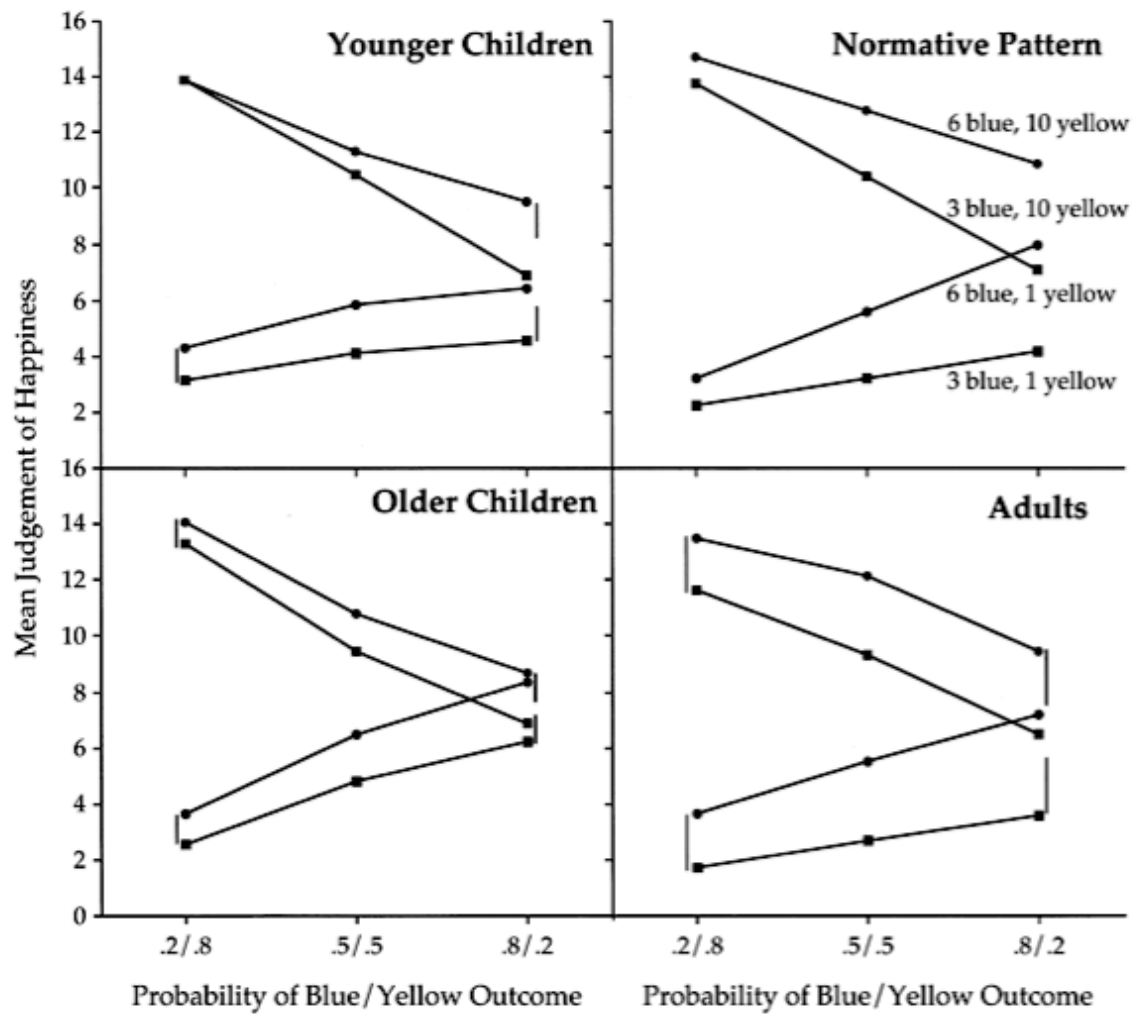


Figure 4

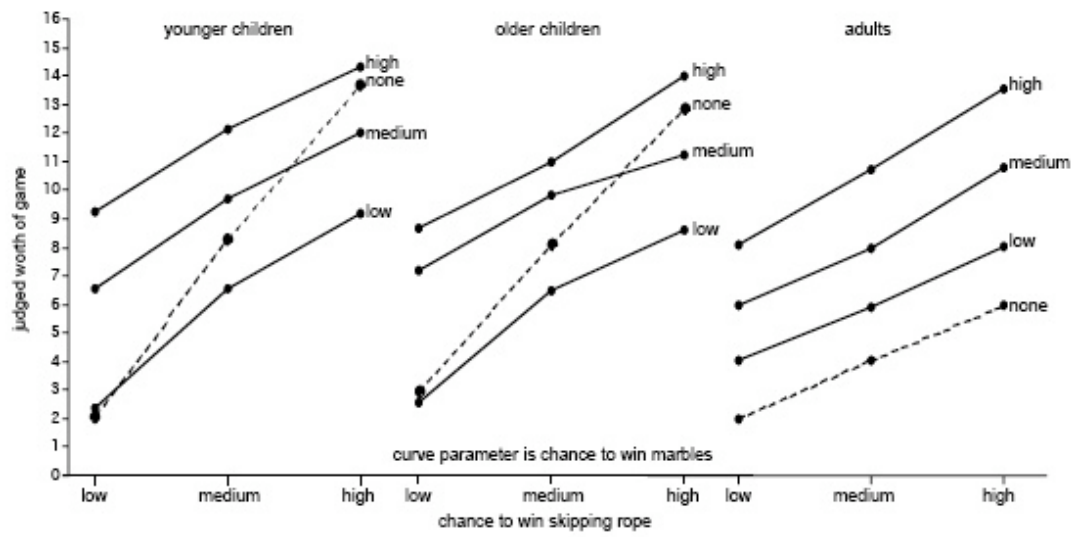


Figure 5

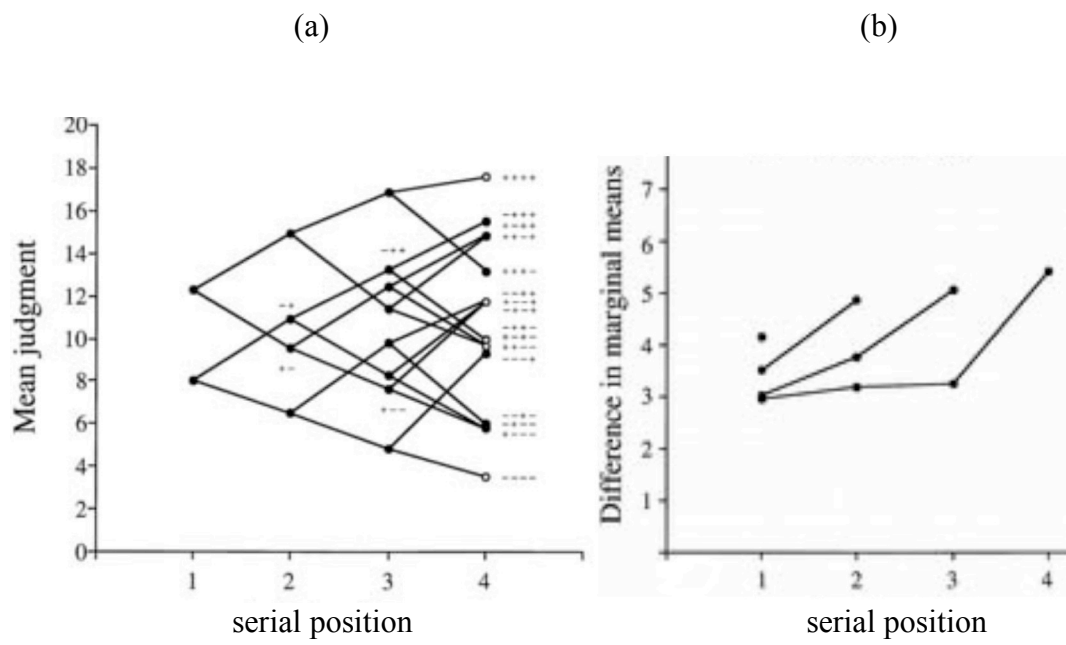


Figure 6