Running Head: DEVELOPMENT OF RISK SENSITIVITY

Development of Adaptive Risky Decision Making:

Risk Sensitivity in Judgment and Choice

Bernd Figner

Columbia University

Simone Schaub

University of Zurich

Author Note

Bernd Figner, Center for the Decision Sciences and Department of Psychology, Columbia University.

Simone Schaub, Department of Psychology, University of Zurich.

This research was supported by a fellowship from the Swiss National Science Foundation (PA001 – 115327 and PBZH1 – 110268) and a grant by the US National Science Foundation (SES – 0720932), awarded to the first author.

Correspondence concerning this article should be addressed to Bernd Figner, Center for the Decision Sciences and Department of Psychology, Columbia University, 406 Schermerhorn Hall, Mail Code 5501, 1190 Amsterdam Avenue, NY 10027, USA. Electronic mail may be sent to bf2151@columbia.edu.

Acknowledgments

We thank Andreas F. Rapp, Elke U. Weber, and several research assistants for their help in various aspects of the study.

Working Paper Version - Please do not cite without permission! -- January 23, 2009 --

Abstract

The ability to make risk-sensitive choices and judgments was examined in 5-, 6-, 8-, and 10-year-old children and adults (*N* = 129). All age groups chose advantageously between high and low outcome variability options in a novel board game varying expected value and aspiration level (i.e., they adaptively chose the option with the higher probability of reaching the aspiration level). Advantageousness of choices and sophistication in information use increased with age. However, judgment and choice patterns came closest to the mathematical normative probabilities in 6-year-olds, followed by adults. Results point to remarkably sophisticated intuitive risk-sensitivity and choice strategies in young children in a mathematically complex task, long before the underlying probabilities can be calculated.

Keywords: cognitive development, risky decision making, risk sensitivity, risky choice, judgment, probability, information integration

Development of Adaptive Risky Decision Making: Risk Sensitivity in Judgment and Choice

Research on children's risky decision making has been relatively sparse and, so far, has focused either on understanding and evaluation of probability and expected value (EV), or on risk taking. Children's understanding of different levels of outcome variability (OV), i.e., risk, and their choices between *two* risky prospects have received virtually no attention, although risk sensitivity, i.e., the ability to adjust one's risk-taking levels to varying circumstances, is an important everyday-life ability for children and adults. In addition, having children choose between two risky options avoids a potential methodological flaw of the typically used choice between a safe riskless and a risky option ("entertainment bias", see below). The current study extends the research on children's risky decision making by drawing from the literature on risk sensitivity in adult humans and non-human species and investigates risk sensitivity in choice and judgment in children ranging from 5 to 10 years in age, and in adults.

Development of Understanding of Probability and Expected Value

Most studies agree that as children get older, they exhibit growing sophistication in understanding and dealing with probabilities and risky situations (Schlottmann & Wilkening, in press). However, there is considerable disagreement about the timing of the developmental trajectory. As we will argue, apparent contradictions in findings can be reconciled by assuming that early *intuitive* competence precedes explicit competence, and that different types of tasks assess different types of competence. Work by Piaget and Inhelder (1975) suggested that children exhibit a mature understanding of probabilities only in the stage of formal operations. More refined follow-up studies largely confirmed the original claims (e.g., Davies, 1965; Hoemann & Ross, 1971; see Reyna & Brainerd, 1994 for an overview). These studies typically used *choice methodology*, e.g., children were asked to select the set with a higher proportion of winning elements out of two sets with winning and losing elements. According to these studies, children up to 4 to 5 years take into account the number of winning elements only (they "center" on the number of winning elements) and appear to have no concept of probability; 6- to 9-year-olds begin to consider the relation between winning and losing elements, but it is not until age 11 or later that children apply the correct proportional strategy consistently.

In contrast, studies using *judgment methodology* and, in particular, functional measurement techniques, demonstrated a much earlier competence for children's intuitive understanding of probability and EV (e.g., Acredolo, O'Connor, Banks, & Horobin, 1989; Anderson, 1980; Schlottmann, 2001; Schlottmann & Tring, 2005; see also Hommers, 1980 and Reyna & Ellis, 1994 for studies using choice and judgment within-subject). Schlottmann (2001), for example, found that children as young as 5 followed the normatively correct multiplicative integration to combine probability and reward magnitude in forming EV, assessed via rating scale. These information integration studies used judgments as the main dependent variable, except Schlottmann and Tring (2005) and Reyna and Ellis (1994) who used judgments and choices in a within-subject design (see below). Recent work by Levin and colleagues, on the other hand, used choice methodology to investigate children's information use in risky decision making (Levin, Weller, Pederson, & Harshman, 2007; see also Levin & Hart, 2003). Levin et al. designed the "cups task" consisting of three types of trials, with the risky option either having a higher, the same, or a lower EV than the riskless option. The authors found that younger children (aged 5-7) showed low sensitivity to differences in EV, i.e., they did not choose the more advantageous option (in their study, the option with the higher EV). Older children (aged 8-11) were more sensitive to EV in their choices but still did not perform as well as adults.

Falk and Wilkening (1998) used a probability-adjustment task and found an even more protracted development of probability understanding. Only children aged 13 used the correct proportional strategy to integrate the number of winning and losing elements. The youngest children (aged 6) relied on only one dimension, and 9- and 10-year-olds integrated the two factors in a non-normative additive fashion. The authors attributed the late development of children's performance to characteristics of their task which trigger more analytical modes of information processing than judgment tasks which trigger mostly intuitive modes. According to Falk and Wilkening, judgment tasks are located at the intuitive end of a dimension that ranges from intuitive to analytical processing. According to them, even young children are able to intuitively evaluate one single option on a meaningful (typically affect-related) scale, such as the happiness of a puppet playing a risky gamble. Choices are considered to be more difficult and require more analytical processing, as the information of at least two options has to be considered. In addition, children not only have to evaluate the options but also have to *choose* one of them, which is assumed to require more analytical processing than evaluation.

Correspondence of Judgment and Choice

Although the method used to investigate children's understanding can strongly influence the findings, comparisons of judgment and choice in the domain of probability and decision making are rare. Reyna and Ellis (1994) and Schlottmann and Tring (2005) assessed choices and judgments in a within-subject design (see also Hommers, 1980). Reyna and Ellis did not report the correspondence of the two response modes but overall found very similar patterns, suggesting substantial correspondence. Schlottmann and Tring tested whether judgments predicted choices, and found little correspondence. The authors even stated that their judgment data highlighted children's sophisticated abilities in decision making, while their choice data highlighted their weaknesses. Finally, Wilkening and Anderson (1982) compared Siegler's rule assessment (a choice methodology) to functional measurement techniques (a judgment methodology) and reported higher diagnostic sensitivity in the judgment method, which could further contribute to the tendency of studies using choice methods to report later development.

Risk Taking and Risk Sensitivity

As outlined above, studies on children's judgment and risky decision-making performance typically focused on EV, i.e., the first moment of a probability distribution. In contrast, the *variability* of a probability distribution, i.e., its second moment, has received far less attention in developmental psychology, although risk sensitivity is an important everyday ability. The few studies that investigated children's risk sensitivity typically gave them the choice between a safe sure-thing option and a simple gamble option (e.g., Harbaugh, Krause, & Vesterlund, 2002; Levin & Hart, 2003; Levin et al., 2007; Reyna & Ellis, 1994). Consistent with the common definition of risk taking as choosing the option with the higher OV, these studies tested whether children are risk seeking or risk averse by giving them the choice between a constant (i.e., safe, riskless) and a variable (risky) option. Reyna and Ellis (1994; see also Reyna, 1996) observed increased risk taking in children in a task that independently varied OV and EV. Additionally, they observed risk sensitivity only from second grade on, but not yet in preschool. As their safe and risky option had the same EV, risk taking was not disadvantageous per se (in the sense of choosing the option with the lower EV). Similarly, studies by Harbaugh et al. and Levin et al. found increased risk taking in children, particularly when it was disadvantageous (see also Huizenga, Crone, & Jansen, 2007). In contrast, Schlottmann and Tring (2005) did not find age differences in risk taking, but children's performance in risky choice was in general suboptimal. All studies agreed, however, on increasing sophistication in risky decision making, such that older children, relative to younger ones, differentiated more between levels of risk or adjusted their choices more strongly with respect to differences in EV (see Boyer, 2006 for an overview of the development of risk taking). Taken together, the studies suggest that children's *choice*, when given a riskless option and a risky option, is far from optimal and their risk sensitivity is immature, compared to adults. However, studies using the choice between a riskless and a risky option to assess risk taking potentially suffer from a methodological problem, particularly if the task is presented as a game and immediate feedback is given about the outcomes. It might be that children are not more risk seeking, per se, but that they simply prefer the more entertaining option. Children might avoid the riskless option, as they always know what they get and no element of chance is involved in learning the outcome. While this does not necessarily argue against the validity of the task, it would be interesting to know if such a mechanism might be at work in producing children's suboptimal performance in risky choice.

A second shortcoming of the existing tasks and studies is that, in everyday-life, risky situations do not always offer a riskless option. Instead, decisions must be made between two (or more) risky options. How different levels of risk (i.e., non-zero OV) affect children's decisions deserves closer investigation, given that OV has been shown to be highly relevant for risky decisions in adult humans (Weber, Shafir, & Blais, 2004) and animals (Kacelnik & Bateson, 1997). In animal research, risk sensitivity, i.e., whether and how animals adjust their behavior to differences in OV, has been widely studied in the framework of risk-sensitive foraging theory (Kacelnik & Bateson, 1997; Stephens, 1981). Only recently has the framework of risk sensitivity been applied to (adult) human decision-making. Several studies have demonstrated that adult humans, similarly to animals, adaptively adjust their decisions depending on the OV and EV of the given options and the required aspiration level (AL). In animal studies, the aspiration level was typically manipulated by their need for caloric intake, e.g., by depriving them of food for shorter or longer periods of time. In human research, the AL can be manipulated, e.g., by varying the number of good lottery outcomes needed to win a prize (Pietras & Hackenberg, 2001; Rode, Cosmides, Hell, & Tooby, 1999; Weber, Shafir, & Blais, 2004).

How expected value (EV) and outcome variability (OV) are taken into account and interact with other variables, most prominently the aspiration level (AL), has led to the "energy-budget rule" in animal foraging behavior (Stephens, 1981). This rule describes a simple decision algorithm that assumes for low ALs (AL ≤ EV), the decision maker will choose options with lower OVs, keeping EV constant (i.e., they will be risk averse). In contrast, for high ALs (AL > EV) the decision maker will choose options with higher OVs, keeping EV constant (i.e., they will be risk seeking). The energy-budget rule describes not only the remarkably sophisticated risky decisions of animals, but also aspects of adult human behavior (Pietras & Hackenberg, 2001; Rode et al., 1999). (Note that the energy-budget rule is similar to the reflection effect in prospect theory (Kahneman & Tversky, 1979) when AL is interpreted as prospect theory's reference point; i.e., outcomes above the AL are gains, and outcomes below the AL are losses). Given the results on animals' and adult humans' risk sensitivity and children's early intuitive competence in probability and EV, it seems plausible that even young children might be able to choose the more advantageous of two risky options differing in OV, adjusting their choices adaptively to levels of EV and AL (note that, in our study, we use the term advantageousness in the sense of the energy-budget rule, i.e., choosing the option with the higher probability of meeting the required AL).

The main goal of the present study was to investigate the development of risk sensitivity by testing whether children of different age groups are able to adjust their choice between two risky options (differing in OV) to varying levels of EV and AL. Simultaneously, using the choice between *two* risky options instead of between a riskless and a risky option allows us to gauge the extent to which children's previously reported suboptimal choice behavior might have been caused by an "entertainment bias" (the tendency to avoid the "boring" safe option rather than a greater risk-taking propensity). Finally, to investigate whether skills of risk sensitivity are present earlier in children's judgments than in their choices, we assessed both in a within-subject design.

Methods

Participants

A total of 105 children and 24 adults took part in the experiment. Children were divided into four roughly equal-sized age groups; 24 5-year-olds (12 girls and 12 boys, 5;1 to 5;11 years old, *M* = 5;6), 25 6-year-olds (12 girls and 13 boys, 6;0 to 7;0 years old, *M* = 6;6), 27 second graders (14 girls and 13 boys, 7;10 to 10;0 years old, *M* = 8;9), and 29 fourth graders (16 girls and 13 boys, 9;3 to 11;7 years old, *M* = 10;7). The adult group consisted of 15 females and 9 males (19 to 53 years old, *M* = 32). Despite some overlap in age ranges, children from second and fourth grade will be referred to as 8- and 10-year-olds, respectively. Children were recruited from kindergarten, primary school, and a day nursery in the German speaking part of Switzerland and were tested individually in sessions of 30 to 45 minutes. Informed assent from children and informed consent from children's parents was obtained prior to participation. Adults were recruited by word of mouth from the University of Zurich and tested in a university laboratory. Typical of the local population, participants came from primarily Caucasian middle- to upper-middle income families and were fluent in the Swiss-German language.

Procedures and Materials

To ensure children's understanding of the task and motivation, we chose a relevant-involvement method (Falk & Wilkening, 1998). The decision-making task was presented as an attractive board game in which children had to help a mouse escape from a cat. Participants' choices were incentive compatible: At the end of the task, six trials were played for real outcomes and participants won a prize for each favorable outcome of their decisions. During the task, however, participants did not receive feedback on their choices or judgments. The general framework of the task was to rescue a mouse from a cat: In each trial, a two-colored die is rolled 6 times, and the cat moves one field forward with each roll, irrespective of the outcome of that roll. The mouse moves only if the roll shows the winning color. If the mouse is not in its safe house by the end of the 6th roll, the cat catches the mouse (see Figure 1). The 30 trials of the task differed on two factors, AL and EV, which varied according to a factorial design.

----- Insert Figure 1 about here -----

Aspiration level (AL) was operationalized by how far the mouse's starting point was away from its safe house, either 1, 3, or 5 moves. I.e., depending on AL, the mouse was saved if, out of the 6 rolls, at least 1, 3, or 5 came up with the winning color. *Expected value (EV)* was operationalized by the die used in a given trial. The die with the highest EV had 5 sides with the winning color and 1 side with the losing color. I.e., the EV for getting the winning color in 6 rolls was 5. The other levels of EV were represented by dice with 4, 3, 2, and 1 winning color sides, with the remaining sides having the losing color. The 30 trials consisted of the 3 levels of AL combined with the 5 levels of EV, with each of the 15 different combinations presented twice, once in each of two blocks. Trials were presented in one of three pseudo-random orders. *Outcome variability (OV)* was incorporated in the experimental design by giving the children, for each trial, the choice between rolling the known die of the current trial (low OV option) or drawing blindly a die from a bag containing 7 different dice, each one with 0, 1, 2, 3, 4, 5, and 6 sides of the winning color (high OV option). We refer to the known die as the low OV option and to the bag as the high OV option because, irrespective of which known die is presented in a trial, the OV of the known die is always lower than the OV of the bag.

Dependent measures. In each trial, the children's *choice* of the low or the high OV option was recorded, followed by their *judgment* of the happiness of the mouse. The latter served as assessment of children's subjective probability estimate that the mouse will be saved. These ratings were done with the use of a 100 cm long wooden bar that had endpoints labeled with a happy and a sad face; a technique successfully used in functional measurement studies with children as young as 4 years of age (Anderson, 1980). After each choice, children rated only the happiness of the mouse for the low OV option, irrespective of their actual choice. This was done because only the EV of the low OV option differed from trial to trial. Happiness ratings for the high OV option were acquired in the first three trials, which were chosen to represent all three ALs. In these three trials, children rated the low and the high OV option. Detailed instructions including practice trials were given to children prior to the start of experimental trials, introducing the different aspects and rules of the game in a child-appropriate stepwise manner. All children appeared to understand the point of the game. Depending on the combination of AL and EV, either the high or low OV option had the higher probability of reaching the required AL, and thus was more advantageous (see Table 1 and Figure 2 for the normative solutions and the associated probabilities). To avoid possible effects of outcomes on subsequent trials and possible "entertainment bias" effects, participants did not receive feedback during experimental trials; i.e., they never actually drew from the bag or rolled any die until the end of all trials.

----- Insert Table 1 about here -----

Results

Data were analyzed with respect to three main aspects: (1) Advantageousness and risk taking of participants' choices, (2) information use underlying choices and judgments, analyzed at the group and the individual level, and (3) correspondence between choices and judgments. As trial order did not reveal significant effects on participants' choices or judgments, data have been collapsed across orders for all analyses [choice: $F(2, 126) = 0.88$, $p = .42$, $\eta^2 = .01$; judgment: $F(2, 126) = 0.04$, $p = .96$, $\eta^2 = .001$; partial η^2 values are reported throughout the paper].

Advantageousness and Risk Taking

Advantageousness of choices. First, we looked at how often participants chose the more advantageous of the two available options (i.e., the option with the higher probability of meeting the AL). As can be seen in Table 2, the number of advantageous choices increased monotonically with age from, on average, 18.75 advantageous choices in 5-year-olds to 26.79 in adults [$F(4, 124) = 23.14$, $p < .001$, $\eta^2 = .43$]. The increase was steeper among the younger age groups and flattened out towards the older age groups. This pattern was reflected by the significant linear and quadratic trends in the proportion of advantageous choices across age groups ($p_{\text{linear}} < .001$; $p_{\text{quadratic}} < .01$). As can be seen from the separate columns for advantageous and disadvantageous choices of the low and the high OV option, age trends were uniform for high and low OV options. I.e., younger children's less optimal performance cannot be explained by avoiding (or being attracted by) either the high or low OV option. Although the 5-year-olds clearly made the least number of advantageous choices, their number of advantageous choices was still significantly above chance, as the choices of the other age groups (for all age groups' one-sample t-tests $p < .001$; testing against 15 advantageous choices expected by guessing).

Risk taking. Consistent with the common definition of risk taking as choosing the option with the higher OV and consistent with other developmental studies (e.g., Levin et al., 2007), we operationalized risk taking as how often participants chose the high OV option (irrespective of whether it was advantageous or disadvantageous). As can be seen in Table 2, all age groups exhibited fairly equal levels of risk taking. Consistent with this impression, an analysis of variance (ANOVA) on the number of high OV choices per age group revealed no significant effect of age [$F(4, 124) = 0.14$, $p = .97$, $\eta^2 = .005$]. As can be seen from Table 2, disadvantageous choices in all age groups consisted of approximately 50% high OV choices, suggesting little developmental differences. In fact, the proportion of high OV choices in disadvantageous choices was highest in adults (1.92 disadvantageous bag choices out of 3.21 disadvantageous choices in total), indicating that children's suboptimal choices could not be explained by increased risk-taking tendencies.

----- Insert Table 2 about here -----

Information Use and Decision Strategies

To gain insight into children's and adults' strategies and information use underlying their responses, we first analyzed their choices and judgments at the level of age groups, followed by individual-level analyses of information use and choice strategies.

Group-level analysis. Information integration methodology (Anderson, 1996; Figner, Mackinlay, Wilkening, & Weber, in press; Figner & Voelki, 2004) was used to investigate which variables (AL and/or EV) influenced participants' judgments and choices. Judgments and choices were analyzed separately for each age group by using repeated-measures ANOVAs, all following a 2 (Block) × 3 (AL) × 5 (EV) design. Block never yielded a significant main effect in any age group and showed only rare and small interactions with AL or EV that did not appear to follow a consistent pattern. Accordingly, results for Block will not be reported. Participants' *choices* were analyzed as the proportion of low OV choices (i.e., proportion of die choices), with choice of high OV option coded as 0 and choice of low OV option coded as 1. For the *judgments*, we focused on the ratings for the low OV option, because only this option varied with respect to EV from trial to trial (for results for the ratings of the high OV option, see below).

As shown in Figure 2 and Table 3, *5-year-olds'* responses, of all the age groups, deviated most strongly from the normative pattern. For their judgments, both visual inspection and statistical analysis suggest that they took into account both AL and EV. However, the pattern does not follow the normative barrel pattern, confirmed by the non-significant interaction of AL × EV. In addition, the differentiation between the factor levels was small and the whole pattern was shifted towards the upper end of the rating scale, i.e., towards increased happiness. Their choice pattern was even more irregular than their judgments, and the effect for AL did not reach statistical significance, suggesting they centered on EV. For the *6-year-olds*, we found clear effects for the different levels of AL and EV in both judgment and choice, with patterns closely following the normative barrel shape. This visual impression is reflected in significant effects for AL, EV, and their interaction in both response modes. The *8- year-olds* and *10-year-olds* differed little from each other. Both showed clear effects of AL and EV in their choices as well as their judgments. However, the interaction of AL and EV was significant only in their choices, not their judgments (with the interaction term for 8-year-olds' judgments reaching marginal significance). This is reflected in the respective graphs for their judgments and choices, forming rather parallel lines for the judgments but following the barrel shape for the choices. *Adults'* choices and judgments were again closer to the normative pattern with clear and significant effects of AL, EV, and their interaction. However, compared to both the normative pattern and the 6-year-olds' judgments, adults' judgments were more parallel and less barrel-shaped.

As can be seen in Figure 2 (bottom graph), participants' *choices* in none of the age groups matched the normative pattern exactly. From a normative perspective, participants should always choose the option with the higher probability of reaching the AL, i.e., they should be *maximizing*. However, the observed patterns are somewhat similar to the normative mathematical probabilities of the low OV option (see bottom right panel of the upper graph). Choosing proportional to the probabilities of success instead of maximizing success has been widely shown in both human and non-human species ("probability matching"; Herrnstein, 1970). In our case, the patterns do not perfectly reflect either the normative choice pattern (maximizing) or the normative probabilities (probability matching), but seem to be a mix between the two strategies with a tendency for maximizing becoming more dominant with increasing age.

----- Insert Table 3 and Figure 2 about here -----

Happiness ratings for the high OV option were analyzed separately for each age group using repeated-measures ANOVAs with the ratings for AL 1, 3, and 5 as dependent variables. As shown in Table 4, the judgments of the 5-year-olds (among all age groups) reflected most closely the decrease of probabilities across the three ALs (see right-most column of Table 1 for normative probabilities). Consistent with this impression, we found a significant effect for AL only for the 5-year-olds but not for the other age groups [in order of ascending age: $F(2, 46) = 5.02$, $p < .05$, $\eta^2 = .18$; $F(2, 48) = 1.51$, $p = .23$, $\eta^2 = .06$; $F(2, 52) = 0.29$, $p = .75$, $\eta^2 = .01$; $F(2, 56) = 2.28$, $p = .11$, $\eta^2 = .08$; $F(2, 46) = 1.27$, $p = .29$, $\eta^2 = .05$]. Except for the 5-year-olds, participants barely adjusted their ratings to the different ALs and rated the happiness more or less constantly at 50%. It is not clear why participants barely adjusted their ratings to the ALs. As we always assessed these ratings right after participants had judged the low OV option, this might have led to contrast effects and/or confusion about the second rating. However, it is still somewhat puzzling that of all age groups, the 5-year-olds gave the most normative ratings. Irrespective of the reasons for these findings, because these ratings appear not to be very valid and are not central to the remaining analyses, the judgments for the high OV option will not be considered further.

----- Insert Table 4 about here -----

Individual-level analysis. Here we tested for each participant separately how the information on AL and EV influenced their responses. *Judgments* were analyzed similarly to the group-level analysis, following standard functional measurement methodology (Anderson, 1996). For each participant, we computed an ANOVA to determine how AL and EV influenced his/her ratings. Following Falk and Wilkening (1998), we adopted a *p*-level of .1 in order to avoid overlooking more complex strategies. Starting with the simplest judgment strategy, we categorized the participants as follows: If neither AL nor EV were significant, we categorized them as random responders. However, there were 4 5-year-old children who constantly gave "totally happy" ratings (and accordingly had nonsignificant effects for AL and EV). These participants were categorized not as random responders but as "Constant Happy" (see Figure 3, top graph). Participants with a significant effect only for AL (i.e., their ratings were significantly influenced only by the different ALs) were categorized as centering on AL ("AL Only"). Participants with a significant effect only for EV were categorized as centering on EV ("EV Only"). Participants with significant effects for AL and EV but a non-significant interaction were categorized as having a subtractive strategy ("Diff EV-AL"). The ratings of these participants were consistent with the energy-budget rule described further above (Stephens, 1981), comparing AL and EV by subtracting AL from EV to determine whether AL was higher, lower, or the same as EV. Lastly, participants with significant main effects and a significant interaction of AL and EV were categorized as having a normative multiplicative strategy ("Normative").

Choices had to be analyzed differently, because the binary choice format did not lend itself to be analyzed by individual-level ANOVAs. To analyze participants' choice strategies, we used an approach similar to Falk and Wilkening (1998). For each participant, we tested separately which of several models provided the best fit to their choice data by counting the number of correct predictions each model made. To be able to compare categorizations across response modes, we used models as analogous as possible to the ones in the judgment analysis. As criterion for a significant fit, the model had to make at least 22 correct predictions. This was based on the reasoning that 18 correct predictions for the normative model would result by simply always choosing the die (see Table 1). The probability to make 4 additional correct choices by guessing was $(0.5)^4 = .0625$. Therefore, the criterion of 22 correct predictions resulted in a threshold similar to the $p < .10$ used in the individual-level judgment analysis.

The tested choice models included the normative choice model plus five simplified choice strategies, similar to the ones in the judgment analyses (see Figure 3). The least sophisticated information-use category was random responding; this category was assigned when none of the tested models met the significance criterion. The next two models were choosing always the low OV option ("Constant Die") and choosing always the high OV option ("Constant Bag"). These categories were assigned when at least 27 choices followed the respective model. The reasoning for this threshold is analogous to the 22 correct prediction criterion mentioned above, such that a deviation of 4 or more choices from the constant choice model constituted a significant deviation. Centering on AL was modeled as taking the low OV option as long as the AL was below or equal to 3 ("AL Only"). Centering on EV was modeled as taking the low OV option as long as its EV was equal to or higher than 3 ("EV Only"). The difference strategy was modeled after the energy-budget rule: For AL ≤ EV, the low OV option should be chosen. For AL > EV, the high OV option should be chosen. The best-fitting model for each participant was determined by the following criteria: (1) 22 or more correct predictions (27 or more for the constant models, see above). (2) If more than one model was identified by criterion (1), the model with the highest number of correct predictions was chosen. (3) When criterion (2) created ties, the most complex model was chosen. To test whether these somewhat arbitrary rules introduced biases in the developmental patterns, we inspected the results separately (a) when we did not restrict the model testing by the above-mentioned criteria (i.e., more than one model could be assigned per participant), (b) when participants with more than one significant model were excluded, and (c) when, in the case of ties, the least complex instead of the most complex model was counted. Although in some of the cases (a) to (c), absolute frequencies for the models were different from what we are reporting here, the developmental trends reported below stayed the same. We take this as evidence for the relative robustness of our analysis.

Figure 3 (top) shows the categorization of participants per age group for judgments; Figure 3 (bottom) shows the respective graph for choices. Aside from notable differences between the two response modes, there were similarities in the developmental patterns which can be described by three trends: (a) The simplest strategies (random and constant responding) decreased sharply within the youngest three age groups and were not found beyond age 8. (b) The normative strategy increased across age groups, with the exception of a decrease in the 10-year-olds. (c) Simplified strategies (centering on EV, centering on AL, difference strategy) predominantly followed an inverted U-shaped pattern. Perhaps the most striking difference between the two response modes can be seen in the frequencies of the difference and the normative strategy, most exemplary in the adult sample. In their choices, most adults were consistent with the normative strategy, whereas in their judgments, most appeared to follow the difference strategy. There are at least two possible reasons for this. Participants (particularly adults) may indeed have had more complex choice than judgment strategies. Alternatively, the difference in apparent complexity of choice vs. judgment strategies might be due to the fact that the individual-level ANOVAs of the judgments lacked the statistical power to diagnose the interaction of AL and EV. Considering the information-use graphs for judgment and choice at the group level (Figure 2), the former explanation appears to be more likely accurate given that the interactions were clearer in choice than judgment. This point will be investigated more closely in the next section.

----- Insert Figure 3 about here -----

Correspondence between Judgment and Choice

We examined the correspondence between choice and judgment in two ways. First, we categorized participants' choice and judgment strategies into four broad categories of increasingly sophisticated information use and tested the ordinal correspondence of the categorizations across response modes. If choices and judgments were indeed utilizing similar decision processes, we would expect a high correspondence between the choice and judgment categorizations. Further, we used the categorizations to test whether one response mode diagnosed more sophisticated decision strategies earlier than the other. As a second approach, we used participants' judgments to predict their choices. If choices and judgments captured similar underlying decision strategies, we would expect judgments to predict choices well.

Correspondence of choice and judgment categorizations. Participants' response strategies were categorized into 4 broad categories based on the individual-level analyses reported above, each for choice and judgment. The four categories were, in the order of increasing complexity of information use, (1) unclassifiable (i.e., random or otherwise nonsystematic responses), (2) constant, non-adaptive responses (e.g., always choosing the same option or always judging "totally happy"), (3) centration (either on AL or EV), and (4) integration (taking into account both AL and EV, comprising both the difference and the normative strategy). The strength of association between strategies across response modes was tested for each age group, as well as for the whole sample, with Kendall's τ. As can be seen in Table 5, overall, participants' choices and judgments corresponded significantly. Looking at the age groups separately, it becomes evident that the correspondence between the two response modes was low for the 5-year-olds and the adults, and high for the three middle age groups, i.e., the 6- to 10-year-olds. The non-correspondence in the 5-year-olds might not be surprising given that their judgments and choices both appeared to be somewhat noisy. In the adult group, the low correspondence might have been caused by lack of variance in the categorizations. Most adults were categorized as using an integrating strategy (71% for choice and 88% for judgment) with the rest being categorized as centration on EV. This skewed distribution together with the somewhat blurred diagnosis of the normative versus "EV Only" choice strategy —the models differed only in 4 of 30 trials — presumably was at least partly responsible for the low τ in adults.

To test whether one response mode was more sensitive to diagnose more complex strategies than the other, we compared the medians for the choice and judgment categorizations in each age group. As can be seen in Table 5, the medians (and means) differed only slightly between the choice and the judgment categorizations within age groups. This impression was reflected by a non-significant sign test ($Z = -1.01$, $p = .31$).

----- Insert Table 5 about here -----

Predicting choices by judgments. An even stricter test for the correspondence between choices and judgments is to use the judgment of an option to predict whether it is chosen or not (Schlottmann & Tring, 2005). Ideally, we would have compared the judgments of the high and the low OV option for each single trial to predict that the option with the higher happiness rating would be chosen. However, this approach was not feasible given that ratings for the high OV appeared problematic, as determined by the group-level analyses. Because the choice predictions for all 30 trials would have depended on these three unreliable ratings of the high OV option, this would have introduced noise to our choice predictions. As an alternative, we chose an approach that seemed feasible given our data and at the same time was conservative with respect to producing false positives. As a prediction of whether participants would choose the high or low OV option, we looked at the happiness rating for the low OV option alone. If the rating was 50% or higher (indicating happiness rather than unhappiness with the low OV option), we predicted choice of the low OV option. If the happiness rating was 49% or lower (indicating unhappiness with the low OV option), we predicted choice of the high OV option. Using this criterion, participants' judgments predicted 69.6% of their choices correctly. These numbers were fairly similar across age groups with 72.4%, 74.3%, 68.3%, 67.5%, and 66.1% for 5-, 6, 8-, 10-year-olds, and adults, respectively. The predictions were above chance in all age groups, tested against 50% correct predictions that would be expected by random guessing (all *p*s < .001).

Discussion

We investigated the development of risk sensitivity and risky decision-making in children between 5 and 10 years of age and adults. The novel features of our study were that (a) children had to integrate outcome variability (OV), expected value (EV), and aspiration level (AL) in choosing between two risky options, (b) we assessed choices and judgments in a within-subject design, and (c) we investigated not only children's decision but also their underlying information use in a mathematically complex task that followed relevant-involvement methods. Our main findings can be summarized as follows: (1) Even at 5 years of age, children were performing above chance in making advantageous choices, and performance improved further with age. (2) All age groups exhibited virtually identical levels of risk taking, implying that the advantageousness of choices was independent of risk preferences across age groups. (3) Sophistication of information use, analyzed at the individual-participant level and at the age-group level, generally improved with age. However, at the group-level, 6-year-olds' judgments came closest to the normative probabilities, followed by adults', then 8-year-olds' responses. There was evidence for a transient decline in performance in 8- to 10-year-olds in both judgment and choice. (4) Correspondence between participants' choices and judgments was substantial and, overall, there was no evidence for decision-making skills being detectable earlier in judgment than choice. In the following sections, these results will be discussed in more detail.

Advantageousness of Choices and Risk Taking

The finding that even 5-year-olds were able to make risk-sensitive choices stands in contrast to Levin et al. (2007), where children of the same age barely adjusted their choices to differences in EV between options. The authors did not report a direct test of whether children's choices were above chance level with respect to advantageousness, but these children chose the risky options most of the time irrespective of EV differences between options. Similarly to our study, Levin et al. found an age-related increase in the advantageousness of choices, with the older children in their study (8-11 years) adjusting their choices to differences in EV, yet not as much as adults. The contrary findings in their and our youngest age group might appear surprising, given that our task was probably more complex than Levin et al.'s, both with respect to the mathematically correct solution and because children in our task had to choose between *two* risky options (whereas Levin et al.'s cups task involved choices between a riskless and a risky option). The two studies differ further with respect to age differences in risk taking. While Levin et al. found increased risk taking in children (see also Harbaugh et al., 2002 and Reyna & Ellis, 1994), we observed virtually identical levels of risk taking across all age groups. There are at least three possible explanations for the differences in Levin et al.'s (2007) and our findings. First, with respect to advantageousness, they used only choices, whereas our task involved judgments *and* choices. Schlottmann and Tring (2005; see also Wilkening & Anderson, 1982) argued that judgments show earlier competence than choices. Thus, it might be that our judgments triggered sophisticated estimation processes which, in turn, led to improved choice performance. Second, Levin et al. gave children the choice between a riskless and a risky option (as did Harbaugh et al., 2002 and Reyna & Ellis, 1994). Perhaps for children the risky option was the more interesting and entertaining one (because it involves an uncertain outcome) and thus they fell prey to an "entertainment bias," leading them to (a) choose the risky option more often than adults, and (b) make more disadvantageous choices (because they neglected the EV of the options). In our task, both options were risky, i.e., involved uncertain outcomes. Thus, they probably differed less in their entertainment value. This explanation is consistent with the fact that we did not observe age differences in choice frequency of the risky option. Third, again in contrast to Levin et al., participants in our task did not receive immediate feedback. This presumably led to an even more uniform entertainment value of the two choice options in our task. In addition, it has been shown that feedback reliably triggers neural activity associated with affective processing (Shohamy et al., 2004). This is relevant because a study comparing adolescents' and adults' risk taking in affect-based versus deliberative risky choice (Figner et al., in press) showed that only *affect-based* decision making led to increased risk taking in adolescents as compared to adults. When predominantly *deliberative* processes were involved (in a task *without* feedback), adults and adolescents did not differ. While a comparison of children and adolescents in this respect might not be so straightforward — the relevant neural substrates undergo substantial changes during adolescence and early adulthood (Casey, Getz, & Galvan, 2008) — it is possible that the affective processes triggered in the cups task led to the observed increased risk taking in children. Because our task triggered presumably less affect-laden reasoning processes, this might have led to uniform levels of risk taking across age groups.

Information Use

We analyzed which variables were taken into account to make the choices and judgments. Analyzed at the group level, we found that all age groups, except 5-year-olds in their choices, took into account both relevant variables, AL and EV. Particularly the 6-year-olds exhibited an astonishing early competence as their judgment and choice patterns came strikingly close to the normative probabilities. This peak in closeness to the mathematical probabilities was followed by a decline in sophistication in the 8- and particularly 10-year-olds and then again improved in the adults. The results were slightly different for the individual-level analysis. Here, the developmental increase in information-use sophistication was closer to a monotonic trend. However, the 10-year-olds still showed a decrease. This developmentally transient deterioration might be indicative of a shift in the predominant processing mode used to solve the task, such that switching to a processing mode not used before to solve such tasks leads to a transient deterioration in performance (Kokis, MacPherson, Toplak, West, & Stanovich, 2002; Reyna & Brainerd, 1995). However, this decline was present only in information use, not in the number of advantageous choices, showing some dissociation of these two aspects of decision making.

That 6-year-olds' judgments were closest to the normative probabilities is consistent with the claim that humans, analogous to many non-human species, might possess an intuition-based decision algorithm similar to the energy-budget rule (Rode et al., 1999). According to this view, 6-year-olds presumably relied on intuitive reasoning to solve the task, leading to good performance, while 8- and 10-year-olds started to rely on analytical strategies, triggered perhaps by experiences in school. Because 8- and 10-year-olds' analytical reasoning about probabilities is not yet mature, this might have led to a temporary deterioration of information use. Alternatively, fuzzy-trace theory (Reyna & Brainerd, 1995) might argue that the improvement in information use after age 8-10 was not necessarily due to improved analytical reasoning but could be due to an improvement of and developmentally increased reliance on intuitive gist-based reasoning. That a transiently increased reliance on analytical processing caused the deteriorated performance in 8- to 10-year-olds would be consistent with Lovett and Singer (1991; as cited in Reyna & Brainerd, 1994), who observed an inverted U-shaped developmental trajectory in using mathematical explicit strategies in solving a probability problem.

Judgment and Choice

The correspondence between children's judgments and choices has not received much attention in the development of risky decision-making. Schlottmann and Tring (2005) found little correspondence between the two response modes, and it has been suggested that, in general, judgments might be able to detect earlier competence than choices (Wilkening & Anderson, 1982). In contrast to this claim, but consistent with Reyna and Ellis (1994), we found substantial correspondence between our judgment and choice measures. However, correspondence was not uniform across the different analyses and age groups. In the group-level analysis, information use was more sophisticated in 5-year-olds' *judgments* than choices. In contrast, information use in 8- and 10-year-olds was more sophisticated in *choice*. In the correspondence analysis, Kendall's τ was highest in our middle age groups (6-, 8-, and 10-year-olds), while it was low in 5-year-olds and adults. Finally, when predicting choices by judgments, 5- and 6-year-olds had the highest percentage of correctly predicted choices. All in all, the results do not unequivocally show that judgments are always more sensitive than choices. Rather it appears that in our study, overall correspondence between judgments and choices was fairly substantial.

Summary and Conclusion

We demonstrated risk sensitivity and advantageousness of risky choice in children as young as 5 years, contrasting studies which used mathematically less complex tasks but found later development (e.g., Levin et al., 2007). A novel contribution was the use of a task that incorporated varying levels of outcome variability (OV), i.e., risk, expected value (EV), and aspiration level (AL) and the choice between *two* risky options. Older tasks using the choice between a risky versus a riskless option can have methodological drawbacks, particularly if combined with immediate feedback as children might not prefer the risky option but rather avoid the "boring" safe option. An interesting topic for future studies would be age-related changes in reliance on different information-processing modes (e.g., intuitive versus analytical) and its consequences on risky decisions.

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. (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. (1996). *A functional theory of cognition*. Hillsdale, NJ: Erlbaum.

Boyer, T. W. (2006). The development of risk-taking: A multi-perspective review. *Developmental Review, 26*, 291-345.

Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review, 28*, 62-77.

Davies, C. M. (1965). Development of the probability concept in children. *Child Development, 36*, 779-788.

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. (in press). Affective and deliberative processes in risky choice: Age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Figner, B., & Voelki, N. (2004). Risky decision making in a computer card game: An information integration experiment. *Polish Psychological Bulletin, 35*, 135-139.

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*, 53-84.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243-266.

Hoemann, H. W., & Ross, B. M. (1971). Children's understanding of probability concepts. *Child Development, 42*, 221-236.

Hommers, W. (1980). Information processing in children's choices among bets. In F. Wilkening, J. Becker & T. Trabasso (Eds.), *Information integration by children* (pp. 99-112). Hillsdale, NJ: Erlbaum.

Huizenga, H. M., Crone, E. A., & Jansen, B. J. (2007). Decision-making in healthy children, adolescents and adults explained by the use of increasingly complex proportional reasoning rules. *Developmental Science, 10*, 814-825.

Kacelnik, A., & Bateson, M. (1997). Risk-sensitivity: Crossroads for theories of decision-making. *Trends in Cognitive Sciences, 1*, 304-109.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-292.

Kokis, J. V., MacPherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83*, 26-52.

Levin, I. P., & Hart, S. 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., Weller, J. A., Pederson, A. A., & Harshman, L. A. (2007). Age-related differences in adaptive decision making: Sensitivity to expected value in risky choice. *Judgment and Decision Making, 2*, 225-233.

Lovett, S. B., & Singer, J. A. (1991, April). *The development of children's understanding of probability: Perceptual and quantitative conceptions.* Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children.* Oxford, England: Norton.

Pietras, C. J., & Hackenberg, T. D. (2001). Risk-sensitive choice in humans as a function of an earnings budget. *Journal of the Experimental Analysis of Behavior, 76*, 1-19.

Reyna, V. F. (1996). Conceptions of memory development with implications for reasoning and decision making. *Annals of Child Development, 12*, 87-118.

Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgment: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 239-272). New York, NY: Wiley.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1-75.

Reyna, V. F., & Ellis, S. C. (1994). Fuzzy-trace theory and framing effects in children's risky decision-making. *Psychological Science, 5*, 275-279.

Rode, C., Cosmides, L., Hell, W., & Tooby, J. (1999). When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. *Cognition, 72*, 269-304.

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

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.

Schlottmann, A., & Wilkening, F. (in press). Early developments in judgment and decision: Probability and expected value, serial processing, heuristics and biases. In M. K. Dhami, A. Schlottmann & M. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development, and evolution*. Cambridge, UK: Cambridge University Press.

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: Converging data from neuroimaging and neuropsychology. *Brain, 127*, 851-859.

Stephens, D. W. (1981). The logic of risk-sensitive foraging preferences. *Animal Behaviour, 29*, 628-629.

Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review, 111*, 430-445.

Wilkening, F., & Anderson, N. H. (1982). Comparison of two rule-assessment methodologies for studying cognitive development and knowledge structure. *Psychological Bulletin, 92*, 215-237.

Table 1

Probabilities of Meeting or Exceeding the Required ALs (in Parentheses: Optimal Choice for Each of the 15 Combinations of EV and AL).

	Low OV Option (Die)					High OV Option (Bag)
	EV 1	EV 2	EV 3	EV 4	EV 5	
AL 1	66.51 (H)	91.22 (L)	98.44 (L)	99.86 (L)	> 99.99 (L)	79.43
AL 3	6.23 (H)	31.96 (H)	65.62 (L)	89.99 (L)	99.13 (L)	56.13
AL 5	0.07 (H)	1.78 (H)	10.94 (H)	35.12 (L)	73.68 (L)	31.66

Note. AL = Aspiration Level; EV = Expected Value. Probabilities are given in percent. Letters in parentheses indicate for each AL/EV combination whether it is advantageous to choose the low OV option (L) or the high OV option (H).

Table 2

Mean and Standard Deviations of Number of Advantageous and Disadvantageous Choices and Risk Taking per Age Group.

	Advantageous Choices			Disadvantageous Choices			Risk Taking
	Die	Bag	Overall	Die	Bag	Overall	Bag
Age Group	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M*	*M*	*M*	*M* (*SD*)
5 Years	12.21 (5.93)	6.54 (4.44)	18.75 (4.42)	5.46	5.79	11.25	12.33 (9.50)
6 Years	14.48 (3.90)	8 (4.06)	22.48 (4.17)	4	3.52	7.52	11.52 (6.78)
8 Years	15.33 (2.40)	9.52 (1.55)	24.85 (2.90)	2.48	2.67	5.15	12.19 (2.81)
10 Years	15.52 (1.90)	9.93 (1.75)	25.45 (2.29)	2.07	2.48	4.55	12.41 (2.85)
Adults	16.08 (1.74)	10.71 (1.57)	26.79 (1.84)	1.29	1.92	3.21	12.63 (2.76)

Note. *SD*s for disadvantageous die choices are identical to *SD*s for advantageous bag choices. *SD*s for disadvantageous bag choices are identical to *SD*s for advantageous die choices. Overall *SD*s for disadvantageous choices are identical to overall *SD*s for advantageous choices.

Table 3

Group-Level Analysis (ANOVA). Effects of AL, EV, and AL × EV in Choice and Judgment Data per Age Group.

Effect	DV	5 Years F (η^2)	6 Years F (η^2)	8 Years F (η^2)	10 Years F (η^2)	Adults F (η^2)
AL	C	1.44 (.06)	15.75*** (.40)	13.61*** (.34)	21.20*** (.43)	25.04*** (.52)
	J	7.19**(.24)	16.12*** (.40)	39.27*** (.60)	29.65*** (.51)	46.52*** (.67)
EV	C	10.98*** (.32)	31.35*** (.57)	125.07*** (.83)	177.73*** (.86)	223.25*** (.91)
	J	21.77*** (.49)	62.32*** (.72)	180.65*** (.87)	160.99*** (.85)	281.16*** (.92)
AL × EV	C	1.59 (.07)	3.51** (.13)	4.23*** (.14)	5.03*** (.15)	10.24*** (.31)
	J	1.78† (.07)	5.24*** (.18)	1.95† (.07)	1.59 (.05)	4.03*** (.15)

Note. †$p < .1$ *$p < .05$ **$p < .01$ ***$p < .001$. AL = Aspiration Level; EV = Expected Value; DV = Dependent Variable; C = Choice data; J = Judgment data.

Table 4

Mean and Standard Deviations of Happiness Ratings for the High OV Option (Bag) per Age Group and Aspiration Level.

	AL 1	AL 2	AL 3
Age Group	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
5 Years	59.58 (36.58)	48.17 (36.39)	35.92 (35.42)
6 Years	42.20 (33.60)	54.28 (33.81)	52.88 (34.65)
8 Years	49.04 (22.12)	47.48 (21.56)	51.26 (16.19)
10 Years	53.45 (13.69)	54.59 (12.79)	48.79 (6.85)
Adults	53.88 (22.20)	47.25 (24.44)	43.46 (27.84)
Normative	79.43	56.13	31.66

Note. Normative = Mathematically normative probabilities in percent.

Table 5

Comparison of Diagnosed Information Use in Choices and Judgments. Median and Mean Levels of Information Use in Choice and Judgment and Association of Categories Across Response Mode.

Age Group	Choice Strategy Median *(M)*	Judgment Strategy Median *(M)*	Strength of Association τ
5 Years	2.5 (2.5)	3 (2.63)	.21
6 Years	3 (3.12)	3 (3.12)	.46**
8 Years	4 (3.44)	4 (3.52)	.83***
10 Years	3 (3.41)	3 (3.41)	.43*
Adults	4 (3.71)	4 (3.88)	.04
Overall	3 (3.25)	3 (3.32)	.51***

Note. *$p < .05$ **$p < .01$ ***$p < .001$. Information use was categorized as 1 = unclassifiable; 2 = constant, non-adaptive responding; 3 = centration; 4 = integration.

Figure Captions

Figure 1. Picture of the board game. AL was operationalized by the number of steps the mouse was away from its safe house (number of black circles in front of the mouse; left display AL = 1; middle display AL = 3; right display AL = 5). EV was operationalized as the number of winning sides of the low OV option given to play the game (number of black circles in the rectangle with 6 circles depicted on the left of the mouse; left display EV = 3; middle display EV = 5; right display EV = 2).

Figure 2. Information use in participants' judgments (top graph) and choices (bottom graph): Effects of AL (x axis) and EV (curves numbered 1 to 5) on happiness ratings for low OV option (top graph) and proportion of choice of low OV option (bottom graph), displayed for each age group separately. Within both graphs, bottom right panels show the normative solution. For the normative judgment solution, probabilities are shown in percent. For the normative choice solution, values are either 0 or 1; i.e., jitter has been added to avoid laying lines on top of one another.

Figure 3. Judgment strategies (top graph) and choice strategies (bottom graph), derived from individual-level analyses. Non-classifiable random responses are not shown as a separate category, but can be inferred from bars not reaching 100%.

Figure 1

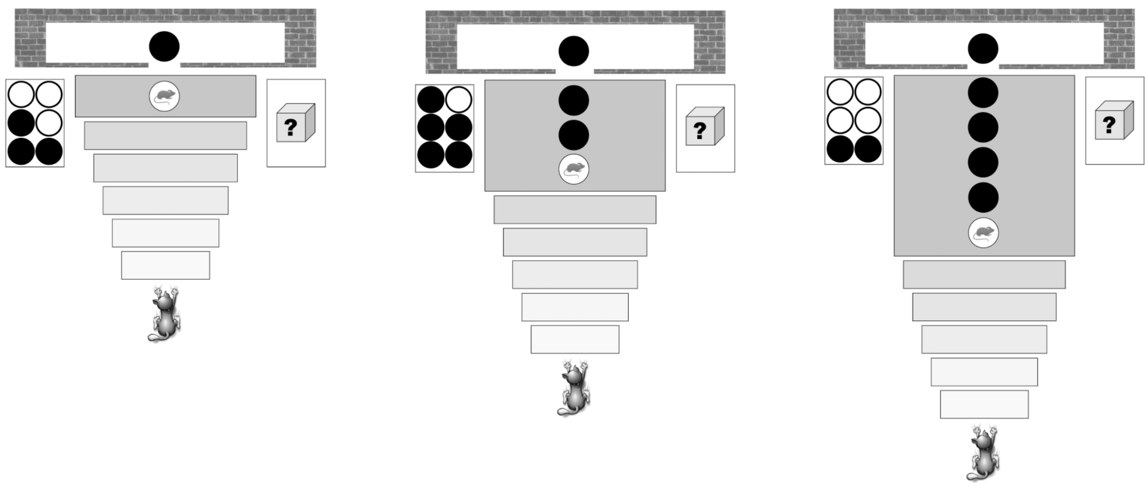


Figure 2

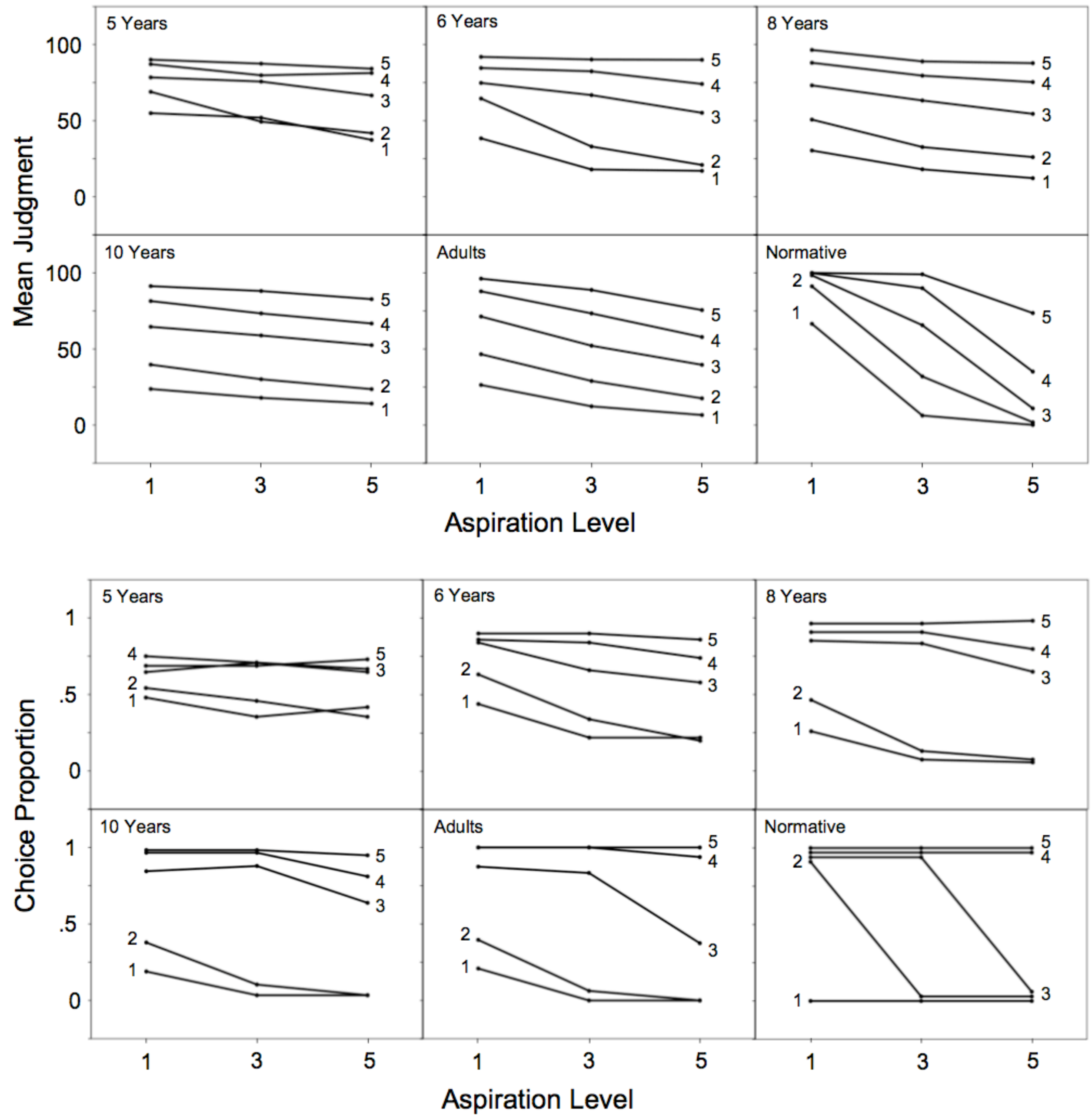
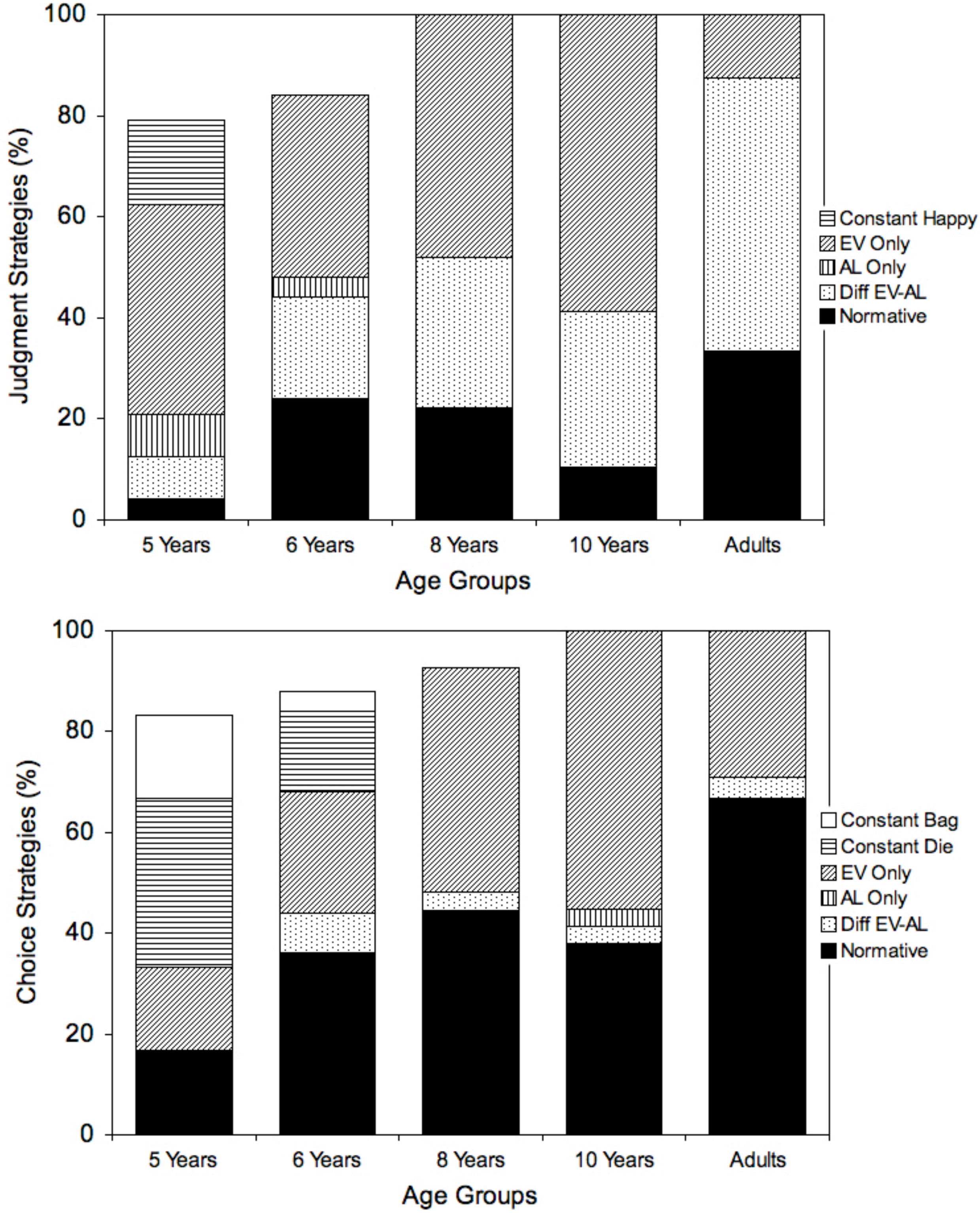

Mean Judgment
100
50
0
5 Years
6 Years
8 Years
10 Years
Adults
Normative
1
3
5
Aspiration Level
Choice Proportion
1
.5
0
5 Years
6 Years
8 Years
10 Years
Adults
Normative
1
3
5
Aspiration Level

Figure 3



100
80
60
40
20
0
Judgment Strategies (%)
5 Years
6 Years
8 Years
10 Years
Adults
Age Groups
Constant Happy
EV Only
AL Only
Diff EV-AL
Normative
100
80
60
40
20
0
Choice Strategies (%)
5 Years
6 Years
8 Years
10 Years
Adults
Age Groups
Constant Bag
Constant Die
EV Only
AL Only
Diff EV-AL
Normative