Children’s Probabilistic Judgment and Decision Making

Dissertation

zur Erlangung des akademischen Grades
Doktorin der Philosophie (Dr. phil)

der
Erziehungswissenschaftlichen Fakultät
der Universität Erfurt

vorgelegt von
Anna Lang

Erfurt 2018

Erstes Gutachten:	Prof. Dr. Tilmann Betsch (Universität Erfurt)
Zweites Gutachten:	Prof. Dr. Claudia Steinbrink (Universität Erfurt)
Drittes Gutachten:	Prof. Dr. Mirjam Ebersbach (Universität Kassel)
Tag der Disputation:	26.10.2018
Datum der Promotion:	26.10.2018
urn:	nbn:de:gbv:547-201800630

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich Unterstützungsleistung von folgenden Personen erhalten:

Prof. Dr. Tilmann Betsch

Weitere Personen waren an der geistigen Herstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin bzw. eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit oder Teile davon wurden bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde als Dissertation vorgelegt. Ferner erkläre ich, dass ich nicht bereits eine gleichartige Doktorprüfung an einer Hochschule endgültig nicht bestanden habe.

Zusammenfassung

Das übergreifende Thema dieser Dissertation ist die Frage nach der Entscheidungskompetenz von Kindern. Da bei vielen Entscheidungen Wahrscheinlichkeiten eine Rolle spielen, habe ich in einer Reihe von Studien untersucht, ob und wann Kinder probabilistische Inferenzentscheidungen bewältigen. Dies sind Entscheidungen bei denen Hinweisreize aus der Umwelt, sogenannten *Cues*, genutzt werden können, um zukünftige Entscheidungskonsequenzen vorherzusagen. Diese Konsequenzen treten allerdings nicht mit Sicherheit ein, sondern nur mit einer gewissen Wahrscheinlichkeit, die in Abhängigkeit von der Stärke des stochastischen Zusammenhangs zwischen Cue und Entscheidungskonsequenz bestimmt werden kann. Diese Konstellation ahmt reale Entscheidungssituationen nach, in denen Hinweisreize genutzt werden, um Entscheidungskonsequenzen zu antizipieren. Zum Beispiel, nutzen wir nonverbale Signale wie Mimik und Körpersprache um vorherzusehen, ob andere Menschen freundlich auf Kontaktaufnahme reagieren werden. Oder wir nutzen Merkmale eines Produktes, wie Preis, Marke und Verarbeitungsqualität, um dessen Langlebigkeit abzusehen.

Bislang war nicht bekannt, ob und wie Kinder im Vor- und Grundschulalter mit diesen Entscheidungen umgehen können. Erste Forschung mit 6- und 9-jährigen zeigte, dass Kinder die probabilistischen Zusammenhänge zwischen Cues und Entscheidungskonsequenzen nicht nutzen, um Entscheidungen zu optimieren (Betsch & Lang, 2013; Betsch et al., 2014; siehe auch Mata, van Helversen, & Rieskamp, 2011 für Kinder ab 9 Jahren). Während 6-Jährige Wahrscheinlichkeiten gar nicht berücksichtigten, nutzten 9-Jährige sie nur teilweise. Dies ist gemäß dem Stufenmodel der kognitiven Entwicklung nach Piaget zu erwarten, da bis zum Alter von 7-8 Jahren kein Verständnis von Zufall oder Wahrscheinlichkeit und bis zum Alter von 11-12 Jahren nur ein eingeschränktes Verständnis angenommen wird (Piaget & Inhelder, 1975). Weitere Forschung zeigte allerdings über viele Situationen hinweg, dass Kinder durchaus in der Lage sind Wahrscheinlichkeiten zu nutzen, um ihr Verhalten zu optimieren (z.B. Denison, Bonawitz, Gopnik, & Griffiths, 2010; Gonzalez & Girotto, 2011; Pasquini, Corriveau, Koenig, & Harris, 2007; Schlottmann, 2001). Dies führte zu der Annahme, dass Kinder bereits sehr früh in der Lage sind, Wahrscheinlichkeiten auf intuitive Weise zu nutzen (z.B. Schlottmann & Wilkening, 2012). Die Vernachlässigung von Wahrscheinlichkeiten bei probabilistischen Inferenzentscheidungen steht somit im Widerspruch zu dieser Auffassung und der sie stützenden Befunde.

Die Studien der Dissertation ergründen deshalb, warum Kinder Wahrscheinlichkeiten vernachlässigen und fokussieren dabei auf die Rolle von Feedback. In Entscheidungen mit

Feedback erlebt der Entscheider die Konsequenz seiner Wahl unmittelbar. Dies kann sich potentiell negativ auf die Nutzung von Wahrscheinlichkeiten auswirken, wenn Feedback eine interferierende Informationsquelle darstellt und von Wahrscheinlichkeiten ablenkt (z.B. Harvey, 2011). Es kann sich aber auch positiv auswirken, wenn es die Salienz des probabilistischen Zusammenhangs zwischen Cue und Entscheidungskonsequenzen erhöht und ermöglicht, angemessenes Entscheidungsverhalten zu lernen (z.B. Rieskamp & Otto, 2006).

Artikel 1: Vernachlässigung von Wahrscheinlichkeiten in kindlichen Entscheidungen mit und ohne Feedback.

Der erste Artikel prüft, ob Kinder Feedback im Vergleich zu expliziter Wahrscheinlichkeitsinformation bevorzugen. Die Darbietung von Feedback könnte somit Abweichung von normativen Entscheidungsmodellen, die eine systematische Nutzung von Wahrscheinlichkeiten verlangen, erklären.

6-Jährige, 9-Jährige und Erwachsene trafen eine Reihe von Entscheidungen, bei denen Vorhersagen probabilistischer Cues genutzt werden konnten, um Entscheidungskonsequenzen zu optimieren. Die Verfügbarkeit von Feedback wurde variiert: Kinder wurden entweder nach jeder Entscheidung sofort über die positive oder negative Konsequenz der getroffenen Entscheidung informiert, oder erfuhren dies erst nachdem alle Entscheidungen getroffen waren.

Die Ergebnisse zeigen, dass jüngere Kinder besonders sensitiv auf negative Entscheidungskonsequenzen reagieren und ihre Entscheidungen somit von unmittelbar erlebten Feedback verzerrt sind. Allerdings nutzten Kinder Feedback nicht auf systematische Weise, um Entscheidungen zu treffen. Unabhängig von der Verfügbarkeit von Feedback, vernachlässigten alle jüngeren und ein Großteil der älteren Kinder Wahrscheinlichkeiten bei ihren Entscheidungen. Wenn überhaupt Entscheidungsstrategien von Kindern angewendet wurden, dann stützten sich diese häufig auf irrelevante Informationen und führten nicht zu optimalen Entscheidungen.

Dies zeigt, dass die Vernachlässigung von Wahrscheinlichkeit robust ist, unabhängig von Feedback stattfindet und erst ab dem Schulalter abnimmt.

Artikel 2: Nutzung und Überwindung nicht-adaptiver Entscheidungsstrategien bei Kindern.

Artikel 2 erfasst adaptive und nicht-adaptive Entscheidungsstrategien bei 6-Jährigen, 9-Jährigen und Erwachsenen und untersucht wie häufig, wie präzise und wie beharrlich sie genutzt werden. Bislang wurde angenommen, dass Kinder unter 9 Jahren eher zufällig entscheiden, ohne jede Systematik. Allerdings lag der Fokus der Forschung auch darauf adaptive, wahrscheinlichkeitsbasierte Entscheidungsstrategie aufzuspüren. Nicht gesucht wurde nach Strategien, die normativ irrelevante Informationen nutzen und somit auch nicht-adaptiv sind und zu suboptimalen Entscheidungen führen.

Die Studie untersuchte deshalb Entscheidungsstrategien in unterschiedlichen Feedbackumwelten unter Berücksichtigung nicht-adaptiver Strategien. Die Ergebnisse zeigen, dass ungefähr die Hälfte aller Kinder Entscheidungsstrategien konsistent nutzt, wenn nicht-adaptive Strategien berücksichtigt werden. 6-Jährige nutzten fast ausschließlich nicht-adaptive Strategien, 9-Jährigen sowohl adaptive als auch nicht-adaptive Strategien. Wenn Kinder umfassend über alle Entscheidungskonsequenzen informiert wurden, also über Konsequenzen der gewählten und nicht gewählten Option, wichen sie zunehmend von ihren nicht-adaptiven Strategien ab.

Kinder sind also schon im Vorschulalter in der Lage, Entscheidungenstrategien systematisch anzuwenden. Analog zur Strategieentwicklung in anderen Bereichen, nutzen sie allerdings zunächst ineffiziente Entscheidungsstrategien, die sie mit zunehmenden Alter und Erfahrung ablegen.

Artikel 3: Lernen Kinder einfache, aber adaptive Entscheidungsstrategien durch Feedback?

Artikel 3 untersucht inwiefern Feedback adaptive Strategien bei Kindern fördern kann. In der ersten Studie, lernten 7- und 9-Jährige Kinder eine simple, wahrscheinlichkeitsbasierte Entscheidungsstrategie. Sie lernten sie weniger gut als eine erwachsene Vergleichsgruppe und besser, wenn konkurrierende Strategien häufiger zu Misserfolgen führten.

In der zweiten Studie verstärkte Entscheidungsfeedback entweder dieselbe simple, wahrscheinlichkeitsbasierte Strategie oder eine ebenso einfache Strategie, die aber nicht auf Wahrscheinlichkeiten basierte. Jünger Kinder lernten letztere besser, ältere Kinder lernten beide Strategien gleich gut. In beiden Studien nutzten Kinder adaptive Entscheidungs-strategien nur selten, wenn kein Feedback verfügbar war, obwohl deren Anwendung durch die Dispersion der verfügbaren Wahrscheinlichkeiten indiziert war.

Unter idealen Umständen, einer eindeutigen Feedbackumwelt und wenn einfache Entscheidungsstrategien verstärkt werden, kann Feedback also adaptive Entscheidungsstrategien fördern und somit die Entscheidungen von Kindern erheblich verbessern.

Artikel 4: Nutzung probabilistischer Cues in Urteilen von Kindern

Artikel 4 untersucht die Nutzung der probabilistischen Zusammenhänge von Cues und Entscheidungskonsequenten in Urteilen von 6- und 9-jährigen Kindern und Erwachsenen. Die Ergebnisse von vier Studien zeigen, dass die Nutzung dieser Wahrscheinlichkeitsinformation in Urteilen vergleichsweise spät beginnt und demselben Entwicklungsverlauf folgt wie in Entscheidungen: 6-Jährigen berücksichtigen Wahrscheinlichkeiten nicht, 9-Jährige berücksichtigen sie nur teilweise. Die Urteile von Kinder zeigen, dass ihre Erwartungen an Entscheidungskonsequenzen bis zum Alter von 9 Jahren nicht an Wahrscheinlichkeiten orientiert sind.

Die Ergebnisse sind über mehrere Studien hinweg und unabhängig von Erfahrungen der Teilnehmer oder Skalen zur Erfassung der Urteile stabil. Die Befunde widersprechen einer frühen und intuitiven Nutzung von Wahrscheinlichkeiten bei Urteilen und zeigen persistierende Defizite in wahrscheinlichkeitsbasierten Urteilen von Kindern bis zu 9 Jahren.

Literaturverzeichnis

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology*, *115*, 108–125. https://doi.org/10.1016/j.jecp.2012.11.003

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica*, *153*, 74–86. https://doi.org/10.1016/j.actpsy.2014.09.008

Denison, S., Bonawitz, E. B., Gopnik, A., & Griffiths, T. L. (2010). Preschoolers sample from probability distributions. In R. Camtrabone & S. Ohlsson (Eds.), *Proceedings of the 32nd annual conference of the Cognitive Science Society* (pp. 2272-2277). Austin, TX: Cognitive Science Society.

Harvey, N. (2012). Learning judgment and decision making from feedback. In M. K. Dhami, A. Schlottmann, & M. R. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development and evolution* (pp. 199–226). Cambridge: University Press.

Mata, R., von Helversen, B., & Rieskamp, J. (2011). When easy comes hard: The development of adaptive strategy selection. *Child Development*, *82*(2), 687–700. https://doi.org/10.1111/j.1467-8624.2010.01535.x

Girotto, V., & Gonzalez, M. (2008). Children's understanding of posterior probability. *Cognition, 106*, 325–344. https://doi.org/10.1016/j.cognition.2007.02.005

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, *43*(5), 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. London: Routledge & Kegan Paul.

Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *Child Development, 72*, 103–122. https://doi.org/10.1111/1467-8624.00268

Schlottmann, A., & Wilkening, F. (2012). Judgment and decision making in young children. In M. K. Dhami, A. Schlottmann, & M. R. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development and evolution* (pp. 55–83). Cambridge: University Press.

Rieskamp, J., & Otto, P.E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*(2), 207–236. https://doi.org/10.1037/0096-3445.135.2.207

Contents

Overview

Over the last years, I investigated the development of decision making in children. This research resulted in several articles and one book chapter on the topic. It was conducted as part of a research program on children's decision making at the Department for Social, Organisational and Economic Psychology supervised by Prof. Tilmann Betsch.

I thank Prof. Tilmann Betsch for his support and encouragement, my co-workers Anne Lehmann, Dr. Stefanie Lindow, and Dr. Johannes Ritter for critical discussions on the topic, and Dr. Niels Haase for very helpful comments on several of the articles.

The research was supported by a grant of the Deutsche Forschungsgemeinschaft to Tilmann Betsch (Grant No. BE 2012/11-1) and several grants for young researchers from the Faculty of Education at the University of Erfurt.

Under review:

Lang, A. (2018). Utilization of probabilistic cues in children's judgments. *Paper under review.*

Lang, A., & Betsch, T. (2018). Children learn simple, adaptive decision strategies from probabilistic feedback. *Paper under review.*

Lang, A., & Betsch, T. (2018). Children use but overcome non-adaptive decision strategies. *Paper under review.*

Published in peer-reviewed journals:

Lang, A. & Betsch, T. (2018) Children's neglect of probabilities in decision making with and without feedback. *Frontiers in Psychology, 9*, 191. https://doi.org/10.3389/fpsyg.2018.00191.

Lindow, S., **Lang, A.**, & Betsch, T. (2017). Holistic information integration in child decision making. *Journal of Behavioral Decision Making, 30*, 1131–1146. https://doi.org/10.1002/bdm.2029.

Betsch, T., Lehmann, A., Lindow, S., **Lang, A.**, & Schoemann, M. (2016). Lost in search: (Mal-)adaptation to probabilistic decision environments in children and adults. *Developmental Psychology, 52* (2). 311–325. https://doi.org/10.1037/dev0000077

Betsch, T., **Lang, A.**, Lehmann, A., & Axmann, J.M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica, 153,* 74–86. https://doi.org/10.1016/j.actpsy.2014.09.008

Betsch, T., & **Lang, A.** (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. J*ournal of Experimental Child Psychology, 115*, 108–125. https://doi.org/10.1016/j.jecp.2012.11.003

Book chapters:

Betsch, T., Ritter, J., **Lang, A.**, & Lindow, S. (2016). Thinking beyond boundaries. In L. Macchi, M. Bagassi, & R. Viale, (Eds.). *Cognitive unconscious and human rationality*. Cambridge, MA: MIT Press.

Four of these articles contribute to this dissertation. All of which investigate the influence of decision feedback on children's probabilistic judgment and decision making. Altogether, they entail six empirical studies. Article 1 has been published in an international, peer-reviewed, academic journal. Articles 2-4 are currently reviewed.

Research Highlights

Article 1

- 6-year-olds, but not 9-year-olds' decisions were biased by, but did not systematically follow recent negative outcomes.
- 6-year-olds fully and 9-year-olds partly neglected probabilistic information in decisions with and without feedback.
- Neglect of probabilities is robust and independent of feedback.

Article 2

- Half of 6-year-olds and 9-year-olds applied decision strategies systematically.
- 6-year-olds only used non-adaptive strategies focusing on irrelevant information.
- 9-year-olds used adaptive as well as non-adaptive decision strategies.
- Children abandoned non-adaptive strategies when given full feedback about decision outcomes.

Article 3

- Feedback that reinforces simple decision strategies increases adaptive strategy use in children.
- Strict feedback benefits children slightly more than lenient feedback.
- 7-year-olds learn an Equal Weight-strategy better than a Lexicographic strategy; 9-year-olds learn both equally well.

Article 4

- Probability utilization for judgments is absent in 6-year-olds and emerging in 9-year-olds.
- Judgments remain unaffected by participant's experience with decisions, decision feedback, or scale formats used to assess judgments.
- Deficits in probabilistic judgments persist until the age of nine and follows the same developmental trajectory as in decisions.

Scope

Data collection. The dissertation presents six studies with more than 1.200 child and adult participants (Table 1). Each participant was supervised by a trained experimenter during the whole procedure. Child data were collected either in separate rooms of kindergartens and after-school day-care centers or in the lab. For lab studies, groups of children were invited to participate in “Science Workshops” at the University of Erfurt. For that purpose, we developed several workshops for different age groups. All workshops were designed to promote children’s understanding of science and of experimental methods.

Parents and teachers were informed in advance about the procedure and purpose of the data collection and consented to participation. Child participants were rewarded with two to five prizes contingent on their performance. In the lab, children were debriefed according to their age. In schools, we informed children, parents, and teachers about the studies’ purpose through posters or flyers.

Adult participants were students of the University of Erfurt, consented to participation, and were debriefed in accordance with the guidelines of the Hermann-Ebbinghaus-Laboratory at the University of Erfurt. They were rewarded with money contingent on performance. All procedures were approved by the Ethics Committee at the University of Erfurt.

Table 1

Overview of sample sizes for Study 1-6

Study	**Child Data**	**Pre-schoolers 5-6 years**	**First Graders 6-7 years**	**Third Graders 8-10 years**	**Adult Controls 20-22 years**	**Analyzed *N***
1	Lab	69	—	56	56	181
2	Lab	80	—	62	53	195
3	Day-Care	81	—	91	85	257
4	Lab	—	110	107	99	316
5	Lab	—	92	96	77	265
6	Day-Care	42	—	41	—	83
Total		272	202	453	370	1.297

Selection of age groups. In each study, we compared two child groups, children who attended preschool or grade 1 (5-7 years) and children at middle elementary school age (8-9 years). This age comparison is especially illuminating when investigating the development of probabilistic decision making. Piaget's stage theory claims that children's idea of chance starts to develop at age six to eight, but remains impaired until late elementary school age (Piaget & Inhelder, 1975). In contrast, current research suggests intuitive probability utilization at preschool age (e.g., Schlottmann & Wilkening, 2012). Accordingly, we concentrated in all studies on the interesting period between preschool and elementary school age.

Research questions & experimental designs. All articles investigate how decision feedback affects children's utilization of probabilistic information. While Article 1 focuses on detrimental effects of feedback on decisions, Article 2 and 3 investigate the beneficial effects of feedback on decisions and decisions strategies. Article 4 investigates utilization of probabilistic information in children's judgments and the moderating role of feedback in this process.

Table 2

Overview of experimental factors

Article & Studies	Age Factor	Additional Factor
Article 1 including decision data from Studies 1 and 2	3 (Age Group: 6-year-olds vs. 9-year-olds vs. Adults)	2 (Feedback: No feedback vs. Feedback)
Article 2 including decision data from Study 3	3 (Age Group: 6-year-olds vs. 9-year-olds vs. Adults)	3 (Feedback: No feedback vs. Selective Feedback vs. Full Feedback)
Article 3 including decision data from Studies 4 and 5	3 (Age Group:7-year-olds vs. 9-year-olds vs. Adults)	3 (Feedback: No feedback vs. Lenient Feedback vs. Strict Feedback) & 3 (Feedback: No feedback vs. Feedback for LEX vs. Feedback for Equal Weight)
Article 4 including judgment data from Studies 1-3 and Study 6	3 (Age Group: 7-year-olds vs. 9-year-olds vs. Adults)	2 (Scale Format: Intuitive vs. Analytical)

Author Contributions

Article 1 Lang, A., & Betsch, T. (2018). Children's neglect of probabilities in decision making with and without feedback. *Frontiers in Psychology, 9*, 191. https://doi.org/10.3389/fpsyg.2018.00191

Conceived and designed the experiments: AL, TB

Performed the experiment & analyzed the data: AL

Wrote the paper: AL

Approved the final draft: AL, TB

Article 2 Lang, A., & Betsch, T. (2018). Children use but overcome non-adaptive decision strategies. *Paper under review.*

Conceived and designed the experiment: AL, TB

Performed the experiment & analyzed the data: AL

Wrote the paper: AL

Approved the final draft: AL, TB

Article 3 Lang, A., & Betsch, T. (2018). Children learn simple, adaptive decision strategies from probabilistic feedback. *Paper under review.*

Conceived and designed the experiments: AL, TB

Performed the experiment & analyzed the data: AL

Wrote the paper: AL

Approved the final draft: AL, TB

Article 4 Lang, A. (2018). Utilization of probabilistic cues in children's judgments. *Paper under review.*

References

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. London: Routledge & Kegan Paul.

Schlottmann, A., & Wilkening, F. (2012). Judgment and decision making in young children. In M. K. Dhami, A. Schlottmann & M. R. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development and evolution* (pp. 55–83). Cambridge: University Press.

General Introduction

As there a few sure things in life, most decision situations entail a fair amount of chance. Consequences of decision seldom occur for sure, but are more or less probable. If we are lucky, we have some notion of how probable outcomes are. For example, when deciding which of our friends to ask for help, past experience tells us who will be most likely to help out. When we decide, what medicine to take for a sore throat, probabilities of negative outcomes, such as side effects, are stated in the package insert. Since probabilities are so crucial, competent decision making inevitably includes utilizing probability in judgments and decisions (Bruine de Bruin, Parker, & Fischhoff, 2007; Finucane & Gullion, 2010; Finucane, Slovic, Hibbard, Peters, Mertz, & MacGregor, 2002; Parker & Fischhoff, 2005; Stanovich & West, 1998; Weller, Levin, Rose, & Bossard, 2011).

If we want to determine whether and when children are competent in making their own decisions, we must therefore ask at what age children utilize probability in judgment and decision making. From an educational perspective, this begs the question, whether and when we should teach children to appropriately utilize probabilities (Bryant & Nunes, 2012); especially, since without formal education, some misconceptions about probability persist into adulthood (e.g., Tversky & Kahnemann, 1974).

Probabilistic Decision Making in Children

Do children utilize probabilities in their decisions or judgments? According to Piaget & Inhelder (1975), children's understanding of chance and probability is limited. Their stage model of cognitive development claims that children have no notion of probability until reaching the concrete-operational stage (7-8 years), and cannot fully consider probability in a normative fashion until reaching the formal-operational stage (11-12 years). While early research partially supported this view (see Reyna & Brainerd, 1994, for an overview), manifold research has shown that preschool aged children—around the age of five—adapt behavior in line with probability over a wide range of situations (Acredolo, O'Connor, Banks, & Horobin, 1989; Boyer, 2007; Brainerd, 1981; Davies, 1965; Denison, Bonawitz, Gopnik, & Griffiths, 2010; Girotto, Fontanari, Gonzalez, Vallortigara, & Blaye, 2016; Girotto & Gonzalez, 2008; Goldberg, 1966; Gonzalez & Girotto, 2011; Hoemann & Ross, l992; Pasquini, Corriveau, Koenig, & Harris, 2007; Reyna & Brainerd, 1994; Zhu & Gigerenzer, 2006; Zmyj, Buttelmann, Carpenter, & Daum, 2010). Moreover, even infants seem to base their expectations on probability (Denison, Reed, & Xu, 2013; Denison & Xu, 2010; Teglas, Girotto, Gonzalez, & Bonatti, 2007; Teglas, Vul, Girotto, Gonzalez, Tenenbaum, & Bonatti, 2011; Xu & Denison, 2009; Xu & Garcia, 2008). This supports the view that intuitions about probability emerge early in life and allow even very young children to utilize probability in judgments and decisions (Anderson & Schlottman, 1991; Dension & Xu, 2014; Reyna & Brainerd, 1994; Schlottmann, 2001, Schlottmann & Wilkening, 2012). Empirical findings remain nevertheless heterogeneous: At the age of five, children can for example estimate posterior probability (Girotto & Gonzalez, 2001) but cannot use probabilistic relations to predict distributions of attributes (Kalish, 2010); they can judge expected values of gambles (Schlottmann, 2001) but not always choose successfully between them (Girotto, et al., 2016; Levin & Hart, 2003; Levin, Hart, Weller, & Harshman, 2007; Levin, Weller, et al. 2007).

In line with this mixed findings, dual-process models assume that two distinct systems are involved in judgment and decision making (Glöckner & Betsch, 2008; Hogarth, 2001, 2005; Kahneman, 2003; Kahneman & Frederick, 2002; Stanovich & West, 2000; Schneider & Shiffrin, 1977). The first system—described as intuitive, automatic, or experiential—handles decision making intuitively, which allows effortless and quick decisions (see Glöckner & Wittmann, 2010, and Evans, 2008, for more detailed discussions of dual-process theories). This intuitive system is available early, potentially since birth (Jacobs & Klaczynski, 2002; Klaczynski, 2005; Schlottmann & Wilkening, 2012). The second system—described as controlled, deliberate, or analytical—handles decisions in an analytical, effortful and slower fashion and is thought to develop slowly with age (Schlottmann & Wilkening, 2012; Stanovich, Toplak, & West, 2008; cf. Reyna & Brainerd, 2011). Children might utilize probability in judgment and decisions if they can do so intuitively, but may not be able to do so if the analytical system is required. Accordingly, the question is not if children utilize probability but when they do so intuitively or analytically.

Probabilistic Inference Decisions

In this thesis, I investigate children's judgment and decision making in probabilistic inference situations. Many real-world decisions allow to predict future consequences probabilistically by currently available cues. For example, in social decisions, when reactions of social partners can be predicted by present behavior such as mimics or body language (Brunswick, 1955); or in consumer decisions, where the durability of a product, such as a smartphone, is predicted by its prize and manufacturer. This structure is mimicked in probabilistic inferences decisions: Available cues predict outcomes of options probabilistically, and this probability varies, some cues predict outcomes better than others. The probabilistic cue-outcome relation—the cue validity—can either be available as summarized description provided by others (*This cue predicts 70% of cases correctly*; e.g., Bröder, Glöckner, Betsch, & Link, 2013) or can be learned gradually by personal observing cue-outcome contingencies (e.g., Lagnado, Newell, Kahan, & Shanks, 2006). Cue validities and predicted values can then be integrated into a judgment for each alternative and lastly allow a decision between choice alternatives. All probabilistic inference decisions thus require to utilize probabilistic cues to predict a criterion. Still, they vary on a multitude of facets, such as the type of the criterion variable (continuous vs. categorical), the relation between criterion variable and cues (linear vs. non-linear; same vs. different variable), the display of options (option as cue-compounds vs. as distinct entities) and the framing of the task (city size task vs. weather prediction task; e.g., Gigerenzer & Goldstein, 1991; Harvey & Fischer, 1997; Lagnado et al., 2006).

How do adults master such decisions? The research has produced a variety of theories. Cue Abstraction models assume that adults form mental representations of each cue validity (Einhorn, Kleinmuntz, & Kleinmuntz, 1979). These representations can either be informed by stated cue validities and or be learned from observing cue-outcome contingencies as long as they are not overly complex (e.g., Brehmer, 1980; Harvey & Fisher, 1997; Lagnado et al., 2006, see Karealaia & Hogarth, 2008, and Harvey & Fischer, 2005, for overviews). Normative decision theory, its subjective variants, and network models then assume that decision makers integrate cue validities and values using a universal mechanism: they weight all cue values with their (subjective) validity and add them up to an expected value (Edwards, 1954; Tversky & Kahnemann, 1981; see Glöckner & Betsch, 2008, for an automatic integration in a network model).

In contrast, adaptive decision models assume that not all information is integrated, but that decision makers apply simple strategies using different integration mechanisms. For example, considering values of the most valid cue and ignoring others (Gigerenzer, Todd, & the ABC Research Group, 1999; Payne et al., 1988).

Exemplar-models take a completely different approach (Juslin, Jones, Olsson, & Winman, 2003): They neglect a mental representation and rule-based integration of cue validities. Instead, they assume that decision makers encode instances of cue value patterns and outcomes, store them, and retrieve them during judgments or decision making as a function of similarity between previous and current cue patterns.

Empirical findings are inconclusive on the superiority of any theoretical approach. Potentially, adult decision making rely on all mechanisms depending on environmental factors, such as task complexity and framing (Bröder, Newell, & Platzer, 2010; Karlsson, Juslin, & Olsson, 2007).

Probabilistic Inference Decisions in Children

Children also face probabilistic inference decisions. Imagine for example, when the child wishes to maximize suspense of the bedtime story and can choose which parent will tell it. The child knows that one of the parents—for example, the mother—is more talented in telling stories. Accordingly, suspense is more likely to be maximized in her story. However, tiredness is another predictor, which is negatively correlated with suspense of the story. Since the mother seems to be more tired than the father, both observable cues point to different options. Tiredness may be a more valid predictor for an exciting bed time story than general talent of the storyteller and therefore this cue's prediction should receive more weight in the decision.

Research on children's probabilistic inferences shows that children until the age of nine fail to consider differences in cue validities when making these decisions: At the age of five, all children, at the age of nine most children do not utilize cue validities in their decisions, and even older children's decisions diverge from adult controls (Betsch & Lang, 2013; Betsch, Lang, Lehmann, & Axmann, 2014; Mata, van Helversen, & Rieskamp, 2011). However, reasons for this probability neglect are still unknown, and are further investigated in this thesis.

Feedback in Decision Making

In most real-world decisions, we experience some consequences when making a decision. The feedback we receive might not be immediate or of urgent importance, but can affect future decisions. Despite this conceivable notion, the role of feedback had initially been neglected in judgment and decision making theory, due to its origin in economic theories of rational choice. These theories were occupied with fully described one-shot decisions, that is, probabilities were provided in a summarized fashion and the decision was made only once, not recurrently. Expected value theory and its subjective variants claimed that decision makers integrate (subjective) stated probability and (subjective) stated value to an (subjective) expected value in a normative fashion ($EV = p \times v$ *and* $EU = p \times u$; Edwards, 1954; Kahnemann & Tversky, 1979; Savage, 1954; von Neumann & Morgenstern, 1947; see Machina, 1987, for an overview).

Feedback was thus theoretically not incorporated until the idea of adaptive decision making (Gigerenzer et al., 1999; Payne, et al., 1988). It postulates a multitude of decision strategies apart from normative probability-value-integration and thus requires a mechanism of strategy selection. Decision feedback can function as such a mechanism: It informs the decision maker how accurately a strategy performs and allows to adaptively select strategies depending on their performance (Beach and Mitchel, 1978; Gigerenzer et al., 1999; Payne et al., 1988; Rieskamp, 2008; Rieskamp & Otto, 2006).

In addition, research focused on the origin of decision-making routines (Betsch & Haberstroh, 2005), on feedback processing in pure experience-based decision (that is, when feedback is the only source of information, for example in the IOWA gambling task; Bechara, Damasio, Damasio, & Anderson, 1994), and on differences in decisions based on stated probabilities compared with probabilities learned from feedback (description-experience gap; Barron & Erev, 2003; Erev et al., 2010; Erev & Barron, 2005; Hertwig, Barron, Weber, & Erev, 2004).

The direction and size of feedback's effect on judgment and decision making is theoretically and empirically non-uniform. Feedback can either be helpful or harmful depending on its properties, such as its complexity, timing, and format and on its functions (Bangerts-Drowns, Kulik, Kulil, & Morgan, 1991; Fyfe & Rittle-Johnson, 2015; Hattie & Timperley, 2007; Karelaia & Hogarth, 1998; Kluger & deNisi, 1996; Narciss & Huth, 2004; Shute, 2008). Generally, feedback can fulfill at least three distinct functions in decision making (but see the aforementioned references for more detailed analysis):

First, feedback enables to learn from trial and error which behavioral response—which choice, for example—produces reward and punishment. In line with classical conditioning (Pavlow, 1926), Thorndike's law of effect (1911), and more specified reinforcement learning models (Sutton & Barto, 1998), reinforced behavioral actions are repeated more frequently over time, and individuals learn, for example, to choose options with larger or more frequent rewards. This learning process is thought to work automatically, explains decision making in simple, uninformed decisions in humans and other animals (see Daw & Tobler, 2014; Dayan & Niv, 2008; for an overview, see Worthy & Maddox, 2014), and improves in humans until adulthood (Cassoti, Aïte, Osmont, Houdé, & Borst, 2014; van den Bos, Güroğlu, van den Bulk, Rombouts, Crone, 2009).

Second, feedback provides the decision maker with information and allows to build up a valid mental model of the environment (Byrnes, Miller & Reynolds, 1999; Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Fyfe & Rittle-Johnson, 2015). The informational value of feedback—it can for example be representative of the feedback distribution or biased (Fiedler, 2008)—and an individual's prior knowledge, beliefs, or feedback attribution affects what can be learned. Thus, the interaction between decision maker and environment determines the match of the mental model and reality, and whether feedback improves judgments or decisions.

Third, feedback can affect motivation, perceptions of self-efficacy and can increase or decrease the cognitive effort individuals are willing to invest (Bandura, 1977; Deci & Ryan, 1985; Weiner, 1974). This illustrates that effects of feedback on decision making are inevitably complex and must account for properties of the decision environment and decision maker.

Feedback in Children's Probabilistic Decision Making

Learning from outcomes is nevertheless a powerful tool to improve judgment and decision making. It is thus an important factor from a developmental perspective, because it might help to overcome deficits, boost development, or even eradicate age-dependent differences in performance.

In children the role of feedback in probabilistic decisions has mainly been investigated in variants of gambling tasks: Children of different ages (the youngest at age three, Bunch, Andrews, & Halford, 2007) repeatedly encounter the same choice-options. Options differ in expected value. Children draw samples, gradually learn the expected value of each option, and maximize outcomes by choosing the option highest in expected value. Importantly, in such uninformed decision situations, feedback is the only source of information to elicit expected value of options (but see van Duijvenvoorde, Jansen, Bredman, & Huizenga, 2012, for informed gamble situations). This research has shown that children learn expected values from feedback, but less efficiently than adults (see Cassoti et al., 2014; Defoe, Dubas, Figner, & Aken, 2015, for developmental reviews). Moreover, they show some biases, such as over-responsiveness to negative outcomes and overweighting of large values (Cassoti et al, 2014; van Duijvenvoorde, et al., 2012).

Research on feedback in probabilistic inference decisions is scarce (Lang & Betsch, 2013; see Mata et al., 2011, for 9-11year olds). This thesis is thus the first attempt to systematically investigate the role of feedback in probabilistic inference decisions from a developmental perspective. In contrast to aforementioned gambling paradigms, decisions are more complex and informed, that is probabilities are stated before judgment or decisions, in form of cue validities. Crucially, stated cue validities and values are sufficient and feedback is unnecessary to maximize decision outcomes or to give accurate judgments.

This mixed-source situation constitutes the—specifically human—conflict between stated probability and self-sampled decision outcomes. When stated probabilities are valid, single decision outcomes do not add informational value to the decision and should be disregarded. In adults, feedback nevertheless detrimentally affects decisions and increases deviance from normative benchmarks (Jessup, Bishara, & Busemeyer, 2008; Lejarraga & Gonzalez, 2011; Newell & Rakow, 2007; Yechiam, Erev, & Barron, 2006).

Beneficial as well as detrimental effects of feedback on children's decision making are conceivable in mixed-source situations. Detrimental to decisions might be that feedback increases task complexity and causes cognitive overload (Harvey, 2011). Further, children might prefer feedback over stated probability, and rely on merely feedback-based decision strategies, like adults (Jessup, Bishara, & Busemeyer, 2008; Lejarraga & Gonzalez, 2011; Newell & Rakow, 2007; Yechiam, Erev, & Barron, 2006). Lastly feedback might encourage exploration of the pay-off structure instead of exploitation of stated probabilities (see Mehlhorn et al., 2015, for the trade-off between exploration and exploitation in decisions). As a result, providing decision outcomes after each choice might impair probability utilization in children.

At the same time, feedback might benefit children: It allows to observe cue-outcome contingencies and might increase saliency of the probabilistic relations between cues and outcomes. Further, it allows to update misbeliefs about the task's structure and about the performance of decision strategies. If children do have misconceptions about how valid each cue is or how well a decision strategy performs, they might acquire more accurate beliefs over time with feedback.

In six studies, I investigated detrimental (Article 1) and beneficial (Article 2, 3) effects of feedback on children's probabilistic decisions and judgments (Article 4).

Research Approach

In all studies, I confronted participants with the same probabilistic inference decision task and varied decision feedback (except in Study 6). The paradigm, first used by Betsch & Lang (2013), adapts the classic MouseLab paradigm (Payne et al., 1988). In the child-friendly adaption, participants choose repeatedly between two options, which can yield positive or negative consequences. Before each decision, participants inspect three probabilistic cues that predict options' outcomes but differ in validity, that is, some cues predict outcomes better than others. Throughout all studies the cues' predictions were organized in a 2 × 3 matrix of an information-board, that is, three cues predict the outcomes of two options (see Figure 1). This results in a simple information-board in comparison to adult decision research (e.g., Payne et al., 1988) and to non-probabilistic child research (Davidson, 1991a, 1991b, 1996; Gregan-Paxton & John, 1995, 1997; Howse, Best, & Stone, 2003; Klayman, 1985; Lindow, Lang, & Betsch, 2017). It is rather complex though, in comparison to research in children's probabilistic decision making, where children are often confronted with winning probabilities and values of one or two gambles (e.g., Levin & Hart, 2003; Schlottmann, 2001).

This high complexity has two methodological benefits: First, it allows to detect individual differences in decision strategies. In particular, decisions with more than two cues allow to differentiate between prominent decision strategies (but see Betsch & Lang, 2013, for decisions with two cues and the following sections for more details on decision strategies). Second, the paradigm offers probabilistic and other kinds of information. It is therefore suited to investigate whether children prefer probabilistic information in the presence of irrelevant information (Lang & Betsch, 2013). Arguably, this decision situation resembles real-world decisions much closer, where probability is hardly ever encountered in isolation.

We also used an open information-board in all studies except Study 1. In comparison to a closed board, which requires to search covered cue values sequentially, it reduces the effortful information search and allows to inspect all cue values at once (Betsch et al., 2014; see Glöckner & Betsch, 2008, for discussions).

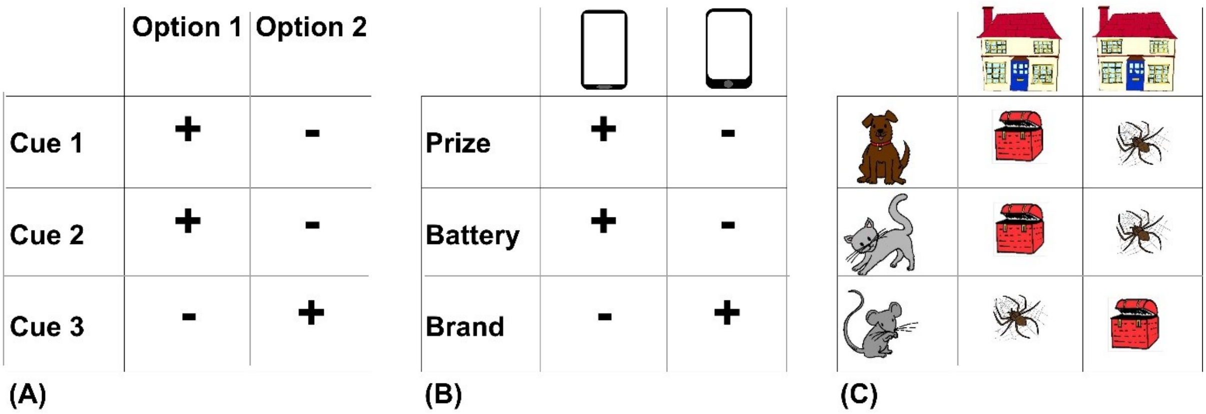


Figure 1. The structure of a probabilistic inference decision with three cues predicting positive (+) or negative (−) outcomes for Option 1 and 2 (A). A real-world example (B): the decision maker wants to buy a durable smart phone, and tries to predict the future durability based on probabilistic cues, such as prize, battery quality and brand. The child-friendly adaption shows the Treasure Hunt (C). Cues are represented by animals and options by houses; positive and negative cue values are pictorial.

Stated Cue Validities

Whereas in adult decision making, probabilities are mainly given in numerical formats—mostly as percentages (80% winning probability) or relative frequencies (8 out of 10)—there are various ways to address this challenge in child participants. The first and most straightforward way is to provide children with a natural event-space that defines the probability. A winning probability of $p = .33$ for a lottery can, for example, be constructed by three boxes with only one containing a prize (e.g., Levin et al., 2007). The second approach is to substitute this event-space by an easy to grasp variable co-varying with probability, for example a magnitude. Accordingly, the lottery's winning probability is illustrated by a spinner wheel consisting of two colored areas, when the winning area takes up one third of the total (e.g., Schlottmann, 2001). Third, frequencies can be observed over time in a defined sample, for example when the lottery wins in one out of three times (e.g., Pasquini et al., 2007). Percentages or other numerical formats can be used at late elementary school age, when children understand them sufficiently (e.g., Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Mata, et al., 2011; van Dujvenvoorde et al., 2012).

We used a relative frequency format to demonstrate each cue's validity to children (Figure 2A-C). Encoding of frequencies is automatic and effortless (Hasher & Zacks, 1979) and frequency formats are simpler to process than probability formats even for adults (Gigerenzer & Hoffrage, 1995). Participants thus observed cue predictions and the respective outcome several times in a defined sample and encoded how often the cue predicted outcomes correctly (e.g., in three out of six times). All variables (that is, cue predictions and outcomes) were binary (positive, negative) and pictorial (treasure and spider) to reduce complexity and facilitate encoding (Gregan-Paxton & John, 1997).

The relation between the cue and the predicted outcome variable is given by the cue validity. A cue can predict a positive outcome correctly (+ | +) or a negative outcome correctly (− | −). It can also incorrectly predict a positive outcome (+ | −) or incorrectly predict a negative outcome (− | +). Exact definitions of cue validities vary between research fields, researchers, and specifics of the task (e.g., Martignon & Hoffrage, 2002; Platzer & Bröder, 2012). We calculated each cue's validity as the probability that the cue predicts the outcome correctly, given by the ratio of the number of correct to all predictions.

$$p = \frac{(+\mid+) + (-\mid-)}{(+\mid+) + (-\mid-) + (+\mid-) + (-\mid+)}$$

Note, that alternative accounts to describe the statistical relation result in similar cue hierarchies.

Importantly, the range of cue validity is constrained in binary worlds: A cue validity of $p = .50$ is uninformative but cue validities below can be inversely used (e.g., $p = .10 \equiv .90$). Cue validities can thus at most vary between $.50 \leq p < 1$, if each cue is probabilistically related to the outcome. However, since we introduced cue validities as frequencies, the range of cue validities was further constraint by the sample size n (i.e., the number of times each cue predicted outcomes and was granted or denied a smart point). Thus, the most dispersed cue validities for $n = 6$, were $p = .50$, $p = .83$, in accordance with three and five correct predictions; and for $n = 7$, $p = .56$, $p = .86$, in accordance with four and six correct predictions.

Expected Value Calculations and Decision Strategies

Normative decision theory prescribes to weight cue values by their validity, add them up to the weighted sum for each option (for example in Figure 2D, $EV_{\text{left option}} = w_1 \times 0 + w_2 \times 0 + w_3 \times 1$ and $EV_{\text{right option}} = w_1 \times 1 + w_2 \times 0 + w_3 \times 0$), and choose the option with the maximum expected value.

It is still debated whether and how cue validities must be transformed to be used as weights in these calculations. For illustration, consider an uninformative cue with $p = .5$. When calculating the expected value, this cue's values should be assigned a weight of 0. A log-transformation of cue validities accounts for that fact, ($log \frac{p}{1\text{-}p}$, Lee, 2014). Other suggestions are to subtract .5, and use these "chance-corrected" cue validities ($p - .50$; Hilbig & Moshagen, 2014; Jekel, Glöckner, Fiedler, & Bröder, 2012), or use cue validities as weights without any correction (Gigerenzer et al., 1999; Rieskamp and Otto, 2006). Transforming weights can lead to different predictions for weighted-additive models. Previous results using the Treasure Hunt suggest that some adults applied a weighted-additive model with uncorrected cue validities (which prescribes to choose the left option in Figure 2D, E, but the right option in Figure 2F; Betsch et al., 2014).

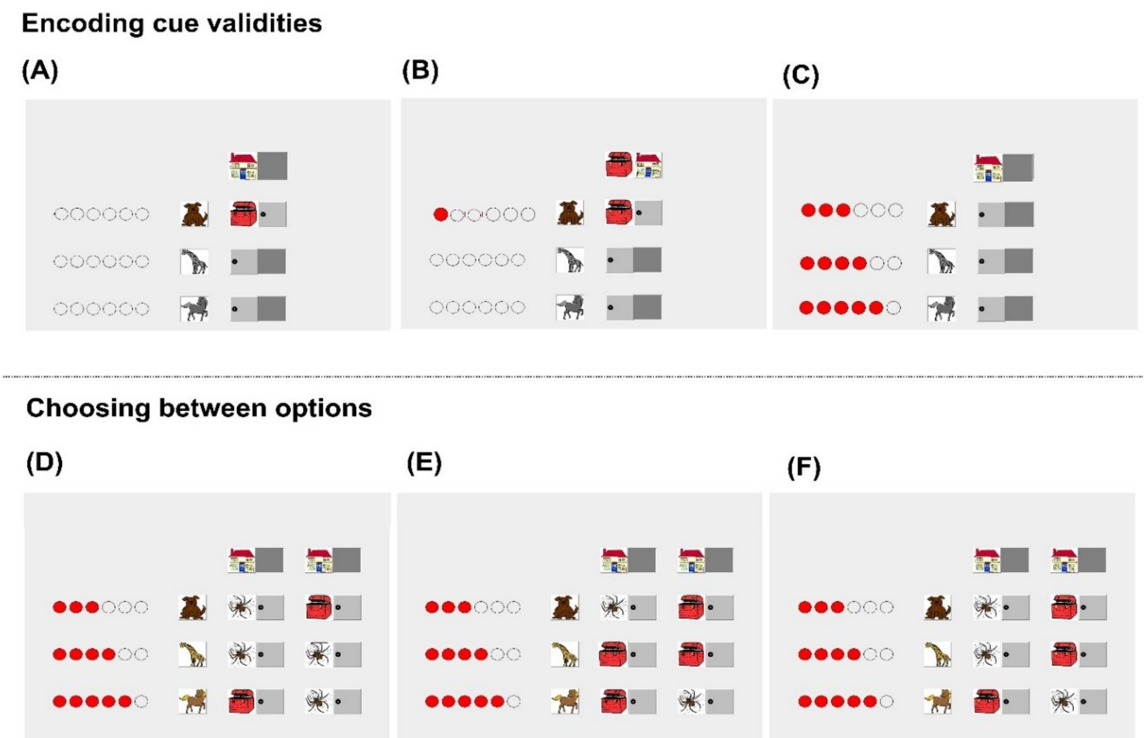


Figure 2. Figure shows the Treasure Hunt with n = 6 and p_1 = .50, p_2 = .66, p_3 = .83. Participants encoded cue validities prior to decision making (A-C). They observed the prediction of one cue and the outcome. If the cue predicted correctly, they granted a smart point to the cue. Cue validities were thus depicted graphically as a magnitude of smart points. Participants then made several decisions between two options and could inspect varying predictions of three cues (D-F).

The bounded rationality approach to decision making suggests that expected value calculations are cognitively effortful and time consuming (Gigerenzer et al., 1999; Payne et al., 1988; Simon, 1954). Thus, humans rely on simplifying decision strategies (Gigerenzer, et al, 1999). The adult decision making literature has produced a variety of strategies that can be applied to probabilistic inferences apart from the normative weighted-additive model. The most prominent is the Lexicographic or take-the-best strategy (Fishburn, 1974; Gigerenzer & Goldstein, 1996). It prescribes to consider only the values of the most valid cue and to choose the option for which this cue's value is positive (e.g., the left option in Figure 2D). When the cue is indifferent, the strategy prescribes to go to the next highest valid cue or to guess between options.

Another prominent example, is the Equal Weight strategy (Payne et al., 1988). It neglects differences in cue validities, only counts the number of positive values for each option, and choses the option with the maximum positive values (e.g., right option in Figure 2F) or guesses in case of a tie.

Dispersion of Cue Validities

The dispersion of cue validities and thus whether a strategy was adaptive varied between studies. In all studies except Study 5, we created highly dispersed decision environments. In order to maximize decision outcomes, a Lexicographic strategy that only considers the most valid cue’s values is sufficient and allows fast and frugal decision making (Gigerenzer & Goldstein, 1996; Gigerenzer, et al, 1999). More complex strategies, such as weighted-additive yield equally good results, but request more time and cognitive efforts (but see Glöckner & Betsch 2008, for automatic, non-effortful implementation of a weighted-additive-like process). Accordingly, the environment invites the application of simple strategies.

Assigning the same validity to both the low valid cues in Study 4, further increased the salience of the most valid cue. Study 5 implemented a low dispersion of cue validities, in which decision strategies like Equal Weight were also adaptive.

Decision Feedback

In all studies, except Study 6, we additionally varied the presence of feedback on decisions outcomes as an experimental factor (Table 3). Control conditions offered no immediate feedback after decisions. Participants were confined to stated cue validities to inform judgments or decisions. When feedback maintained cue validities, the stated cue validities were reproduced in the feedback structure, for example, a cue with a stated validity of $p = .50$ correctly predicted outcomes in 50% of decisions. When feedback reinforced decision strategies, one defined strategy yielded the best outcomes while competing strategies performed worse. To achieve these differences in reinforcement, cue validities were allowed to vary over the course of decisions. Note, that feedback was normatively irrelevant to maximize outcomes. All relevant information, that is cue validities and values, were available in advance.

To sum up, we used a mixed-source paradigm that stated probabilistic information prior to decision making in form of probabilistic relations between cues and outcomes—the cues’ validities—and varied immediate feedback on decision outcomes.

Table 1

Overview of decision environments in Study 1-6

Study	Stated probabilities	Feedback		
		None	Maintain	Reinforce
1	$p1 = .5, p2 = .66, p3 = .83$	X	X	—
2	$p1 = .5, p2 = .66, p3 = .83$	X	X	—
3	$p1 = .5, p2 = .66, p3 = .83$	X	X	—
4	$p1 = .57, p2 = .57, p3 = .86$	X	X	X
5	$p1 = .72, p2 = .72, p3 = .86$	X	—	X
6	$p1 = .5, p2 = .66, p3 = .83$	—	X	—

Note. Maintain = feedback maintains stated probabilities; Reinforce = feedback reinforces a specified decision strategy; probabilities are allowed to diverge from stated values over time.

Measuring decisions, decision strategies and judgments. In Articles 1-3, we analyzed participants' decisions. The approach was two-folded: Choice data at the group level was compared with a normative benchmark. In addition, conclusions about each individual's strategy in a series of decisions were drawn from individual data patterns, using an outcome-based maximum likelihood classification method (Bröder & Schiffer, 2003). Other methods, such as verbal protocols or tracking of active information search might bias strategy application (e.g., Betsch, Funke, & Plessner, 2011; Glöckner & Betsch, 2008 for discussions).

Article 4 analyzes probability utilization in predictive judgments. Participants predicted event frequencies that should be based on cue validities. There, I likewise analyzed judgments on group level, comparing each age groups to a normative benchmark, and additionally individual judgments.

Summary of the Research. Article 1 investigates whether children prefer feedback over stated probabilistic information and whether interference of feedback can explain deviations from normative models of choices. Six-year-olds', 9-year-olds', and adults' decision making was examined in a decision environment where probabilistic information about choice outcome had to be actively searched ($N = 166$) or was available without search before choices ($N = 183$). This information was provided as predictions of cues differing in validity. The presence of outcome feedback was varied.

Six-year-olds, but not 9-year-olds were over-responsive to negative outcomes leading to choices biased by recent feedback. However, children did not systematically utilize feedback in choices. Irrespective of feedback, 6-year-olds fully and 9-year-olds partly neglected stated probabilistic information in their choices. When 6-year-olds chose systematically, they only relied on irrelevant information, which did not maximize outcomes. Nine-year-olds still applied invalid choice rules, but also choice rules based on probability.

Results suggest that neglect of probabilities in complex decisions is robust, independent of feedback, and only starts to subside at elementary school age.

Article 2 assessed both adaptive and non-adaptive decision strategies in 6-year-olds, 9-year-olds and adults to investigate how often, accurately and persistently children apply them. Previously, children aged nine and younger seemed to decide randomly without any systematic plan in probabilistic decision environments. Research focused on detecting adaptive decision strategies that consider probabilistic information embedded in the environment and are used by adults. However, it ignored non-adaptive decision strategies that focus on irrelevant information and might be used by children.

Results show that half of the children in each age group applied a decision strategy accurately. Six-year-olds predominantly used non-adaptive strategies, while 9-year-olds used adaptive strategies as well. Full feedback positively affected strategy use: children abandoned non-adaptive strategies, or applied them less consistently.

Results suggest that children are able to systematically apply decision strategies very early in life. Similar to other cognitive areas, they initially use inefficient strategies, but overcome them with increasing age and experience.

In Article 3, we investigated 7- and 9-year-old children's learning of simple adaptive strategies from feedback. In Study 1 ($N = 316$), both age groups learned the selective Lexicographic decision strategy, which only focused on the most valid cue, equally good. Strategy learning was slighty better in a stricter compared with a more lenient feedback condition and worse than in adult controls. In Study 2 ($N = 259$), feedback either reinforced the selective Lexicographic strategy or the holistic Equal Weight strategy. Younger children learned the holistic strategy better, while older children learned both equally well. In both studies, children rarely used adaptive decision strategies without feedback, although a priori stated dispersion of probabilities suggested their use.

Under ideal circumstances, that is, in highly dispersed probabilistic environments and when adaptive strategies are not too complex to perform, feeding back decision outcomes improves children's, otherwise very deficient, decision making.

Article 4 investigates whether children at the age of six and nine utilize probabilistic relations between cues and outcomes in predictive judgments. It shows that probability utilization in judgments emerges late and follows the same developmental trajectory as in decisions: Probability utilization is absent in 6-year-olds and emerging in 9-year-olds. Children's judgments show that their expectations of outcomes are not informed by probability until late elementary school age. Results remain consistent across studies and are unaffected by experience of participants or scale formats used to assess judgments.

The findings contradict the notion that children utilize probabilities in judgments earlier than in decisions and highlight that deficits in probabilistic judgments persist until the age of nine.

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. https://doi.org/10.2307/1131034

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.110–134). Hillsdale, NJ: Erlbaum.

Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin, 3,* 410–433. https://doi.org/10.1037/0033-2909.106.3.410

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61,* 213–238. https://doi.org/10.3102/00346543061002213

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making, 16*, 215–233. https://doi.org/10.1002/bdm.443.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*, 7–15. https://doi.org/10.1016/0010-0277(94)90018-3.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review, 3*, 439-449. https://doi.org/10.5465/amr.1978.4305717

Betsch, T., Funke, J., & Plessner, H. (2011). *Denken - Urteilen, Entscheiden, Problemlösen*. Berlin: Springer.

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology*, *115*, 108–125. https://doi.org/10.1016/j.jecp.2012.11.003

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica*, *153*, 74–86. https://doi.org/10.1016/j.actpsy.2014.09.008

Boyer, T. W. (2007). Decision making processes: Sensitivity to sequentially experienced outcome probabilities. *Journal of Experimental Child Psychology, 97*, 28–43. https://doi.org/10.1016/j.jecp.2006.12.008

Brainerd, C. J. (1981). Working memory and the developmental analysis of probability judgment. *Psychological Review, 88,* 463–502. https://doi.org/10.1037/0033-295X.88.6.463

Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica, 45*. 223–241. https://doi.org/10.1016/0001-6918(80)90034-7

Bröder, A., Glöckner, A., Betsch, T., Link, D., & Ettlin, F. (2013). Do people learn option or strategy routines in multi-attribute decisions? The answer depends on subtle factors. A*cta Psychologica, 143*, 200–209. https://doi.org/10.1016/j.actpsy.2013.03.005

Bröder, A., Newell, B. R., & Platzer, C. (2010). Cue integration vs. exemplar-based reasoning in multi-attribute decisions from memory: A matter of cue representation. *Judgment and Decision Making, 5,* 326–338. Retrieved from http://journal.sjdm.org/10/10614a/jdm10614a.pdf

Bröder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making*, *16*(3), 193–213. https://doi.org/10.1002/bdm.442

Bruine de Bruin W., Parker A., & Fischhoff, B. (2007) Individual differences in adult decision-making competence (A-DMC). *Journal of Personality and Social Psychology, 92*, 938–956. https://doi.org/10.1037/0022-3514.92.5.938

Brunswik, E. (1952). *The conceptual framework of psychology.* Chicago: University of Chicago Press.

Brunswik, E. (1955). Representative design and the probability theory in a functional psychology. *Psychological Review, 62*, 193–217. https://doi.org/10.1037/h0047470

Bryant, P., & Nunes, T. (2012). Children's understanding of probability: A literature review (full report). A report to the Nuffield Foundation. Retrieved from http://www.nuffieldfoundation.org/sites/default/files/files/Nuffield_CuP_FULL_REPORTv_FINAL.pdf

Bunch, K. M., Andrews, G., & Halford, G. S. (2007). Complexity effects on the children's gambling task. *Cognitive Development*, *22*(3), 376–383. https://doi.org/10.1016/j.cogdev.2007.01.004

Byrnes, J. P., Miller, D. C., & Reynolds, M. (1999), Learning to Make Good Decisions: A Self-Regulation Perspective. *Child Development, 70*, 1121–1140. https://doi.org/10.1111/1467-8624.00082

Cassotti, M., Aïte, A., Osmont, A, Houdé, O., & Borst, G. (2014). What we learned about the processes involved in the Iowa Gambling Task. *Frontiers in Psychology*, 5, 915. https://doi.org/10.3389/fpsyg.2014.00915

Davidson, D. (1991a). Children's decision-making examined with an information-board procedure. *Cognitive Development, 6,* 77–90. https://doi.org/10.1016/0885-2014(91)90007-Z

Davidson, D. (1991b). Developmental differences in children's search of predecisional information. *Journal of Experimental Child Psychology, 52*, 239–255. https://doi.org/10.1016/0022-0965(91)90061-V

Davies, C. M. (1965). Development of the probability concept in children. *Child Development, 36,* 779–788. https://doi.org/10.1111/j.1467-8624.1965.tb05338.x.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron. 69,* 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. G. (2015). A meta-analysis on age differences in risky decision making: Adolescents versus children and adults. *Psychological Bulletin, 141*(1), 48-84. https://doi.org/10.1037/a0038088

Denison, S., Bonawitz, E. B., Gopnik, A., & Griffiths, T. L. (2010). Preschoolers sample from probability distributions. In R. Camtrabone & S. Ohlsson (Eds.), *Proceedings of the 32nd annual conference of the Cognitive Science Society* (pp. 2272-2277). Austin, TX: Cognitive Science Society.

Denison, S., Reed, C., & Xu, F. (2013). Probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology, 49(2)*, 243–249. https://doi.org/10.1037/a0028278

Denison, S., & Xu, F. (2010). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science,13,* 798–803. https://doi.org/10.1111/j.1467-7687.2009.00943.x

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*, 380–417. https://doi.org/10.1037/h0053870

Einhorn, J. H., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review, 86,* 465–485. https://doi.org/10.1037/0033-295X.86.5.465

Erev, I., & Barron, G. (2005). On adaptation, maximization and reinforcement learning among cognitive strategies. *Psychological Review, 112*, 912–931. https://doi.org/10.1037/0033-295X.112.4.912.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., … Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. J*ournal of Behavioral Decision Making, 23*, 15–47. https://doi.org/10.1002/bdm.683.

Evans, J. S. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13*, 378–395. https://doi.org/10.3758/BF03193858

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*, 186–203. https://doi.org/10.1037/0278-7393.34.1.186.

Finucane, M. L., & Gullion, C. M. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychology and Aging*, *25*(2), 271–288. https://doi.org/10.1037/a0019106

Finucane, M. L., Slovic, P. , Hibbard, J. H., Peters, E. , Mertz, C. K. and MacGregor, D. G. (2002), Aging and decision-making competence: An analysis of comprehension and consistency skills in older versus younger adults considering health-plan options. *Journal of Behavioral Decision Making, 15,* 141-164. https://doi.org/10.1002/bdm.407

Fishburn, P. C. (1974). Lexicographic orders, utilities and decision rules: A survey. *Management Science, 20*(11), 1442–1471. Retrieved from http://www.jstor.org/stable/2629975

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology,108*(1)*,* 82–97. https://doi.org/10.1037/edu0000053

Girotto, V., & Gonzalez, M. (2008). Children's understanding of posterior probability. *Cognition, 106*, 325–344. https://doi.org/10.1016/j.cognition.2007.02.005

Girotto, V. Fontanari, L. Gonzalez, M., Vallortigara, G., & Blaye A. (2016) Young children do not succeed in choice tasks that imply evaluating chances, *Cognition, 152,* 32–39, https://doi.org/10.1016/j.cognition.2016.03.010

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650–669. https://doi.org/10.1037/0033-295X.103.4.650

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684–704. https://doi.org/10.1037/0033-295X.102.4.684

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

Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(5),* 1055-1075. https://doi.org/1055-1075

Gregan-Paxton, J., & John, D. R. (1995). Are young children adaptive decision makers? A study of age differences in information search behavior. *Journal of Consumer Research, 21,* 567–580. https://doi.org/10.1086/209419

Gregan-Paxton, J., & John, D. R. (1997). The emergence of adaptive decision making in children. *Journal of Consumer Research, 24,* 43–56. https://doi.org/10.1086/209492

Gonzalez, M. & Girotto, V. (2011). Combinatorics and probability. Six- to ten-year-olds reliably predict whether a relation will occur. *Cognition, 120*, 372-379. https://doi.org/10.1016/j.cognition.2010.10.006

Goldberg, S. (1966). Probability judgments by preschool children: Task conditions and performance. *Child Development, 37,* 157-167. https://doi.org/10.2307/1126436.

Harvey, N. & Fisher, I. (2005). Development of experience-based judgment and decision making: The role of outcome feedback. In T. Betsch & S. Haberstroh (Eds.), *The Routines of Decision Making.* Mahwah, NJ: Lawrence Erlbaum.

Harvey, N., & Fischer, I. (1997). Taking advice: accepting help, improving judgment and sharing responsibility. *Organizational Behavior and Human Decision Making Process, 70,* 117–133. https://doi.org/10.1006/obhd.1997.2697

Hasher, L., & Zacks, R. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*(3), 356–388. https://doi.org/10.1037/0096-3445.108.3.356

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15*, 534–539. https://doi.org/10.1111/j.0956-7976.2004.00715.x.

Hilbig, B. E., & Moshagen, M. (2014). Generalized outcome-based strategy classification: Comparing deterministic and probabilistic choice models. *Psychonomic Bulletin & Review, 21,* 1431–1443. http://dx.https://doi.org/10.3758/s13423-014-0643-0

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, NY: Springer.

Hogarth, R. M. (2001). *Educating intuition*. Chicago, IL: University of Chicago Press.

Hogarth, R. M. (2005). Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thought. In T. Betsch & S. Haberstroh (Eds.), *The routines of decision making*. Mahwah, NJ: Lawrence Erlbaum Associates.

Howse, R., Best, D.L., & Stone, E.R. (2003) Children's decision making: the effects oftraining, reinforcement, and memory aids. *Cognitive Development, 18,* 247–268. https://doi.org/10.1016/S0885-2014(03)00023-6

Jekel, M., Glöckner, A., Fiedler, S., & Bröder, A. (2012). The rationality of different kinds of intuitive decision processes. *Synthese, 189*, 147–160. https://doi.org/10.1007/s11229-012-0126-7

Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychological Science, 19*, 1015–1021. https://doi.org/10.1111/j.1467-9280.2008.02193.x.

Juslin, P., Jones, S., Olsson, H., & Winman, A. (2003). Cue abstraction and exemplar memory in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(5), 924–941. https://doi.org/10.1037/0278-7393.29.5.924

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*, 697–720. https://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D. & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York, NY: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291. https://doi.org/10.2307/1914185.

Kalish, C. W. (2010). How children use examples to make conditional predictions. *Cognition, 116*(1), 1–14. https://doi.org/10.1016/j.cognition.2010.03.008

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, *134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Karlsson, L., Juslin, P., & Olsson, H. (2007). Adaptive changes between cue abstraction and exemplar memory in a multiple-cue judgment task with continuous cues. *Psychonomic Bulletin & Review, 14*(6), 1140–1146. https://doi.org/10.3758/BF03193103

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision-making. *Child Development, 72,* 844–861. https://doi.org/10.1111/1467-8624.00319

Klaczynski, P. A. (2005). Metacognition and cognitive variability: A two-process model of decision making and its development. In J. E. Jacobs, & P. A. Klaczynski (Eds.), *The development of decision making in children and adolescents* (pp. 39–76). Mahwah, NJ: Erlbaum.

Klayman, J. (1985). Children's decision strategies and their adaptation to task characteristics. *Organizational Behavior & Human Decision Processes, 35,* 179–201. https://doi.org/10.1016/0749-5978(85)90034-2

Kluger, A., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254.

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*(1), 26–52. https://doi.org/10.1016/S0022-0965(02)00121-2.

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General, 135*(2), 162–183. https://doi.org/10.1037/0096-3445.135.2.162.

Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. O*rganizational Behavior and Human Decision Processes, 116,* 286–295. https://doi.org/10.1016/j.obhdp.2011.05.001

Lee, M. D. (2014). Bayesian outcome-based strategy classification. *Behavior Research Methods, 48*(1), 29–41. https://doi.org/10.3758/s13428-014-0557-9

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*(5), 397–413. https://doi.org/10.1002/bdm.453

Levin, I. P., Hart, S. S., Weller, J. A., & Harshman, L. A. (2007). 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. https://doi.org/10.1002/bdm.552

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. Retrieved from http://journal.sjdm.org/jdm7404.pdf

Lindow, S., Lang, A., & Betsch, T. (2017) Holistic information integration in child decision making. *Journal of Behavioral Decision Making*, *30*(5), 1131–1146. https://doi.org/10.1002/bdm.2029

Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision, 52*, 29–71. https://doi.org/10.1023/A:1015516217425

Mata, R., von Helversen, B., & Rieskamp, J. (2011), When easy comes hard: The development of adaptive strategy selection. *Child Development, 82*, 687–700. https://doi.org/10.1111/j.1467-8624.2010.01535.x

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., … Gonzalez, C. (2015). Unpacking the Exploration–Exploitation Tradeoff: A Synthesis of Human and Animal Literatures. *Decision, 2*(3), 191–215. https://doi.org/10.1037/dec0000033

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner, & R. Brunken (Eds.), *Instructional design for multimedia learning* (pp. 181–195). Munster, NY: Waxmann.

Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin & Review, 14*, 1133–1139. https://doi.org/10.3758/BF03193102

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, *43*(5), 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual-differences approach. *Journal of Behavioral Decision Making, 18(1)*, 1–27. https://doi.org/10.1002/bdm.481

Pavlov, I.P. (1926) *Conditional Reflexes.* New York: Dover Publications.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review, 31*, 180–206. https://doi.org/10.1016/j.dr.2011.07.004

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). Oxford: Wiley.

Savage, L. J. (1954). *The foundations of statistics*. Oxford, GB: Wiley.

Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *Child Development, 72*, 103–122. https://doi.org/10.1111/1467-8624.00268

Schlottmann, A., & Wilkening, F. (2012). Judgment and decision making in young children. In M. K. Dhami, A. Schlottmann & M. R. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development and evolution* (pp. 55–83). Cambridge: University Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Shute, Valerie. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69,* 99–118. Retrieved from http://www.jstor.org/stable/1884852

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*, 645–726.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2011).The complexity of developmental predictions from dual process models. *Developmental Review, 31*, 103–118. https://doi.org/10.1016/j.dr.2011.07.003.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Teglas, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences of the United States of America, 104*(48), 19156–19159. https://doi.org/10.1073/pnas.0700271104

Teglas, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332*, 1054–1058. https://doi.org/10.1126/science.1196404

Tversky, A., & Kahnemann, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–31. https://doi.org/10.1126/science.185.4157.1124

van den Bos, W., Güroğlu, B., van den Bulk, B. G, Rombouts, S. A. R. B., & Crone, E. A. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience*, *3*, 52. https://doi.org/10.3389/neuro.09.052.2009

van Duijvenvoorde, A. C. K, Jansen, B. R. J., Bredman, J. C., & Huizenga, H. M. (2012). Age-related changes in decision making: Comparing informed and noninformed situations. *Developmental Psychology*, *48*(1), 192–203. https://doi.org/10.1037/a0025601

von Neumann, J. & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Worthy, D. A., & Maddox, W.T. (2014). A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to W. K. Estes, *Journal of Mathematical Psychology*, 59, 41–49. https://doi.org/10.1016/j.jmp.2013.10.001

Weiner, B. (1974). *Achievement motivation and attribution theory*. Morristown, NJ: General Learning Press.

Weller, J. A., Levin, I. P., Rose, J. P., & Bossard, E. (2012), Assessment of Decision-making Competence in Preadolescence. *Journal of Behavioral Decision Making*, *25(4)*, 414–426. https://doi.org/10.1002/bdm.744

Yechiam, E., Erev, I., & Barron, G. (2006). The effect of experience on using a safety device. *Safety Science, 44*, 515–522. https://doi.org/0.1016/j.ssci.2005.11.006.

Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: the role of representation in mental computation. *Cognition, 98*, 287–308. https://doi.org/10.1016/j.cognition.2004.12.003.

Zmyj, N., Buttelmann, D., Carpenter, M. J., & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology, 106*(4), 208–220. https://doi.org/ 10.1016/j.jecp.2010.03.002

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition, 112*, 97–104. https://doi.org/10.1016/j.cognition.2009.04.006

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America, 105*(13), 5012–5015. https://doi.org/10.1073/pnas.0704450105

Article 1

Children’s Neglect of Probabilities in Decision Making with and without Feedback

Children's Neglect of Probabilities in Decision Making with and without Feedback

Anna Lang and Tilmann Betsch
University of Erfurt, Germany

Author Note

Anna Lang, Department of Psychology, University of Erfurt, Germany; Tilmann Betsch, Department of Psychology, University of Erfurt, Germany.

This research was supported by a grant from the Deutsche Forschungsgemeinschaft to the second author (Grant No. BE 2012/11-1).

Correspondence concerning this article should be addressed to Anna Lang, Department of Psychology, University of Erfurt, Nordhaeuser Strasse 63, 99089 Erfurt, Germany. E-mail: anna.lang@uni-erfurt.de

Declarations of interest: none

Word count: 10598

Abstract

We investigated whether children prefer feedback over stated probabilistic information and whether interference of feedback can explain deviations from normative models of choices. 6-year-olds', 9-year-olds', and adults' decision making was examined in a decision environment where probabilistic information about choice outcome had to be actively searched ($N = 166$) or was available without search before choices ($N = 183$). This information was provided as predictions of cues differing in validity. The presence of outcome feedback was varied. Six-year-olds, but not nine-year-olds were over-responsive to negative outcomes leading to choices biased by recent feedback. However, children did not systematically utilize feedback in choices. Irrespective of feedback, 6-year-olds fully and 9-year-olds partly neglected stated probabilistic information in their choices. When 6-year-olds chose systematically, they only relied on invalid information, which did not maximize outcomes. Nine-year-olds still applied invalid choice rules, but also choice rules based on probability. Results suggest that neglect of probabilities in complex decisions is robust, independent of feedback, and only starts to subside at elementary school age.

Keywords: child decision making, probabilistic inference, feedback, win-stay-lose-shift, information board

Children's Neglect of Probabilities in Decision Making with and without Feedback

Children, like adults, must make decisions in an uncertain world. Distributions of behavioral outcomes are governed by the rules of probability. It is a developmental challenge to become sensitive and responsive to the probabilistic relations between choices and outcomes in order to achieve decision competence. Information about these relations can be conveyed prior to choices in a summarized form in terms of stated probabilities, for example by stating winning probabilities of gambles (Edwards, 1954; Kahneman & Tversky, 1979). Such information can also be gradually acquired when choices are performed and met with feedback from the environment. When sequentially sampled feedback is valid and properly stored in memory, it can provide a powerful source for subsequent intuitions and decisions (Hogarth, 2001).

There is large body of research on the utilization of stated probabilities—not only in the adult but also in the child decision making literature (e.g., Girotto & Gonzalez, 2008; Levin, Hart, Weller, & Harshman, 2007; Schlottman, 2001). There is also a great number of studies on experience-based decision making where children learn from feedback (e.g., Boyer, 2007; Bunch, Andrews, & Halford, 2007; Van den Bos, Güroglu, van den Bulk, Rambouts, & Crone, 2009). Mixed-source paradigms in which decision makers have access to both stated probabilities and feedback, however, are rare in child decision making research (Betsch, Lang, Lehmann, & Axmann, 2014; Van Duijvenvoorde, Jansen, Bredman & Huizenga, 2012; Mata, van Helversen, & Rieskamp, 2011).

In addition, so far research lacks the systematic variation of access to both kinds of information to understand their respective influence on decision making in children. As a first attempt to fill this gap, we varied the presence of feedback in an environment that provided 6-year-olds and 9-year-olds with probabilistic cues and compared their performance to adults.

Children's Utilization of Stated Probabilistic Information in Decision Making

In a variety of paradigms, children are confronted with stated probabilistic information in child-friendly formats, such as relative frequencies of options' wins and losses, which they can subsequently utilize in choices. According to the classical Piagetian view, children up to 8 years of age completely lack an understanding of probability and thus the ability to properly utilize such probabilistic information (Piaget & Inhelder, 1951). However, more recent research has documented sensitivity for probabilistic information in 8-months-old infants who distinguished probable from less probable samples based on statistical properties of their environment (Xu & Garcia, 2008). Preschool-aged children are able to utilize stated probabilistic information for decisions in social contexts when choosing whom to trust or imitate (Pasquini, Corriveau, Koenig, & Harris, 2007; Zmyj, Buttelmann, Carpenter, & Daum, 2010) but also to judge gambles with different expected values. Children as young as 5 years consider winning probabilities and values and even integrate these variables in a multiplicative fashion (Schlottmann, 2001; Schlottmann & Tring, 2005).

When it is required to not only judge but to choose between two or more risky options the findings have been strikingly different. Employing an information board paradigm, Betsch and co-workers (Betsch & Lang, 2013; Betsch et al., 2014; Betsch, Lehmann, Lindow, Lang, & Schoemann, 2016) assessed probabilistic inference decisions in preschoolers (6-year-olds) and elementary schoolers (9-year-olds). Preschoolers were not able to systematically use stated probabilistic information as decision weights. Moreover, only about one third of elementary schoolers was able to do so. This is consistent with studies that have shown that children up to 8 years of age do not systematically consider stated probabilistic information when choosing between lotteries (Levin & Hart, 2003; Levin, Weller, Pederson, & Harshman, 2007, but see Levin, Hart, et al., 2007). Thus, the evidence on children's abilities to utilize stated probabilistic information for decision making is mixed. Judgment tasks so far suggest that children at preschool age consider such information (Acredolo, O'Connor, Banks, & Horobin, 1989; Anderson & Schlottmann, 1991; Schlottmann, 2001; Schlottmann & Tring, 2005); results in choice tasks are rather inconclusive with some demonstrating its utilization prior to school age (Levin, Hart, et al., 2007; Pasquini et al. 2007; Zmyj et al., 2010), while others do not (Betsch & Lang, 2013; Betsch et al., 2014, 2016; Levin & Hart, 2003; Levin Weller, et al. 2007).

Children's Utilization of Feedback in Decision Making

Children's ability to improve their decisions through feedback has been studied with gambling tasks in which participants repeatedly choose between options of different expected value (e.g., decks of cards), and only learn about frequencies and magnitudes of associated gains and losses through experiencing outcomes. Children up to 13 years of age typically fail to properly learn from feedback to avoid inferior options when multiple pieces of information such as gains and losses for different options have to be considered to make a choice (forbreviews see, Cassotti, Aïte, Osmont, Houde, & Borst, 2014; Defoe, Dubas, Figner, & van Aken, 2014). The high complexity of this type of game presumably accounts for the poor performance (but cf. Crone, Bunge, Latenstein, & van der Molen, 2005). In simpler versions, children at preschool age learn to choose the advantageous option more frequently (Boyer, 2007; Kerr & Zelazo, 2004). When options only differ on one dimension (e.g., only gains) even 3-year-olds can adapt their choices and improve their decisions over time (Bunch et al., 2007).

Further, children's deficits in feedback processing are, at least partly, caused by their over-responsiveness to negative outcomes. When choosing between options, children are highly sensitive towards negative feedback (Crone et al., 2005; Eppinger, Mock, & Kray, 2009; Huizenga, Crone, & Jansen, 2007) and tend to switch behavioral responses after experiencing failure (Cassotti, Houdé, & Moutier, 2011; Van Duijvenvoorde et al., 2012). In a probabilistic environment where even the best performing behavior sometimes provides negative outcomes, this tendency leads children to abandon the superior response and thus prevents optimal choice performance (see Brehmer, 1980, for similar findings in adults)

Mixed-Information Paradigms

Studies that provide children with both stated information about probabilistic relations between choices and outcomes and self-sampled experience are rare. Mata et al. (2011) demonstrated that 9-year-olds can adapt their decision strategies to feedback in an information-board paradigm, when probabilistic information is provided before choices. Van Duijvenvoorde et al. (2012) observed children's abilities to identify the advantageous option in a gambling task to be much improved when outcome probabilities and values were made explicit. However, even then children up to 13 years could not overcome the tendency to switch options after experiencing failure.

While these findings show that decisions can improve when the probabilistic properties of the environment are stated, research in child decision making provides little insight into how feedback influences choices that should be based on stated probabilistic information. Importantly, even in adults, feedback does not always improve decision making (see Karelaia & Hogarth, 2008, for a review) and can increase deviations from normative models (Barron & Erev, 2003; Newell & Rakow, 2007). Betsch and co-workers (Betsch & Lang, 2013M; Betsch et al., 2014, 2016) found that 6-year-old preschoolers fully neglected stated probabilistic information, available as validities of different cues, and 9-year-olds partly did so when feedback could also be used for subsequent choices. Especially preschoolers switched between options in line with the last experienced outcome (Betsch & Lang, 2013). This leads to the question whether children's neglect of such probabilistic information is due to the presence of feedback. As a second source of information, feedback can be used mal-adaptively by children and decrease the reliance on stated probabilistic information. If this is the case, children's utilization of stated probabilistic information should be increased when no feedback is available. On the other hand, if feedback does not contribute to children's probability neglect, it should be observed with and without feedback.

Research Goal and Approach

In two studies we investigated whether feedback can account for children's neglect of stated probabilistic information in a mixed-source paradigm. We followed the research approach from Betsch et al. (2014) in which probabilistic information was available as validities of predictive cues while feedback about decision outcomes could be experienced after each choice. Crucially, the information was redundant. Feedback reinforced the structure of the cues' validities and reinforced both options equally. To investigate whether children's decision making deteriorates when they can sample choices outcomes themselves, we manipulated the presence of feedback. When feedback was provided we expected children to demonstrate over-responsiveness to negative outcomes, and to prefer feedback information over cue validities, reflected in relying on a simplifying feedback-heuristic, that is, staying with one option until it fails. When feedback was absent, we expected children to increasingly utilize cue validities.

Research Paradigm

In both experiments, we used a computerized version of an information-board paradigm (Mousekids*)* which adapted the classic Mouselab tool (Payne, Bettman, & Johnson 1988). It resembles different tasks applied in adult decision making research, such as probabilistic inferences (e.g., Gigerenzer, Hoffrage & Kleinbölting, 1991), probabilistic category learning (e.g., Lagnado, Newell, Kahan, & Shanks, 2006), and advice taking (e.g., Harvey & Fisher, 1997). Participants recurrently chose between two houses (i.e., decision options) in which either a treasure (i.e., positive outcome) or a spider (i.e., negative outcome) was hidden. Decision makers could inspect the predictions of three different animals (i.e., the cues), which were correct with a certain probability (i.e., cue validity, here p = .50, .66, .83). Participants were informed about cue validities in advance, and could utilize this information to maximize outcomes. Replicating the procedure from previous research (Betsch et al., 2014), participants were confronted with different patterns of cues' predictions. Although not the focus of our analysis, this allowed us to differentiate between probability-based decision strategies (specifically WADD and LEX, see section 1.4.2). Each participant was confronted with each cue pattern eight times in a fixed order resulting in 24 choices (Figure 1). The decision environment was non-compensatory: In order to make good decisions, simple heuristics were sufficient, that is, only the predictions of the high valid cue had to be considered (Lexicographic rule, Fishburn, 1974; take-the-best, Gigerenzer & Goldstein, 1996).

	Option 1	Option 2
Cue 1 p = .50	-	+
Cue 2 p = .66	-	-
Cue 3 p = .83	+	-

(A) Type 1

	Option 1	Option 2
Cue 1 p = .50	-	+
Cue 2 p = .66	+	+
Cue 3 p = .83	+	-

(B) Type 2

	Option 1	Option 2
Cue 1 p = .50	+	-
Cue 2 p = .66	+	-
Cue 3 p = .83	-	+

(C) Type 3

Figure 1. Types of prediction patterns. Each pattern was used four times in the depicted manner and four times in a mirrored version. Types were alternated starting with Type 1. In the first pattern, the low validity cue contradicts the high validity cue's prediction, whereas the medium validity cue is indifferent and predicts a negative outcome for both options. In the second pattern, the medium validity cue predicts a positive outcome for both options and the low and high validity cues contradict each other. In the last pattern, both lower validity cues contradict the high valid cue's prediction.

Following each decision, participants either received selective outcome feedback—that is, they were informed about the consequence of the chosen option—or no feedback. The probabilistic feedback structure during choices was adopted from previous research (Betsch et al., 2014) and matched each cue's validity, that is, for example the prediction of a positive outcome by the high valid cue was correct 83% of the trials while the prediction by the low valid cue was correct in only 50%. Either one or both options could result in positive outcomes in each trial. The feedback schedule further reinforced options equally, that is, each option provided positive and negative outcomes equally often. Thus, no option was superior. Normatively, the decisions should be based on the stated probabilistic information, that is, the cue validities and the encountered cue predictions.

Choice Strategies in Children

In addition to group-level analysis, the inspection of individual strategies can reveal further within-age group variability. Therefore, we analyzed individual choice behavior using an outcome-based strategy classification method to test for a variety of choice strategies in children (Bröder & Schiffer, 2003). In addition to the Lexicographic rule (LEX) predicts reliance on the high valid cue's prediction only, we considered the weighted additive rule

(WADD: integrating weighted predictions of all cues; e.g., Payne et al., 1988)[1] and an option-based win-stay-lose-shift rule (WSLS) for the feedback conditions. According to this decision rule, the cues' predictions are ignored and recent feedback predicts choices. A decision maker would thus stay with one option until it fails once and then switch to the other option until the next failure. This decision rule fits well with children's tendency to switch after experiencing a negative outcome and represents a simple heuristic utilizing recent feedback. Additionally, we considered the strategy of following the low valid, but first selected cue (LVC, see section 2.1.2), since previous research has demonstrated that children sometimes rely on how much they like a cue rather than on its validity (Betsch & Lang, 2013). Finally, we tested a simple alternation strategy, consisting of systematically switching between options regardless of feedback (SW, i.e., choosing the option rejected in the last trial; Brainerd, 1981).

The considered strategies differ with regard to the amount and type of information taken into account. Two strategies are based on cue validities (LEX, WADD); one is based on recent feedback alone and can only be applied in feedback conditions (WSLS). Two rely on invalid information and neglect cue validities as well as feedback (LVC, SW). Based on prior findings (Betsch et al., 2016), we expected large differences between age groups, with the vast majority of adults, some proportion of elementary schoolers and only very few, if any, preschoolers using probability-based strategies when feedback is available. Instead we expected children to mal-adaptively use recent feedback for choices. On a group level, this would be indicated by over-responsiveness to negative outcomes, and on an individual level by the use of WSLS as a choice strategy. Without feedback we expected children of both age groups to systematically rely on cue validities in their choices. However, if probability-based strategies are equally rare, interference by feedback can be ruled out as a cause for children's neglect of cue validities.

Information Presentation

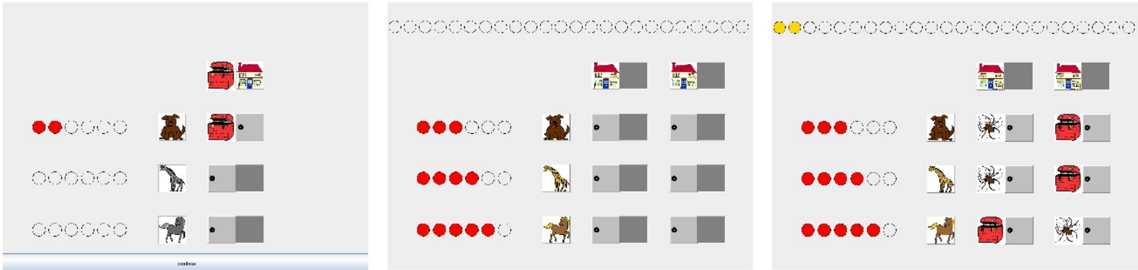
We conducted two studies that used the same factorial design and age groups but differed in the presentation of the cues' predictions. In Study 1 we used a closed information-board: Predictions were hidden and active information search was required (see Figure 2B).

---

[1] For WADD all cue predictions should be weighted by cue validities (CV) and added up to determine the expected value for each option. In order to achieve a non-compensatory decision environment, we included a cue whose predictions were at chance level and could therefore be ignored. However, whether CVs are used as weights or must be transformed is discussed controversially (e.g., correcting for chance level, CV−.5, Hilbig & Moshagen, 2014; Jekel, Glöckner, Fiedler, & Bröder, 2012; or log-transforming, log(CV/1−CV), Lee, 2014). For determining predictions of WADD we used CVs without any correction as weights, thus WADD would predict following the HVC in the Type 1 and Type 2 predictions patterns but not in Type 3 where both other cues overrule predictions of the HVC (Rieskamp & Hoffrage, 1999). Based on previous research (Betsch et al., 2014) we assumed that a substantial proportion of adults would apply an uncorrected form of WADD in this decision environment although it contradicts normative expectations, see also Footnote 2 and 3.

Individuals could access predictions by sequentially opening doors in the board. This format is most common in comparable studies conducted with adults (e.g., Payne et al., 1988), and allowed to investigate children's information search behavior. We tested, whether the presence of feedback affected search behavior. Specifically, we investigated whether children's search behavior was biased by over-responsiveness to negative feedback and systematically guided by cue validities without feedback.

However, prior studies (Betsch et al., 2014; Glöckner & Betsch, 2008) have shown that the presentation format influences strategy application. Thus, in Study 2 we used an open board where all predictions are uncovered right from the start.



(A) (B) (C)

Figure 2. Mousekids during cue observation and choices. Figure shows a screen during the observation of cue validities (A), and choices in the closed presentation format of Study 1 (B) and of the open presentation format of Study 2 (C). Prior to choices, probabilistic information is presented as validities of different cues. Specifically, participants observe relative frequencies of correct predictions of three cues. Each animal predicts outcomes six times; for example, the first animal predicts a positive outcome (i.e., treasure). Smart points at the left of each cue serve as a visual aid of cue validity during decision making (B, C). Participants then choose between two options and can inspect cue predictions. In the closed presentation format, participants uncover cue predictions before making a choice. In the open presentation format, cue predictions are uncovered. For example, participants can assess directly that the low valid cue predicts a negative outcome (spider) for the option on the left and a positive outcome (treasure) for the option on the right. At the top, a treasure point is granted for each treasure found in feedback conditions.

Study 1: Closed Board

Method

The study implemented a 3 (age group: preschoolers vs. elementary schoolers vs. adults) × 2 (feedback: yes vs. no) full factorial design. Pattern was varied within subjects (Type 1 vs. Type 2 vs. Type 3).

Sample. The sample consisted of 69 6-year-old preschoolers (28 female, *M* = 67.9 months, *SD* = 7.3), 56 9-year-old elementary schoolers (25 female, *M* = 107.4 months, *SD* = 5.7), and 56 adults (46 female, *M* = 281.18 months, *SD* = 44.66) who were randomly assigned to feedback conditions. Children were recruited from day care centers and schools with different socio-economic backgrounds from a middle-sized city in Germany. The recruitment procedure, selection of age groups, and target sample size followed considerations from prior work that used the same experimental tool (Betsch et al., 2014). Children took part in the study during a workshop at the lab and were randomly assigned to experimenters. Parents consented to participation prior to data collection. Adult participants were students of different majors.

Eight preschoolers were excluded from analyses because they did not complete the study. Five preschoolers and two elementary schoolers were excluded because they did not identify the high valid cue correctly before choices. The final sample consisted of 56 preschoolers (23 female, M = 68.0 months, SD = 7.3), 54 elementary schoolers (24 female, M = 107.5 months, SD = 5.7) and 56 adults.

Procedure. Each child was supervised by a trained experimenter who first introduced the child to the game's purpose and the cues. The child selected three animals as cues. The first chosen animal was placed on the top row, the last chosen on the bottom row (see Figure 2A). Thus, the favored animal, that is, the first chosen one, was always associated with the lowest validity.

Observing cue validities. The experimenter told the child: “Now we will find out how smart the animals are. The animals will tell you whether there is a treasure or a spider hidden in this house. But the animals are not always right. We will find out how often they are right”. She then started with the low valid cue. The experimenter opened the door next to the animal, interpreted the picture and then opened the house on top of the screen, interpreting the picture again and explained: “The dog knew that there was a treasure in the house. It was right and gets a smart point.” The experimenter clicked on the first of the smart points next to the cue, which turned red. In subsequent trials, the experimenter continued to verbalize the content of each door and house. Each time, the child was asked to indicate whether the animal should get a smart point. If the child did not answer correctly, the experimenter explained it once more. After six trials, the experimenter summarized the performance of the cue by stating that it had received three out of six smart points and continued with the next cue.

Finally, the child had observed each cue's prediction and the particular outcome six times. This allowed us to create three cues that correctly predict outcomes with varying probability, yet perform at least at chance level (i.e., $p \geq .5$). This is important, because otherwise, in a binary world, inverse probability could be used to infer outcomes (i.e., cue validities = .1 is equivalent to cue validity = .9). The first cue's predictions were correct in three, the second cue's predictions were correct in four, and the third cue's predictions were correct in five out of six cases. The children translated this information into smart points, which subsequently served as a cognitive aid for representing the cue's validity. To ensure that all children encoded differences in cue validities they were asked which animal was the smartest as a manipulation check.

Choice phase. The experimenter explained the game's purpose, procedure and checked in two training trials that the child understood the choice procedure and the information board matrix. To ensure that the children's goal was to maximize the number of treasures found, they were informed that treasures could be traded for gifts afterwards. Participants then made 24 decisions between two options represented by two houses (see Figure 2B). Cue predictions were covered by doors and could be uncovered before making each choice to inspect each cue's predictions for the two options. Participants could inspect as many cue predictions as they pleased and as often as they liked. The doors covering the cues' predictions stayed open for 3500 ms. In the feedback conditions, participants were informed about outcomes by opening the house and finding either a treasure or a spider. In conditions without feedback, the participants did not observe outcomes during the choice phase, but were informed about their performance at the end of the game. Participants were unaware of the number of decisions and prediction patterns they would encounter.

Afterwards, the manipulation check was assessed a second time to rule out that children forgot about cue validities during the course of the game. Children were rewarded with two to four prizes contingent on their performance. Additionally, participants answered several questions concerning the cues and their motivation during the game, which we do not address in this paper.

Procedure for adults. Following previous research (Betsch et al., 2014), we used the same procedure for adults as for children with the exception that adults were informed in advance that they served as a control group for children and would receive money according to their performance (€4 on average).

Results

Information search. Children searched less information than adults did. Specifically, preschoolers on average uncovered 3.58 of the six cue predictions (SD = 1.77), elementary schoolers uncovered 3.72 (SD = 1.59) and adults 4.74 (SD = 1.34) predictions. In accordance with this observation, a GLM ANOVA with age group and feedback condition as between factors revealed a main effect for age, $F(2, 154) = 8.49$, $p < .001$, $\eta^2_G = .10$. Withholding feedback did not significantly affect the amount of information searched in any age group, all other $ps \geq .08$.

If search behavior is guided by stated probabilistic information, the search should start on the most important information, that is, the high valid cue's prediction. We analyzed the frequency of searches starting on the high valid cue in a GLM ANOVA with age and feedback as between factors. Adults' searches started in 17.63 (SD = 9.24) of the 24 decisions on the most valid cue, while children did not systematically start their search on the high valid cue, M = 9.72, SD = 6.19; elementary schoolers, M = 11.29, SD = 6.21; $F(2, 154) = 18.25$, $p < .001$, η^2_G =.19. However, large variability in search behavior suggests that individuals in each age group applied different search strategies. Most importantly, searches did not start more often on the high valid cue without feedback in any age group, all other ps $\geq .16$.

To test whether children's over-responsiveness to recent negative feedback biased their information search, we compared how often it started at the non-chosen option after a negative and after a positive outcome in feedback conditions (OR = percentage of searches starting at the non-chosen option after losses/ percentage of searches starting at the non-chosen option after gains). If children were biased by negative feedback, they should more often start their search at the non-chosen option after failure (OR >1). This was only the case in preschoolers, OR = 1.39, 95% Bootstrap CI [1.10, 1.72]; elementary schoolers, OR = 0.94, CI [0.80, 1.07]; adults, OR = 0.90, CI [0.80, 1.00]. Comparison to preschoolers' search behavior without feedback ensured that this was indeed due to experiencing feedback, OR = 0.98, CI [0.81, 1.15], $t(53) = 2.21$, $p = .031$, $d = 0.28$.

In line with previous research, children's information search was not guided by cue validities (e.g., Betsch et al., 2014). In both child groups, information search was equally unsystematic with and without feedback. Interestingly, over-responsiveness to negative outcomes biased younger but not older children's information search.

Choices. We first analyzed choices on the aggregate level with our focus on over-responsiveness to recent outcomes in feedback conditions and an increased use of stated probabilistic information, that is, cue validities, in conditions without feedback. Secondly, we analyzed individual choice patterns, expecting children to use the feedback-based choices rule WSLS with feedback and to use probability-based choice rules substantially more frequently without feedback.

Over-responsiveness to negative outcomes in choices. To test whether children tended to stay with an option and switch between options based on recent feedback, we compared the percentage of option switches after a negative outcome to option switches after a positive outcome (OR = percentage of switches after losses / percentage of switches after gains). Over-responsiveness is reflected in increased switching after losses in feedback conditions ($OR > 1$). In adults, switches were equally likely following positive and negative outcomes, $OR = 0.96$, 95% Bootstrap CI [0.87, 1.04]. Likewise, in elementary schoolers, switches were equally frequent, $OR = 1.11$, CI [0.92, 1.38]. In contrast, preschoolers switched significantly more often after losses than after gains, $OR = 1.31$, CI [1.11, 1.52]. To ensure that this was indeed a result of the experienced outcomes and not due to a strategy independent of feedback, preschoolers' option switching without feedback served as a standard of comparison, $OR = 1.10$, CI [0.98, 1.25], $t(54) = 1.61$, $p = .017$, $d = 0.5$. Thus, only preschoolers' choices were over-responsive to negative outcomes indicating a bias by feedback that was absent in elementary schoolers.

Choices following the high valid cue (HVC). Normatively, in this environment participants should always follow the HVC. Figure 3 depicts the mean frequencies of choices following the HVC's prediction for each age group, separated for feedback conditions and prediction patterns. A GLM analysis of variance with the frequency of following the HVC as dependent variable, and age group, feedback condition (between) and pattern (within) as factors revealed large age and pattern effects as well as an age-pattern interaction, Age, $F(2, 160) = 57.46$, $p < .001$, $\eta^2_G = .22$; Pattern, Huynh-Feldt $F(1.79, 289.38) = 67.89$, $p < .001$, $\eta^2_G = .20$; Age × Pattern, Huynh-Feldt $F(3.58, 288.48) = 18.65$, $p < .001$, $\eta^2_G = .12$; all other ps $\geq .400$.

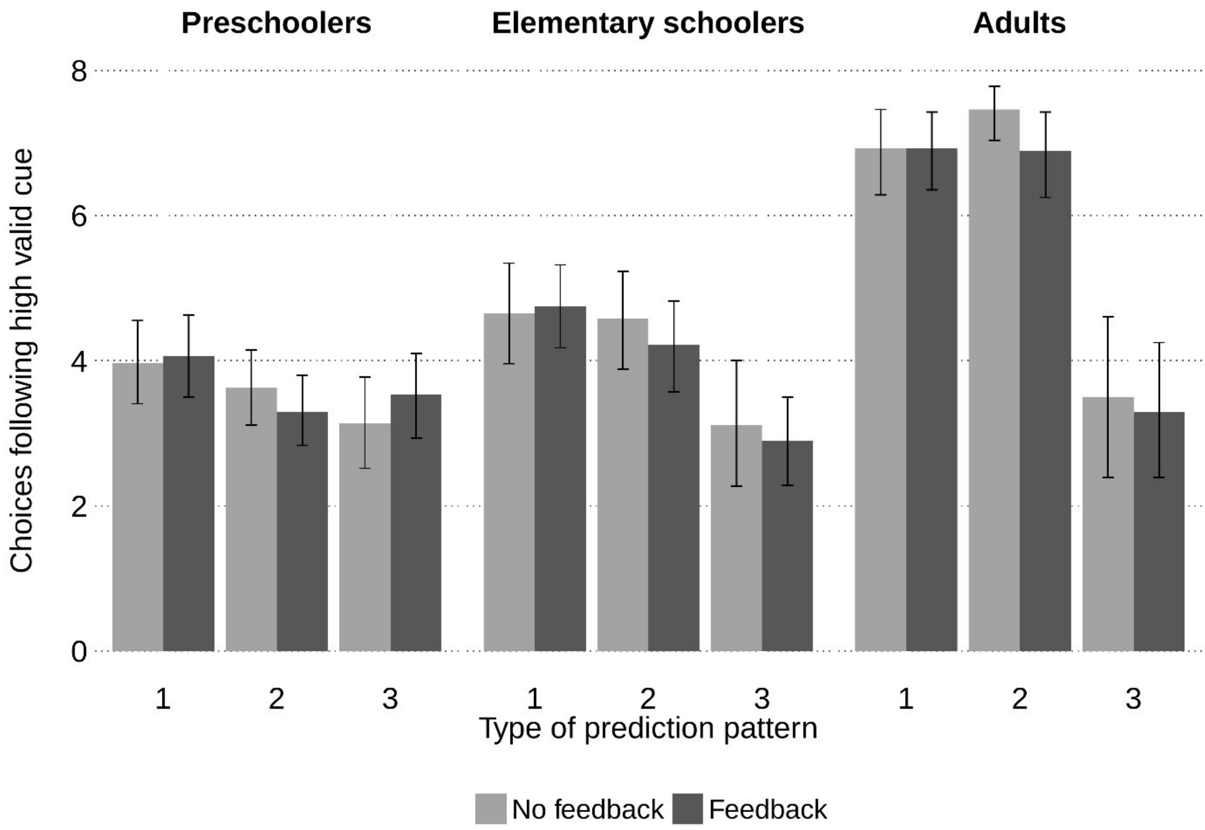


Figure 3. Mean frequencies of choices in line with the high valid cue's predictions for each pattern type and feedback condition in Study 1. Chance level is four out of eight choices. Bars represent 95% bootstrap confidence intervals.

Adults approximated normative decision making in Type 1 and Type 2 prediction patterns with choices following the HVC's predictions in 7 out of 8 cases. However, when both other cues contradicted the HVC this rate dropped to about half the cases. Similarly, elementary schoolers followed the HVC's prediction less often in Type 3 patterns and chose generally less systematically than adults but still above chance level in the first two prediction patterns. Preschoolers, in contrast, performed at chance level in each prediction pattern.

Crucially, children without feedback did not follow the HVC's prediction more often, neither in any of the prediction patterns nor when analyzed over all 24 choices, preschoolers, $M_{\text{No feedback}} = 10.74$, $SD = 2.68$, $M_{\text{Feedback}} = 10.96$, $SD = 3.11$; elementary schoolers, $M_{\text{No feedback}} = 12.34$, $SD = 3.88$, $M_{\text{Feedback}} = 11.86$, $SD = 3.09$; adults, $M_{\text{No feedback}} = 17.89$, $SD = 4.05$, $M_{\text{Feedback}} = 17.11$, $SD = 4.02$; all ps $\geq .83$. Thus, the assumption that feedback prevented children from choosing in line with normative expectations was not supported.

Individual choice strategies. To account for individual differences, we classified participants according to their choice behavior over all patterns using an outcome based strategy classification method (Bröder & Schiffer, 2003). Predictions were derived from five different choice models (LEX; WADD; WSLS; SW; LVC, see method section). For each individual, we calculated the likelihood of the observed choice pattern under each of these models, assuming that strategies are not applied flawlessly but with a constant error maximizing the likelihood. Individuals were classified to a choice model, if their choices fitted the model predictions perfectly or if (a) the likelihood for the classified strategy was higher than for any other strategy and at least twice as high as for a Random Model ($OR > 2$), and (b) at least 66% of choices were successfully predicted by the model. Otherwise participants were classified to the Random Model. Individuals with equal likelihoods for two strategies remained unclassified. Error rates were allowed to vary over participants. Table 1 shows the results of the classification and the mean error rates for each age group and strategy.

The individual strategy classification largely confirmed aggregate findings. Contrary to our expectations, WSLS was not a common choice rule in either child age group, and application of probability-based strategies was not more prevalent without feedback. Most children's choices were captured best by the Random Model (63% in preschoolers, 50% in elementary schoolers), only a minor percentage of children applied a choice rule systematically (16% in preschoolers, 30% in elementary schoolers). Intriguingly, all preschoolers that followed any strategy at all relied on invalid information with a simple switching rule being most common (11%). Elementary schoolers, on the other hand, while employing invalid strategies as often as preschoolers (17%), also and almost as often applied probability-based strategies (13%).

While the majority of adults were classified as users of probability-based strategies, 30% were classified to the Random Model. This was mainly due to inconsistent behavior by these adults in decisions with the Type 3 pattern.[2] Nonetheless, the prevalence of probability-based decision strategies was strongly determined by age, $\chi^2(2, N = 166) = 61.42, p < .001$, Cramer's $V = .61$. Probability-based strategies were mainly used by adults (61%), rarely by elementary schoolers (13%), and not found in preschoolers. Complementary, invalid choice rules were not found in adults and were observed equally often in preschoolers (16%) and elementary schoolers (17%), $\chi^2 (1, N = 110) =.007, p =.93$.

---

[2] The vast majority of adults classified to the Random Model did not choose randomly in all types of predictions patterns. Only in Type 3 patterns their choices followed the HVC at chance level, $Mdn = 4$, interquartile range (IQR): 2.5-6; while in Type 1 and 2 they were closely aligned with normative choice behavior, $Mdn_{\text{Type1}} = 7$, IQR: 5-8; and $Mdn_{\text{Type2}} = 8$, IQR: 6-8 (see also Footnote 1). However, since the classification method assumes a constant application error in all prediction patterns, and therefore cannot account for deviations in only one pattern, these adults were not classified to any of the choices models.

Table 1

Strategy classification in Study 1

	No feedback		Feedback		Overall		
Preschoolers	*n* = 27		*n* = 29		*n* = 56		
	n	%	*n*	%	*n*	%	Error
LEX	—	—	—	—	—	—	
WADD	—	—	—	—	—	—	
LVC	—	—	2	6.9	2	3.6	.29
SW	3	11.1	3	10.3	6	10.7	.18
WSLS			1	3.4	1	1.8	.21
Random	15	55.6	20	69.0	35	62.5	
Unclassified	9	33.3	3	10.3	12	21.4	
Elementary schoolers	*n* = 26		*n* = 28		*n* = 54		
	n	%	*n*	%	*n*	%	Error
LEX	1	3.8	1	3.6	2	3.7	.15
WADD	2	7.7	3	10.7	5	9.3	.21
LVC	1	3.8	—	—	1	1.9	.21
SW	4	15.4	3	10.7	7	13	.15
WSLS			1	3.6	1	1.9	.25
Random	15	57.5	12	42.9	27	50.0	
Unclassified	3	11.5	8	28.6	11	20.4	
Adults	*n* = 28		*n* = 28		*n* = 56		
	n	%	*n*	%	*n*	%	Error
LEX	7	25.0	7	25.0	14	25.0	.08
WADD	10	35.7	10	35.7	20	35.7	.07
LVC	—	—	1	3.6	1	1.8	.25
SW	—	—	1	3.6	1	1.8	.17
WSLS			—	—	—	—	
Random	8	28.6	9	32.1	17	30.4	
Unclassified	3	10.7	—	—	3	5.4	

Note. Unclassified = individuals who could not be classified in one of the considered strategies; Random = random choice. Probability-based choice rules: LEX = Lexicographic rule, WADD = weighted-additive rule. Invalid choice rules: LVC = following low valid cue; SW = switching rule. Feedback-based choice rule, only applicable in feedback conditions: WSLS = win-stay-lose-shift rule. Error = mean error rate, that is, the mean proportion of trials in which strategy inconsistent choices were observed for the corresponding strategy and age group. For example, the error rate of .08 in adults being classified as LEX users means that on average strategy-incongruent choices were observed in 8%, that is, in less than two out of 24 choices. In case of *n* = 1, absolute values are displayed.

Discussion

In prior studies, no preschoolers and only a small proportion of elementary schoolers utilized stated probabilistic information presented as cue validities in their search and choice behavior (Betsch et al., 2014). We suspected that choice feedback led children to rely on recent outcomes rather than cue validities and could thus be the reason for the observed probability neglect. If this assumption were true, over-responsiveness to negative outcomes in search and choice behavior and systematic use of feedback in terms of WSLS as a choice rule should be observed. Withholding feedback should lead to an increased utilization of cue validities in children's searches and choices. Study 1 provided some evidence of the former but no indication of the latter. Preschoolers showed over-responsiveness in search and choice behavior; children's behavior was not increasingly guided by cue validities without feedback. We address both results in the general discussion. However, we first wanted to corroborate the findings and rule out that expected differences in utilization of cue validities were disguised by the presentation format of the information board.

Specifically, the application of choice strategies can be hampered in a closed information board. The predictions of the cues need to be looked up sequentially and kept in memory until a decision is made. Such an environment binds cognitive resources and, hence, may impede successful strategy implementation (Glöckner & Betsch, 2008). This is especially the case in children who have demonstrated significant deficits in searching for relevant information in information board paradigms until late elementary school age (Betsch et al. 2016; Davidson, 1996; Katz, Bereby-Meyer, Assor, & Danziger, 2010). For example, in order to apply a normatively correct strategy like LEX, children have to look for the prediction of the HVC first and keep it in mind until they choose an option. Even when applying WSLS, where a search of cue predictions is not necessary, the delay between choices could nevertheless hinder successful strategy application. Accordingly, the lack of usage of WSLS in feedback conditions and of probability-based strategies without feedback might have been a consequence of a decision environment that impeded application of strategies.

Study 2: Open Board

We replicated the first study in a decision environment without constrained access to information to facilitate the application of choice strategies and thereby the detection of differences in strategy application between feedback conditions.

Method

The design, stimulus material, and procedure were identical to Study 1, with the exception that the information board matrix in the choice phase was open so that all predictions could be inspected at once for every decision (see Figure 2C). Search data is therefore not available.

The sample consisted of 80 preschoolers (49 female, M = 69.2 months, SD = 5.3), 62 elementary schoolers (28 female, M = 104.4 months, SD = 4.7), and 53 adults (39 female, M = 258.59 months, SD = 31.96). Two preschoolers, one elementary schooler, and one adult had to be excluded because they did not complete the study. Seven preschoolers and one adult did not pass the manipulation check. The final sample consisted of 71 preschoolers (46 female, M = 69.1 months, SD = 5.3), 61 elementary schoolers (27 female, M = 104.4, months SD = 4.7), and 51 adults (37 female, M = 267.0 months, SD = 31.2). The recruitment procedure was the same as in Study 1.

Results

Over-responsiveness to negative outcomes in choices. Again, only preschoolers switched options more frequently after negative than after positive outcomes, OR = 1.29, 95% Bootstrap CI [1.11, 1.49]; without feedback, OR = 1.07, CI [0.99, 1.16]; $t(69) = 2.02$, $p = .003$, $d = 0.49$. Elementary schoolers in feedback conditions showed no such tendency, OR = 1.08, CI [0.95, 1.24], neither did adults, OR = 0.78, CI [0.68, 0.90].

Choices following the high valid cue (HVC). Figure 4 depicts the mean frequencies of following the HVC's predictions for every age group, feedback condition and prediction pattern. The results largely matched those in Study 1. Again, feedback did not affect choices in any age group, not in any of the patterns nor when analyzed over all 24 choices, preschoolers, $M_{\text{No feedback}} = 10.09$, $SD = 4.63$, $M_{\text{Feedback}} = 10.78$, $SD = 4.42$; elementary schoolers, $M_{\text{No feedback}} = 13.50$, $SD = 5.02$, $M_{\text{Feedback}} = 14.21$, $SD = 4.35$; and adults, $M_{\text{No feedback}} = 18.52$, $SD = 2.58$, $M_{\text{Feedback}} = 18.70$, $SD = 3.65$; all ps ≥ .40. A GLM ANOVA resulted in similarly large age and pattern effects, and a pattern-age interaction as in Study 1, Age, $F(2, 177) = 54.65$ $p < .001$, $\eta^2_G = .26$; Pattern, Huyn-Feldt $F(1.59, 281.64) = 93{,}24$, $p < .001$, $\eta^2_G = .19$; Age × Pattern, Huyn-Feldt $F(3.81, 281.64) = 42.79$, $p < .001$, $\eta^2_G = .18$, all other ps ≥ .170. Adults followed the HVC's prediction except when both other cues contradicted its prediction. The same pattern effect was observed in elementary schoolers, who followed the HVC's prediction less systematically but above chance level in the first two prediction patterns. Preschoolers performed at chance level, regardless of the prediction pattern.

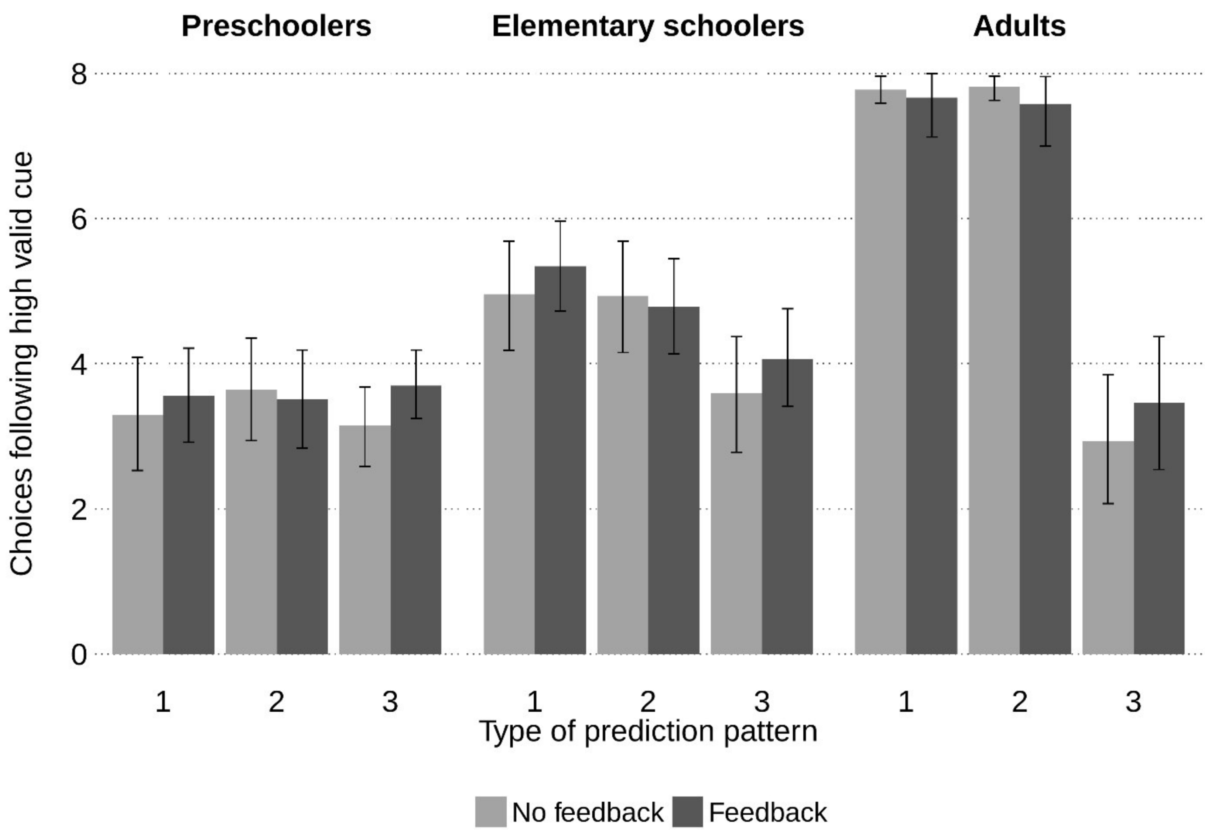


Figure 4. Mean frequencies of choices in line with predictions of the high valid cue for each pattern type and feedback condition in Study 2. Chance level is four out of eight choices. Error bars represent 95% bootstrap confidence intervals.

Individual choice strategies. The results of the individual strategy classification are shown in Table 2. Similar to Study 1, feedback did not influence the prevalence of probability-based strategies in either child age group and WSLS was used rarely.

Adults applied probability-based choice strategies to a large extend (49%, LEX, WADD). However, like in Study 1, a relatively large proportion was classified to the Random Model. Again, that was mainly due to their indifferent choice behavior in Type 3 prediction patterns (43%).[3] Among children that chose systematically, elementary schoolers used valid (16%) as well as invalid strategies (28%, LVC, SW) while preschool-aged children relied predominantly on invalid ones (48%). The prevalence of probability-based strategies was again strongly dependent on age, with 49% of adults (n = 25), and 16% of elementary schoolers (n = 10) using such strategies, but only 3% of preschoolers (n = 1), $\chi^2(2, N = 183) = 43.20$, $p < .001$, Cramer's $V = .47$. Likewise, the use of invalid strategies was most

[3] For adults classified to the Random Model, $Mdn_{\text{Type1}} = 8$, IQR: 8-8; $Mdn_{\text{Type2}} = 8$, IQR: 8-8; and $Mdn_{\text{Type3}} = 3$, IQR: 2-5.

common in preschoolers, (48%), still found at elementary school age (28%) and only sporadically present in adults (2%). However, in contrast to Study 1, invalid choice rules were more common in preschoolers than in elementary schoolers, $\chi^2(2, N = 132) = 4.64$, $p = .031$, Cramer's $V = .19$.

Discussion

In an open board, no active search processes are needed to access information. We tested whether this would foster children's systematic utilization of the provided sources of information—either stated probabilistic information, that is, cue validities, or feedback. However, the findings of Study 2 were strikingly similar to those of Study 1. Thus, we are confident that the lack of systematic utilization of cue validities or feedback is not the result of the presentation format requiring too much of children's limited cognitive resources. However, compared to Study 1 more children of each age group applied systematic choice rules, indicating that open presentation facilitated strategy application, but not utilization of probabilistic information or feedback.

Table 2

Results of strategy classification in Study 2

	No feedback		Feedback		Overall		
Preschoolers	*n* = 34		*n* = 37		*n* = 71		
	n	%	*n*	%	*n*	%	Error
LEX	—	—	1	2.7	1	1.4	.17
WADD	—	—	—	—	—	—	
LVC	3	8.8	4	10.8	7	9.9	.14
SW	16	47.1	11	29.7	27	38.0	.12
WSLS			—	—	—	—	
Random	13	38.2	17	45.9	30	42.3	
Unclassified	2	5.9	4	10.8	6	8.5	
Elementary schoolers	*n* = 32		*n* = 29		*n* = 61		
	n	%	*n*	%	*n*	%	Error
LEX	4	12.5	4	13.6	8	13.1	.12
WADD	2	6.3	—	—	2	3.3	.08
LVC	2	6.3	—	—	2	3.3	.17
SW	10	31.3	5	17.2	15	24.6	.19
WSLS			1	3.4	1	1.6	.29
Random	9	28.1	12	41.4	21	34.4	
Unclassified	5	16.5	7	24.1	12	19.7	
Adults	*n* = 27		*n* = 24		*n* = 51		
	n	%	*n*	%	*n*	%	Error
LEX	3	11.1	5	20.8	8	15.7	.04
WADD	8	29.6	9	37.5	17	33.3	.05
LVC	—	—	1	4.2	1	2.0	.25
SW	—	—	—	—	—	—	
WSLS			—	—	—	—	
Random	13	48.1	9	37.5	22	43.1	
Unclassified	3	11.1	—	—	3	5.9	

Note. Unclassified = individuals who could not be classified in one of the considered strategies; Random = random choice. Probability-based choice rules: LEX = Lexicographic rule, WADD = weighted-additive rule. Invalid choice rules: LVC = following low valid cue; SW = switching rule. Feedback-based choice rule, only applicable in feedback conditions: WSLS = win-stay-lose-shift rule. Error = Mean error rate for the corresponding strategy and age group. In case of *n* = 1, absolute values are displayed.

General Discussion

Achieving decision competence requires, as a prerequisite, that individuals develop sensitivity and responsiveness to the probabilistic relations between choices and outcomes. We employed a research paradigm in which information about these relations was provided prior to choices—predictions of cues with different validities—in addition to self-sampled feedback about the choices' consequences. Previous studies using this paradigm found that preschoolers around the age of six years are virtually insensitive to stated probabilistic information in choices while a minority of 9-year-old elementary schoolers is able to utilize it in a systematic fashion (Betsch & Lang, 2013; Betsch et al., 2014, 2016). This evidence as well as the present studies contradict findings from other domains of developmental research where the utilization of stated probabilistic information prior to school age has been observed (e.g., Pasquini et al., 2007; Schlottmann, 2001). To investigate the potentially interfering effect of experience on utilization of stated probabilistic information, we manipulated the presence of feedback in a closed and open information board paradigm. We hypothesized that experiencing decision outcomes provides an interfering second source of information and can account for children's neglect of stated probabilistic information.

The results, however, show that the removal of feedback from the decision task did not improve children's performance. Normative utilization of probability in choices prescribes to rely only on the high valid cue's predictions. In either study, though, the frequency of following the high valid cue did not differ as a function of the presence of feedback. Compared to conditions with feedback, choices did not follow the high valid cue's prediction more often without feedback in any age group; preschoolers, d_{Study1} = -0.08, 95% CI [-0.61, 0.44], d_{Study2} = -0.15, CI [-0.62, 0.31], d_{overall} = -0.12, CI [-0.47, 0.23]; elementary schoolers, d_{Study1} = 0.14, CI [-0.40, 0.67], d_{Study2} = -0.15, CI [-0.65, 0.36], d_{overall} = -0.01, CI [-0.38, 0.35]; adults, d_{Study1} = 0.19, CI [-0.33, 0.72], d_{Study2} = -0.06, CI [-0.61, 0.49], d_{overall} = 0.07, CI [-0.31, 0.45]. Individual analysis of choice patterns further showed that the prevalence of probability-based strategies was not increased without feedback and that when feedback was available children did not use a feedback-based choice strategy. Tracing information search behavior in Study 1 demonstrated additionally that younger children's over-responsiveness to negative outcomes also affected their search behavior, although searches were equally unsystematic without interfering feedback.

We conclude that children did not neglect stated probabilistic information because they preferred feedback as a source of information. Consequently, the reasons for children's probability neglect must lay elsewhere (see 4.2). Our two studies yielded three important new insights:

Over-Responsiveness to Negative Outcomes but no Systematic Use of Feedback-Based Strategies.

Children overreact to negative feedback and tend to switch responses. This behavior has been attributed to deficits in inhibitory control (Crone & van der Molen, 2007; Van Duijvenvoorde et al., 2012), and linked to the assumption, that negative consequences lead to an affective reaction to avoid this option, which must be actively overruled (Damasio, 1994, but see Brehmer, 1980). At the age of nine, children in our studies no longer showed over-responsiveness. This contrasts prior findings of over-responsiveness to negative outcomes until late school age (Cassotti et al, 2011), even when winning probabilities of options are stated explicitly (Van Duijvenvoorde et al., 2012).

Importantly, although over-responsiveness biased preschool-aged children's search and choice behavior, they did not strategically utilize recent outcomes and consistently rely on an option-based Win-Stay-Lose-Shift rule for choices, which solely considers recent feedback. On individual level, Win-Stay-Lose-Shift did not fit children's choice patterns when compared to other plausible models of choice. This apparent contradiction underlines the important distinction between data analysis at the aggregate and the individual level. It is still possible, that children applied a more sophisticated form of Win-Stay-Lose-Shift as a feedback-based strategy, for example switch and stay in relation to the option's expected value calculated over several trials (e.g., Worthy & Maddox, 2014). However, application of such a strategy is inconsistent with the overall lack of differences between feedback conditions on both aggregated and individual level.

Robustness of Probability Neglect

In line with previous findings, we observed large differences between age groups in the utilization of stated probabilistic information for choices in both studies. The majority of adults used strategies based on stated probabilistic information such as the Lexicographic or Weighted Additive Rule consistently while only few elementary schoolers and only one preschooler did so. Previous studies with the same paradigm have found that such probability neglect is only marginally affected by variations of the decision environment such as information search constraint (Betsch et al., 2016) or lure information (Betsch & Lang, 2013). The present findings further underline the robustness of probability neglect in children's risky decision making. Across two decision environments—one requiring an active search for cue predictions, the other displaying all cue predictions without search—preschool-aged children at the age of six were unable to utilize stated probabilistic information to adapt choices, whereas 9-year-old elementary schoolers were partly able to do so. Further, withholding feedback had no facilitating effect on the utilization of stated probabilistic information at the aggregate or the individual level. Thus, probability neglect is not a consequence of the mixed-source paradigm which offers feedback as a second and possibly preferred source of information.

Our findings contribute to the complex picture of children's ability to utilize stated probabilistic information for decisions and are in line with findings from gambling studies that highlight children's deficits until late school age (Levin & Hart, 2003; Levin, Weller, et al., 2007). However, there is also contradicting evidence stemming from selective trust tasks (Pasquini et al., 2007), judgments tasks (Schlottmann, 2001), and low-complex experienced-based gambling tasks (Bunch et al., 2007), in which utilization at preschool age was observed. Although these paradigms differ much in terms of information presentation and complexity, they all share that probabilities are directly assigned to options. The challenge for children is to make advantageous choices between these options. In a probabilistic inference task, however, the relations between options and outcomes are moderated by cues, to which these probabilities are assigned. Therefore, they pose a greater challenge to children's conceptual understanding of probabilities, which might overburden them and explain the contradictory results. Probability neglect in children might therefore strongly depend on specifics of the probabilistic environment created by the research paradigm.

Development of Decision Strategies

Our results corroborate prior findings regarding the prevalence of probability-based strategies in individuals of different age groups (e.g., Betsch et al., 2014; Mata et al., 2011). At the age of six, children do not rely on probability-based choice rules, at the age of nine this ability is still emerging. This is in line with research from other areas that suggests a strong developmental improvement in cognitive strategy use during that period (see Björklund & Causey, 2018, for an overview). Yet, while most children neglected both sources of potentially valid information—stated probabilistic information and feedback—many nonetheless did not chose randomly between options. Instead, they systematically relied on invalid information, that is, information that was not useful to maximize choice outcomes, such as, which cue was liked the most and which option had been chosen in the preceding decision. This is not only interesting from a theoretical point of view, but holds important implications for the improvement of decision quality in children. Rather than just introducing or teaching appropriate strategies, children's current strategies must be directly addressed and revealed as inappropriate.

Implications of Individual Strategy Classification

In each age group we observed rather large proportions of individuals that could not be classified to a decision strategy (i.e., Random). This finding has two important implications. First, it shows that adults' decision behavior is a more appropriate benchmark for children's decision than normative standards, that is, ideal choices under all circumstances. Second, though, it points to a limitation of our strategy classification method. Many adults remained unclassified because they choose indifferently in Type 3 prediction patterns. We cannot rule out, however, that these participants used different strategies for different prediction patterns, which the applied method cannot detect or might have used decision strategies not considered a priori. For example, participants might have applied an Equal Weight Rule (e.g., Payne et al., 1988), which prescribes to follow the majority of predictions and, in case of a tie, to guess between options. This would lead to guessing in two thirds of the investigated choices and choices in line with the majority in Type 3 prediction patterns. We cannot rule out that single participants might have used this strategy; however, choices in Type 3 prediction patterns are not in line with this model[4]. It is therefore unlikely, that it was widely-used in any age group.

Further, it is conceivable, that participants classified to WADD instead used a compound-strategy that favors the option with more positive predictions and only in case of indifferences considers cue validities. Although, this strategy would lead to exactly the same choices as WADD, the underlying cognitive steps are quite different. Nonetheless, it would still rely on the probabilistic cues. As our main focus was to differentiate choice rules based on on which kind of information was used—probabilities, feedback, or invalid information—the possibility of a compound-strategy does not alter our interpretation of the results.

Beneficial Aspects of Feedback

This paper aimed to investigate potentially interfering effects of feedback on children's utilization of probabilities. However, feedback does not have to be harmful. A study by Mata et al. (2011) suggests that 9-year-olds can adapt decision strategies via feedback learning and our own results demonstrate that even 6-year-olds are generally

[4]In two thirds of all choices, that is, prediction patterns Type 1 and 2, Equal Weight and Random predict guessing, and can therefore not be reliably distinguished from each other on an individual level. In Type 3 prediction patterns, Equal Weight prescribes following the majority (e.g., Option 1 in Figure1). In Study 1, 9% of preschoolers, 11% of elementary schoolers, and 29% of adults followed the majority systematically in choices in this pattern type (i.e., 6 out of 8). In Study 2, 6% of preschoolers, 12% of elementary schoolers, and 20% of adults followed the majority in at least 75% of choices. Accordingly, the proportion of individuals who may have applied an Equal Weight Rule is low, especially in child groups.

responsive to feedback in probabilistic inferences, even if this resulted in over-responsiveness to negative outcomes. Unfortunately, in order to reveal such beneficial aspects of feedback, the number of decisions has to be large enough to enable learning in all age groups while learning rates may vary with age. Since the number of decision was limited in our studies, no conclusion about potential beneficial aspects of feedback can be drawn.

Conclusion

So far, research had demonstrated that children up to late school age fail at utilizing probabilities in a complex environment with multiple predictive cues (Betsch & Lang, 2013; Betsch et al., 2014, 2016). Our studies replicated these findings and ruled out the interference of feedback information as a cause for this probability neglect. Below school age, children's choices are not ruled by probabilities. If at all, irrelevant features of the task guide their choices. At elementary school age, utilization of probabilities is only emerging, yet not consolidated in all children. At that age, children are still far from being competent when decisions are risky and complex.

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. https://doi.org/10.2307/1131034

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.110–134). Hillsdale, NJ: Erlbaum

Barron, G., & Erev, I. (2003). Small feedback based decisions and their limited correspondence to description based decisions. *Journal of Behavioral Decision Making, 16,* 215–233. https://doi.org/10.1002/bdm.443

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology*, *115*, 108–125. https://doi.org/10.1016/j.jecp.2012.11.003

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica*, *153*, 74–86. https://doi.org/10.1016/j.actpsy.2014.09.008

Betsch, T., Lehmann, A., Lindow, S., Lang, A., & Schoemann, M. (2016). Lost in Search: (Mal-)adaption to probabilistic decision environments in children and adults. *Developmental Psychology, 52*, 311–325. https://doi.org/10.1037/dev0000077

Björklund, D. F., & Causey, K. B. (2018). *Children's thinking: cognitive development and individual differences.* Thousand Oaks, CA: Sage Publications.

Boyer, T. W. (2007). Decision making processes: Sensitivity to sequentially experienced outcome probabilities. *Journal of Experimental Child Psychology, 97*, 28–43. https://doi.org/10.1016/j.jecp.2006.12.008

Brainerd, C. J. (1981). Working memory and the developmental analysis of probability judgment. *Psychological Review, 88,* 463–502. https://doi.org/10.1037/0033-295X.88.6.463

Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica, 45*, 223-241. http://doi.org/10.1016/0001-6918(80)90034-7

Bröder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making*, *16*(3), 193–213. https://doi.org/10.1002/bdm.442

Bunch, K. M., Andrews, G., & Halford, G. S. (2007). Complexity effects on the children's gambling task. *Cognitive Development*, *22*(3), 376–383. https://doi.org/10.1016/j.cogdev.2007.01.004

Cassotti, M., Aïte, A., Osmont, A, Houdé, O., & Borst, G. (2014). What we learned about the processes involved in the Iowa Gambling Task. *Frontiers in Psychology*, 5, 915. http://dx.https://doi.org/ 10.3389/fpsyg.2014.00915

Cassotti, M., Houdé, O, & Moutier, S. (2011). Developmental changes of win-stay and loss-shift strategies in decision making. *Child Neuropsychology, 17*, 400–411. http://dx.https://doi.org/10.1080/09297049.2010.547463

Crone, E. A., Bunge, S. A., Latenstein, H., & van der Molen, M. W. (2005). Characterization of children's decision making: Sensitivity to punishment frequency, not task complexity. *Child Neuropsychology, 11*, 245–263. http://dx.https://doi.org/10.1080/092970490911261

Crone, E. A., & van der Molen, M. W. (2007). Development of decision making in school-aged children and adolescents: Evidence from heart rate and skin conductance analysis. *Child Development, 78,* 1288–1301. https://doi.org/10.1111/j.1467-8624.2007.01066.x

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Grosset/Putnam.

Davidson, D. (1996). The effects of decision characteristics on children's selective search of predecisional information. *Acta Psychologica, 92*(3), 263–281. https://doi.org/10.1016/0001-6918(95)00014-3

Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. G. (2014). A meta-analysis on age differences in risky decision making: Adolescents versus children and adults. *Psychological Bulletin, 141*(1), 48–84. https://doi.org/10.1037/a0038088

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*, 380-417. https://doi.org/10.1037/h0053870

Eppinger, B., Mock, B., & Kray, J. (2009). Developmental differences in learning and error processing: Evidence from ERPs. *Psychophysiology*, *46*(5), 1043–1053. http://https://doi.org/10.1111/j.1469-8986.2009.00838.x

Fishburn, P. C. (1974). Lexicographic orders, utilities and decision rules: A survey. *Management Science, 20,* 1442–1471.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650–669. https://doi.org/10.1037/0033-295X.103.4.650

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98*, 506–528. https://doi.org/10.1037/0033-295X.98.4.506

Girotto, V., & Gonzalez, M. (2008). Children's understanding of posterior probability. *Cognition, 106*, 325–344. https://doi.org/10.1016/j.cognition.2007.02.005

Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1055–1075. https://doi.org/10.1037/0278-7393.34.5.1055

Harvey, N. and Fischer, I. (1997). Taking advice: Accepting help, improving judgment and sharing responsibility. *Organizational Behavior and Human Decision Processes, 70*, 7–33. https://doi.org/10.1006/obhd.1997.2697

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13*(12), 517–523. https://doi.org/ 10.1016/j.tics.2009.09.004

Hilbig, B. E., & Moshagen, M. (2014). Generalized outcome-based strategy classification: Comparing deterministic and probabilistic choice models. *Psychonomic Bulletin & Review, 21,* 1431–1443. http://dx.https://doi.org/10.3758/s13423-014-0643-0

Hogarth, R. (2001). *Educating intuition.* Chicago: University of Chicago Press.

Hogarth, R.M. & Soyer, E. (2011). Sequentiylly Simulated Outcomes: Kind Experience Versus Nontransparent Description. *Journal of Experimental Psychology: General, 140*, 434–463. https://doi.org/10.1037/a0023265

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. https://doi.org/10.1111/j.1467-7687.2007.00621.x

Jekel, M., Glöckner, A., Fiedler, S., & Bröder, A. (2012). The rationality of different kinds of intuitive decision processes. *Synthese, 189*, 147–160. https://doi.org/10.1007/s11229-012-0126-7

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–291. https://doi.org/10.2307/1914185

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, *134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Katz, I., Bereby-Meyer, Y., Assor, A., & Danziger, S. (2010). Children's adaptive pre-decisional search behavior: Effects of memory and number of alternatives. *Journal of Economic Psychology*, *31*, 17–24. https://doi.org/ 10.1016/j.joep.2009.09.003

Kerr, A., & Zelazo, P. D. (2004). Development of "hot" executive function: The children's gambling task. *Brain and Cognition*, *55*(1), 148–157. https://doi.org/10.1016/S0278-2626(03)00275-6

Lagnado, D. A., Newell, B. R., Kahan, S. and Shanks, D. R. (2006). Insight and strategy in multiple-cue probability learning. *Journal of Experimental Psychology: General, 135*, 162–183. https://doi.org/10.1037/0096-3445.135.2.162

Lee, M. D. (2014). Bayesian outcome-based strategy classification. *Behavior Research Methods, 48*(1), 29–41. https://doi.org/10.3758/s13428-014-0557-9

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*(5), 397–413. https://doi.org/10.1002/bdm.453

Levin, I. P., Hart, S. S., Weller, J. A., & Harshman, L. A. (2007). 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. https://doi.org/10.1002/bdm.552

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.

Mata, R., von Helversen, B. & Rieskamp, J. (2011). When easy comes hard: The development of adaptive strategy selection. *Child Development*, *82*(2), 687–700. https://doi.org/10.1111/j.1467-8624.2010.01535.x

Newell, B. R., & Rakow, T. (2007). ©*Psychonomic Bulletin and Review, 14*(6), 1133–1139. https://doi.org/1133–1139.10.3758/BF03193102

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, *43*(5), 1216–1226. https://doi.org/10.1037/0012-1649.43.5.1216

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Piaget, J., & Inhelder, B. (1951). *The origin of the idea of chance in children*. London: Routledge & Kegan Paul.

Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics, and how can we tell? In G. Gigerenzer, P. M. Todd, & ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 141-167). New York: Oxford University Press.

Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *Child Development, 72*, 103–122. https://doi.org/10.1111/1467-8624.00268

Schlottmann, A., & Tring, J. (2005). How children reason about gains and losses: Framing effects in judgment and choice. *Swiss Journal of Psychology*, *64*(3), 153–171. https://doi.org/10.1024/1421-0185.64.3.153

Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10*(1), 104–109. https://doi.org/10.1111/j.1467-7687.2007.00571.

Van den Bos, W., Güroğlu, B., van den Bulk, B. G, Rombouts, S. A. R. B., & Crone, E. A. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience*, *3*, 52. https://doi.org/10.3389/neuro.09.052.2009

Van Duijvenvoorde, A. C. K, Jansen, B. R. J., Bredman, J. C., & Huizenga, H. M. (2012). Age-related changes in decision making: Comparing informed and noninformed situations. *Developmental Psychology*, *48*(1), 192–203. https://doi.org/10.1037/a0025601

Worthy, D. A., & Maddox, W.T. (2014). A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to W.K. Estes, *Journal of Mathematical Psychology*, *59*, 41–49. https://doi.org/10.1016/j.jmp.2013.10.001

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(13), 5012–5015. https://doi.org/10.1073/pnas.0704450105

Zmyj, N., Buttelmann, D., Carpenter, M. J., & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology, 106*(4), 208–220. https://doi.org/ 10.1016/j.jecp.2010.03.002

Article 2

Children Use but Can Overcome Non-adaptive Decision Strategies

Children Use but Can Overcome Non-adaptive Decision Strategies

Anna Lang and Tilmann Betsch
University of Erfurt, Germany

Author Note

Anna Lang, Department of Psychology, University of Erfurt, Germany; Tilmann Betsch, Department of Psychology, University of Erfurt, Germany.

This research was supported by a grant from the Deutsche Forschungsgemeinschaft to the second author (Grant No. BE 2012/11-1).

Correspondence concerning this article should be addressed to Anna Lang, Department of Psychology, University of Erfurt, Nordhaeuser Strasse 63, 99089 Erfurt, Germany. E-mail: anna.lang@uni-erfurt.de

Declarations of interest: none

Word count: 3761

Abstract

Previously, children aged nine and younger seemed to decide randomly without any systematic plan in probabilistic decision environments. Research focused on detecting adaptive decision strategies that consider probabilistic information embedded in the environment and are used by adults. However, it ignored non-adaptive decision strategies that focus on irrelevant information and might be used by children. The present study assessed both adaptive and non-adaptive decision strategies in 6-year-olds, 9-year-olds and adults to investigate how often, accurately and persistently children apply them. Participants repeatedly made decisions in a probabilistic environment that either provided no feedback, selective or full feedback about strategies' performances. Results show that half of the children in each age group applied a decision strategy accurately. Six-year-olds predominantly used non-adaptive strategies, while 9-year-olds used adaptive strategies as well. Full feedback positively affected strategy use: children abandoned non-adaptive strategies, or applied them less consistently. Results suggest that children are able to systematically apply decision strategies very early in life. Similar to other cognitive areas, they initially use inefficient strategies, but overcome them with increasing age and experience.

Keywords: decision strategy, strategy development, decision feedback, adaptive decision making

Children Use but Can Overcome Non-adaptive Decision Strategies

Successful probabilistic decision making requires consistently applying decision strategies that consider probabilities embedded in the environment. The Weighted Additive strategy weights outcomes with their probability and maximizes decision outcomes (Payne, Bettman, & Johnson, 1988). In some probabilistic environments, the simpler Lexicographic strategy that considers only the most probable outcome yields equally good outcomes (Fishburn, 1974; Gigerenzer & Goldstein, 1996). In previous developmental research, children up until late elementary school age failed to apply such adaptive strategies. Instead all 6-year-old children and the majority of 9-year-old children seemed to decide randomly without any systematic plan (Betsch & Lang, 2013; Betsch, Lang, Lehmann, & Axmann, 2014, Betsch, Lehmann, Lindow, Lang, & Schoemann, 2016). This research however, concentrated on detecting strategies derived from normative decision theory, such as the above-mentioned; strategies that deviated from normative expectations were ignored. More recent research suggests that children do apply strategies, but use strategies non-adaptive to the decision environment, such as option alternation and following favored but less accurate probabilistic cues (Lang & Betsch, 2018). Assessment of individuals' strategies, however, was not reliable enough to investigate the dimensions of strategic change, such as how often, how accurately, and how persistently children used these strategies. In this paper, we thoroughly measure adaptive and non-adaptive decision strategies to show that non-adaptive strategies are widespread in 6-year-olds, and become less prevalent with age, when, in return, adaptive strategies become more prevalent.

We further investigated how persistently children apply non-adaptive strategies. Monitoring feedback of strategies' performances can initiate changes in strategy use. Decision feedback can be used to update erroneous strategy expectations and eventually elicit shifting to better strategies (Rieskamp & Otto, 2006; Shrager & Siegler, 1998). Children might thus overcome non-adaptive decision strategies when they monitor their performance. To test this assumption, we assessed children's decision strategies in three different environments: One offered no immediate feedback about decision outcomes, therefore did not allow monitoring strategy performance and might have invited non-adaptive decision strategies. The second immediately fed back decision outcomes. Mimicking most real-world decisions, the feedback was selective, that is, it provided the outcome of the actual and not the counterfactual decision. With selective feedback, belief updating can remain fragmentary and biased by the individual's decisions (Fiedler, 2008; Hogarth & Soyer, 2011). A third, full feedback environment—informing about all outcomes—established an ideal environment to monitor and eventually overcome non-adaptive decision strategies.

Method

We confronted children and adult controls with a child-friendly information-board paradigm, in which three probabilistic cues predicted binary outcomes of two options. Participants repeatedly chose between options (i.e., houses) in order to yield as many positive outcomes (i.e., treasures) as possible. Cue predictions varied from choice to choice (Figure1D-F).

Before making choices, participants picked three animals as cues and encoded the relevant probabilistic information; that is, the cue validities. Participants observed that cues differed in validity; the first cue predicted outcomes correctly in three, the second cue in four and the third cue in five out of six cases (corresponding to p = .50, .66, .83). This is further graphically depicted next to each cue as a magnitude of “smart points”; one for each correct prediction.

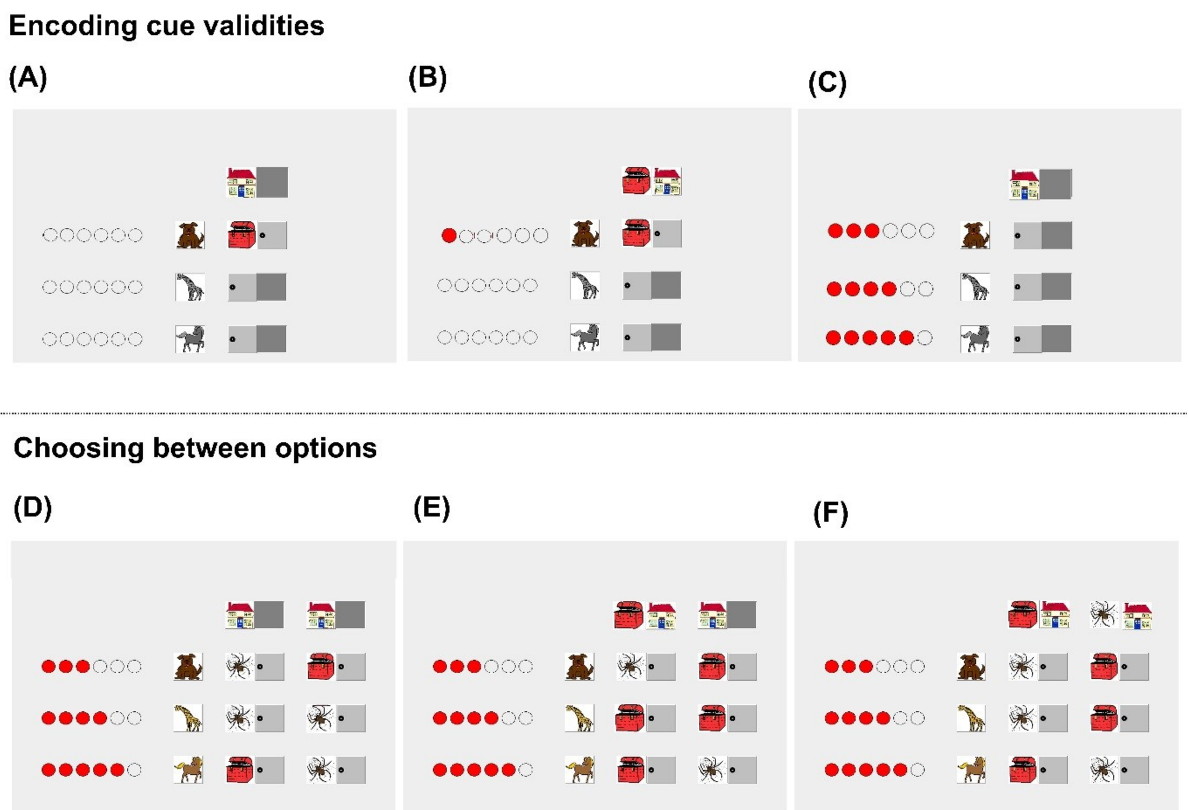


Figure 1. The Decision Game. Figure shows screenshots from the decision game during the encoding of cue validities and decisions between options. Cue predictions varied between decisions (D-F). For example, the first cue predicts a spider (i.e., negative outcome) for the left option and a treasure (i.e., positive outcome) for the right option (D). Each cue prediction pattern was used 12 times in this and in the mirrored version. In feedback conditions, participants were selectively (E) or fully (F) informed about decision outcomes. Feedback maintained each cue's validity and reinforced both options equally.

Decision Strategies

Strategies that are adaptive to this probabilistic decision environment account for the cue validities and predictions: The Lexicographic strategy considers only the predictions of the most valid cue and prescribes to follow them in all choices. The Weighted Additive Strategy weights predictions according to their uncorrected cue validities and prescribes to follow the high valid cue unless both other cues contradict it.[1]

However, the decision environment is perceptively rich and offers additional information, that is irrelevant to but can be utilized for decision making. Thus, children can also solve the decision problem without considering probabilistic information. Recent research suggests three non-adaptive decision strategies (Lang & Betsch, 2018): First, Take-

[1] For Weighted Additive all predictions should be weighted by cue validities and added up to determine the expected value for each option. However, whether and how cue validities must be transformed before they are used as weights is discussed controversially; we used untransformed cue validities (see Lang & Betsch, 2018; Betsch et al., 2014, for a discussion).

your-Favorite, prescribes to follow the predictions of the favored, first-chosen cue despite its low validity (see also Betsch & Lang, 2013). Second, option alternation prescribes to systematically alternate between the choice option on the right and on the left of the decision matrix. It considers which position was rejected in the last trial and neglects all other information (see also Brainerd, 1981). The third strategy, Win-Stay-Lose-Shift, considers only the last decision outcome (and can only be applied in feedback conditions); it prescribes to select the option on the same position as long as it yields positive results, but to switch if it does not. It accounts for children's over-sensitivity to recent negative outcomes (see also van Duijvenvoorde, Jansen, Bredman, & Huizenga, 2012).

Crucially, when participants choose repeatedly, all strategies result in different choice patterns (Figure 1D-F, e.g., the Lexicographic Strategy prescribes choosing the first option in all three decisions; Weighted Additive prescribes choosing the first option in the first two and the second option in the third decision). We applied an outcome-based maximum likelihood strategy classification to identify each individual's decision strategy (Bröder & Schiffer, 2003). Each individual's choices were compared with predictions derived from each of the strategies. If they fitted the strategy predictions flawlessly, the individual was assigned to this strategy. Otherwise, we calculated how likely the individual choice pattern would be produced by each of the strategies. We assumed that the individual applied the strategy with a constant error rate maximizing the likelihood. We assigned individuals to the strategy that produced the highest likelihood, if at least 66% of their decisions were strategy-consistent. Participants were classified to the Random model, if choosing randomly produced the highest likelihood or none of the other strategies fitted the participant's choices well enough. Participants with equal likelihoods for two strategies remained unclassified.

We assumed that participants might adapt strategies in feedback conditions and shift to more adaptive strategies over time. Therefore, we determined each individual's strategy in the last third of the 72 decisions, when feedback effects should have been most pronounced (i.e., adaptive decision strategies should have been most and non-adaptive least prevalent). We ran the analysis a second time with the last two thirds of decisions. The main results did not change, but are more reliable the more choices are included, therefore we report the second analysis.

Design & Sample

The study implemented a 3 (age group: 6-year-old preschoolers vs. 9-year-old elementary schoolers vs. adult controls) × 3 (feedback condition: no vs. selective vs. full feedback) between-subjects design. Participants were randomized to feedback conditions.

The sample consisted of 81 6-year-old preschoolers (*M* = 68.61 months, *SD* = 5.15, Range = 60–78, 34 female), 91 9-year-old elementary schoolers (*M* = 107.25 months, *SD* = 8.64, Range = 84–149, 55 female), and 85 adults (*M* = 253.20 months, *SD* = 38.52, 66 female). Children came from different socioeconomic backgrounds, and were recruited from different German schools and kindergartens. Parents consented to their child's participation. One 6-year-old and one 9-year-old did not complete the study; eight more 6-year-olds and three 9-year-olds failed at selecting the most valid cue as a manipulation check. The final sample consisted of 72 preschoolers (*M* = 68.89 months, *SD* = 5.27, Range = 60–78, 31 female), 87 elementary schoolers (*M* = 106.91 months, *SD* = 8.45, Range = 84–149, 52 female) and 85 adults.

Procedure

Child data were collected by a trained experimenter in a separate room of the school or kindergarten on a touchscreen computer. She explained that the game's purpose was to find treasures that would be traded for presents afterwards. The experimenter started the game with introducing the cues in the following fashion: "You don't have to find the treasures on your own; somebody will be helping you. Do you see the animals on the screen? You are allowed to choose three of them to help you." The first chosen animal was placed on the top row of the screen, the last chosen on the bottom row to ensure that the favored animal was lowest in validity.

Encoding cue validities prior to choices. Cue validity was framed during the game as the smartness of the animals. Accordingly, the experimenter explained that animals would predict hidden outcomes of a house but differed in smartness. Supervised by the experimenter, children encoded differences in cue validities: The experimenter started with the animal placed on the top row; she opened the door next to the cue to uncover the cue's prediction (e.g., "The dog says there is a treasure in the house.", Figure 1A). Then she revealed the actual outcome by opening the house (e.g., "There is a treasure in the house.", Figure 1B), and explained that for this correct prediction the animal was granted a "smart point". In the following five trials, the participant was asked to indicate whether the animal should be granted a smart point. No child had difficulties in identifying correct predictions and granting smart points accordingly. For each smart point, the experimenter clicked on a point next to the cue which turned red. After six trials, the experimenter summarized the cue's validity by counting smart points together with the child and stating that it had received three out of six smart points. Similarly, participants observed the predictions of the second and third cue and corresponding outcomes. Subsequently each cue's validity was displayed by a magnitude of smart points (i.e., three, four and five, cf. Figure 1C). Participants were required to identify the most valid cue as a manipulation check before choices and excluded from analyses if they failed to do so.

Choices between options. Before the choices, the experimenter explained that the child would make several choices between two houses, would receive a treasure point for each treasure found, and ensured in two training trials that the child understood the procedure and the information board matrix. Participants made 72 decisions with varying cue predictions (Figure 1D-F), and two short breaks after the 24th and 48th decision.

In the selective feedback conditions, participants experienced the consequence of their decision by opening the house and finding either a treasure or a spider. The experimenter provided additional verbal feedback ("Oh great, you found a treasure. Now you get a treasure point." vs. "Oh no, a spider. You don't get a treasure point.", Figure 1E) and filled in a treasure point if a treasure was found. When full feedback was provided, the foregone outcome was revealed and verbalized additionally ("There was a treasure/spider in the other house.", Figure 1F). In conditions without feedback, participants decided without observing the outcome. The whole study took 30 to 45 minutes. Children were rewarded contingent on their performance. Additionally, they answered several questions concerning the cues and their motivation during the game, which we do not address here.

Procedure for adults. We informed adults that they served as a control group in a children's study in advance. The procedure was identical, but adults received money contingent on their performance (M = €8).

Results

Table 1 shows frequencies of decision strategies. First, we observed consistent application of decision strategies in all age groups. Half of the children in both age groups applied adaptive or non-adaptive decision strategies. Second, as expected, we obtained striking differences between age groups: Only 6% of 6-year-olds applied adaptive strategies (i.e., Lexicographic, Weighted Additive), but 41% of 9-year-olds and 70% of adults did, $\chi^2(2) = 68.38$, $p < .001$, Cramer's $V = .53$. Non-adaptive strategies (i.e., Favored Cue, Option

Table 1

Participants' decision strategies

	No feedback	Selective feedback	Full feedback	Overall	
6-year-olds	n = 27	n = 23	n = 22	n = 72	Error
Lexicographic	—	9(2)	9 (2)	6 (4)	.06
Weighted Additive	—	—	—	—	
Favored Cue	33 (9)	4 (1)	—	14 (10)	.05
Option Switching	26 (7)	30 (7)	14 (3)	24 (17)	.13
Win-Stay-Lose-Shift		22 (5)	5 (1)	8 (6)	.22
Random	37 (10)	35 (8)	73 (16)	47 (34)	
Unclassified	4 (1)	—	—	1 (1)	
9-year-olds	n = 29	n = 30	n = 28	n = 87	Error
Lexicographic	10 (3)	30 (9)	46 (13)	29 (25)	.06
Weighted Additive	21 (6)	10 (3)	4 (1)	12 (10)	.07
Favored Cue	10 (3)	—	—	3 (3)	.08
Option Switching	17 (5)	7 (2)	7 (2)	10 (9)	.18
Win-Stay-Lose-Shift		—	7 (2)	2 (2)	.20
Random	41 (12)	53 (16)	32 (9)	43 (37)	
Unclassified	—	—	4 (1)	1 (1)	
Adults	n = 28	n = 28	n = 29	n = 85	Error
Lexicographic	36 (10)	50 (14)	35 (10)	40 (34)	.06
Weighted Additive	36 (10)	32 (9)	24 (7)	31 (26)	.05
Random	25 (7)	18 (5)	41 (12)	28 (24)	
Unclassified	1	—	—	1 (1)	

Note. Table shows relative frequencies and absolute frequencies in parentheses. Error = Mean error rate in percentage = the averaged proportion of strategy-incongruent choices. For example, adults diverge from the Lexicographic strategy in 6%, that is, in less than 3 out of 48 choices.

Switching, Win-Stay-Lose-Shift) were most prevalent in 6-year-olds (46%), less prevalent in 9-year-olds (15%) and absent in adults, $\chi^2(1) = 16.73$, $p \leq .001$, Cramer's V = -.32.

Third, we expected decision feedback to encourage adaptive strategies and discourage non-adaptive strategies. Results only support the latter: 6-year-olds used non-adaptive strategies less often when they were fully informed about decision outcomes, $\chi^2(2) = 9.79$, $p = .007$, Cramer's V = .37. The effect was only marginally significant in older children, $\chi^2(2) = 4.88$, $p = .087$, Cramer's V = .24, possibly because non-adaptive strategies were overall less common in this age group. Unexpectedly, the prevalence of adaptive decision strategies remained unaffected by feedback condition in both age groups, all ps $\geq .15$; instead Random decision making increased in 6-year-olds only, $\chi^2(2) = 7.55$, $p = .023$, Cramer's V = .32.

In an exploratory analysis, we investigated how accurately children applied decision strategies: Error rates suggest that children executed adaptive strategies as accurately as adults, deviating from the strategy in 3 out of 48 decisions, M_{Children}= 6%, CI [4; 8][2], M_{Adults}= 5%, CI [4; 7], $p \geq .47$. Children further executed non-adaptive strategies very consistently without feedback, M =7%, CI [3; 10], but less consistently in feedback conditions, $M_{\text{Selective Feedback}}$ = 20%, CI [15; 25], $t(27) = -4.737$, $p < .001$, $g = -1.81$, $M_{\text{Full Feedback}}$ = 14%, CI [6; 21], $t(17) = -1.78$, $p = .092$, $g = -1.23$.

We aimed to reliably identify non-adaptive and adaptive strategies in children. On average, participants' choices were much more likely produced by the assigned strategy than by a Random model or the second-best fitting strategy, compared with Random, $Mdn_{\text{6-year-olds}}$ = 88.99, $Mdn_{\text{9-year-olds}}$ = 71.49, Mdn_{Adults} = 87.51; compared with the second-best strategy, $Mdn_{\text{6-year-olds}}$ = 8.20, $Mdn_{\text{9-year-olds}}$ = 23.22, Mdn_{Adults} = 87.51. For the vast majority of participants, we achieved at least moderate evidence for the assigned strategy (i.e., $OR > 3$ in 92% when compared with Random, in 86% when compared with the second-best strategy, Wassermann, 2000). Thus, we are confident that we could correctly identify strategy use. Analysis of decision times further supports the strategy classification: When children used simple strategies, like alternating options, they needed three seconds to make their decisions, M = 2.9s, CI [2.43;3.57], while more complex strategies such as Weighted Additive needed more time, M = 4.5s, CI [3.6; 5.6].

Each age group included participants whose choices were best fitted by the Random model. This can be expected for children, who might indeed choose randomly or switch between strategies (which would likewise appear Random), but unexpected for adults, who should exclusively apply normative, adaptive decision strategies. A closer look at adults' choices reveals that most individuals not classified to either a Lexicographic or Weighted Additive strategy were in fact switching between the two. Although their decisions did not fit either strategy, they were highly adaptive.[3]

[2] All intervals are 95% confidence intervals bootstrapped with 10000 samples.

[3] This was due to choices when both low validity cues contradicted the high valid cue (Figure 1C)—the Lexicographic strategy prescribes to choose with the high valid cue while the Weighted Additive strategy prescribes to choose against it. Most adults classified as Random chose in line and against the Lexicographic Strategy in half of these choices. Because such systematic deviation leads to low likelihoods for the Lexicographic and Weighted Additive strategy, they were instead assigned to Random. Still, they otherwise chose in line with the Lexicographic strategy. In contrast, children assigned to Random did not chose adaptively: Choices in line with the Lexicographic strategy for individuals classified to Random in different prediction patterns, 6-year-olds, Mdn_{Figure1D} = 37%, inter quartile range (IQR):31–50; Mdn_{Figure1E} = 50%, IQR, 43–56; Mdn_{Figure1F} = 50%, IQR: 37–56; 9-year-olds, Mdn_{Figure1D} = 56%, IQR: 31–56; Mdn_{Figure1E}= 56%, IQR:43–72; Mdn_{Figure1F} = 50%, IQR: 31–56; adults, Mdn_{Figure1D} = 100%, IQR:: 100–100; Mdn_{Figure1E} = 100%, IQR: 95–100; Mdn_{Figure1F}= 43%, IQR: 37–56.

Analysis of decision outcomes further shows that children who used adaptive decision strategies were indeed the most successful decision makers with 77% of choices yielding positive outcomes, CI [76;78], while children applying non-adaptive strategies were less successful, 59%, CI [57;61]. Children assigned to Random succeeded in 65% of decisions, CI [63; 76]. This is in line with options' average success rates of 66%.

Discussion

We investigated 6-year-olds', 9-year-olds' and adult controls' strategies in probabilistic decisions. As expected, strategy use develops considerably from non-adaptive strategies—strategies that focus on irrelevant information and do not maximize outcomes—in 6-year-olds, to co-existence of non-adaptive and adaptive strategies—strategies that focus on probabilistic information and maximize outcomes—in 9-year-olds and adaptive strategies in adults. Strategy use in 6-year-olds and 9-year-olds was not sporadic but prevalent, with 51% in younger children, CI [40; 63], and 56% in older children, CI [46;67]. When children applied adaptive strategies, they executed them very accurately resulting in error rates comparable to adults'. Decision feedback affected the prevalence and execution of strategies: 6-year-olds monitored strategy performance and abandoned their otherwise frequently used non-adaptive strategies; both child groups executed non-adaptive strategies less accurately when receiving feedback.

With increasing age, children's decision strategies align to adults' strategies. In non-probabilistic environments, adult-like strategies emerge at preschool age (Lindow, Lang, Betsch, 2017); in more complex probabilistic environments they emerge but are not yet consolidated at late elementary school age (Betsch & Lang, 2013; Betsch et al., 2016; Lang & Betsch, 2018; Mata, van Helversen, & Rieskamp, 2011). Consistent with strategy development in other domains, children initially use very inefficient strategies (see Björklund & Caussey, 2018, for an overview of strategy development across domains). These are un-tailored to the specific demands of probabilistic decisions, but still allow children to make decisions at all.

Presumably, children cannot apply better strategies because their strategy repertoire does not entail them yet. In probabilistic decision making, adaptive strategies require consideration of probabilistic information, a challenge that young children cannot meet in a complex multi-cue environment (Betsch & Lang, 2013; Betsch et al., 2014, 2016).

Additionally, children might rely on non-adaptive strategies for quite rational reasons: Such strategies might be successful in children's everyday decision making. Children might for example, prefer advice of closer individuals, such as a parent, even when they are less accurate in a specific situation (Corriveau & Harris, 2009; Lucas, Burdett, Burgess, Wood, McGuigan, Harris, & Whiten, 2017). Thus, children might have developed a strong preference for such simple decision rules and apply them to new decisions by default (Gigerenzer & Todd, 1999). Only after having learned that this strategy performs non-satisfyingly, children might switch to better ones. Consistent with this thought, children used non-adaptive strategies less often when they received full feedback informing about the strategy's ineffectiveness. However, we found no evidence of an uptake of adaptive strategies. This is anticipated under the assumption that children do not possess adaptive strategies yet, and thus must fall back to deciding randomly.

Decision strategies are widespread in children. With a thorough measurement of decision strategies, we identified around half of the children as strategic. Still, we might underestimate this proportion. Obviously, our a-priori considered strategy repertoire might lack decision strategies actually used by children. For example, children might have counted the number of positive predictions and chosen the option with most positive predictions (Equal Weight, Payne et al., 1988). Also, children might have switched between multiple strategies, including non-adaptive and adaptive strategies (Siegler, 2007). Our approach cannot account for that, but adult data, demonstrating switching between two strategies (see Footnote 3), support this view. Thus, we assume that the proportion of children deciding strategically might be even higher. Nevertheless, that 6-year-olds can apply decision strategies and even monitor their performance is encouraging for future research attempting to improve children's decision making.

References

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology*, *115*, 108–125. https://doi.org/10.1016/j.jecp.2012.11.003

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica*, *153*, 74–86. https://doi.org/10.1016/j.actpsy.2014.09.008

Betsch, T., Lehmann, A., Lindow, S., Lang, A., & Schoemann, M. (2016). Lost in Search: (Mal-)adaption to probabilistic decision environments in children and adults. *Developmental Psychology, 52*, 311–325. https://doi.org/10.1037/dev0000077

Björklund, D. F., & Causey, K. B. (2018). *Children's thinking: Cognitive development and individual differences.* Thousand Oaks, CA: Sage Publications.

Brainerd, C. J. (1981). Working memory and the developmental analysis of probability judgment. *Psychological Review, 88,* 463–502. https://doi.org/10.1037/0033-295X.88.6.463

Bröder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making*, *16*(3), 193–213. https://doi.org/10.1002/bdm.442

Corriveau, K. & Harris, P. L. (2009). Choosing your informant: weighing familiarity and recent accuracy. *Developmental Science, 12*, 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 186–203. https://doi.org/ 10.1037/0278-7393.34.1.186

Fishburn, P. C. (1974). Lexicographic orders, utilities and decision strategies: A survey. *Management Science, 20,* 1442–1471. https://doi.org/10.1287/mnsc.20.11.1442

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650–669.

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

Hogarth, R. M., & Soyer, E. (2011). Sequentially simulated outcomes: Kind experience vs. non-transparent description. *Journal of Experimental Psychology: General, 140*(3), 434–463. https://doi.org/10.1037/a0023265

Lang, A. & Betsch, T. (2018). Children's neglect of probabilities in decision making with and without Feedback. *Frontiers in Psychology, 9, 191*. https://doi.org/10.3389/fpsyg.2018.00191

Lindow, S., Lang, A., & Betsch, T. (2017) Holistic information integration in child decision making. *Journal of Behavioral Decision Making*, *30*(5), 1131–1146. https://doi.org/10.1002/bdm.2029

Lucas, A. J., Burdett, E. R., Burgess, V., Wood, L. A., McGuigan, N., Harris, P. L., & Whiten, A. (2017), The development of selective copying: Children's learning from an expert versus their mother. *Child Development,* 88. 2026–2042. https://doi.org/10.1111/cdev.12711

Mata, R., von Helversen, B., & Rieskamp, J. (2011). When easy comes hard: The development of adaptive strategy selection. *Child Development, 82*, 687–700. https://doi.org/10.1111/j.1467-8624.2010.01535.x

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*(2), 207–236. https://doi.org/10.1037/0096-3445.135.2.207

Shrager, J. & Siegler, R.S. (1998). A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076

Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. https://doi.org/10.1111/j.1467-7687.2007.00571.x

van Duijvenvoorde, A. C. K., Jansen, B. R. J., Bredman, J. C., & Huizenga, H. M. (2012). Age-related changes in decision making: Comparing informed and noninformed situations. *Developmental Psychology*, *48*(1), 192–203. https://doi.org/10.1037/a0025601

Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology, 44*(1), 92–107. https://doi.org/10.1006/jmps.1999.1278

Article 3

Children Learn Simple, Adaptive Decision Strategies from Probabilistic Feedback

Children Learn Simple, Adaptive Decision Strategies from Probabilistic Feedback

Anna Lang and Tilmann Betsch
University of Erfurt, Germany

Author Note

Anna Lang, Department of Psychology, University of Erfurt, Germany; Tilmann Betsch, Department of Psychology, University of Erfurt, Germany.

This research was supported by a grant from the Deutsche Forschungsgemeinschaft to the second author (Grant No., BE 2012/11-1).

Correspondence concerning this article should be addressed to Anna Lang, Department of Psychology, University of Erfurt, Nordhaeuser Strasse 63, 99089 Erfurt, Germany. E-mail: anna.lang@uni-erfurt.de

Declarations of interest: none

Word count: 8057

Abstract

In two studies, children learned simple but adaptive decision strategies from probabilistic feedback. In a probabilistic multi-cue decision task, we assessed children's decision strategies in different feedback conditions. In Study 1 (*N* = 316), 7 and 9-year-old children learned the selective Lexicographic decision strategy. Strategy learning was slightly better in a stricter compared to a more lenient feedback condition and worse than in adult controls. In Study 2 (*N* = 259), feedback either reinforced the selective Lexicographic strategy or the holistic Equal Weight-strategy. Younger children learned the holistic strategy better, while older children learned both equally well. In both studies, children rarely used adaptive decision strategies without feedback, although a priori stated dispersion of probabilities suggested their use. It illustrates, that under ideal circumstances, that is, in highly dispersed probabilistic environments and when adaptive strategies are not too complex to perform, feeding back decision outcomes improves children's, otherwise very deficient, decision making.

Keywords: decision making, feedback, strategies, development, Lexicographic strategy, Take-the-Best, Equal Weight

Children Learn Simple, Adaptive Decision Strategies from Probabilistic Feedback

Adapting decision making to changing environments is a basic skill that children must acquire as soon as they make decisions on their own. Decision feedback can either lead to option learning—when children learn to adaptively select choice options, or to strategy learning—when children learn to adaptively select decision strategies.

Probabilistic option learning is available quite early in life. When feedback indicates that one option is superior, that is, it provides higher or more frequent pay-offs for this option, children as young as three years learn to select better options more frequently (Bunch; Andrews & Halford, 2007). However, a large amount of research shows that option learning is still deficient in childhood and improves considerably until adolescents and beyond (ee Cassoti et al., 2014; Defoe, Dubas, Figner, & Aken, for developmental reviews).

Probabilistic strategy learning rests upon the idea of an adaptive decision maker that selects an appropriate decision strategy from a repertoire (Payne, Bettman & Johnson, 1988; Gigerenzer, Todd & the ABC Research Group, 1999). The selection of this strategy can either work in a top-down fashion, by analyzing the decision situation and its probabilistic structure (Beach and Mitchel, 1978; Payne et al., 1988; Rieskamp & Otto, 2006), or in a bottom-up fashion via decision feedback (Rieskamp & Otto, 2006): Initially, decision makers select a strategy which they expect to perform well, but update their expectations depending on the feedback the strategy produces. If the feedback does not meet expectations, decision makers switch from non-adaptive strategies to more adaptive strategies.

Empirical evidence shows first evidence of probabilistic strategy learning in 9-11-year-old children (Mata, van Helversen, & Rieskamp, 2011). Younger children aged 5-6 and 8-9 showed not adaption of strategies to feedback (Betsch & Lang, 2013; Betsch et al., 2014; 2016; Lang & Betsch, 2018a and b, Betsch et al., 2018). But children dropped non-adaptive strategies, otherwise frequently used, when decision feedback indicated that they did not suit the decision environment (Lang & Betsch, 2018b).

However, studies either provided limited learning opportunity (Betsch & Lang, 2013; Betsch et al., 2014; 2016, Lang & Betsch, 2018a), required children to learn very complex decision strategies (Betsch et al., 2018), or offered ambiguous feedback (Betsch & Lang, 2013; Betsch et al., 2014; 2016; Lang & Betsch, 2018a and b, Betsch et al., 2018). Due to the probabilistic nature of environments, feedback is inherently ambiguous because even the best strategy does not always produce success, and even the worst strategy does not always produce failure. Nevertheless, probabilistic feedback environments can either strictly reinforce one strategy while other strategies perform noticeable worse—thus, differences in strategy performance are extended—or they can rather leniently provide high pay-offs for various strategies (Hogarth, 2001). In such environments, learning of better decision strategies is not utterly important or easy to achieve (Hogarth, 2001; Hogarth & Soyer, 2011). Strict feedback environments allow to update strategy expectations more efficiently and might thus be more suited to initiate shifts towards more adaptive strategies in young children (Rieskamp & Otto, 2006, cf. Siegler, 2006).

In two studies, we focused on learning of simplifying decision strategies in strict but probabilistic feedback conditions. The first study investigates whether children learn the prominent, simple and highly selective Lexicographic strategy (LEX; Fishburn, 1974; Take-the-Best, Gigerenzer & Goldstein, 1996): The decision maker only considers information of the most important dimension and chooses the option that offers the highest value. When dimension weights are highly dispersed, that is, one dimension is much more important than others, this strategy is very efficient and allows fast and frugal decision making (Gigerenzer & Goldstein 1996; Payne et al., 1988).

Decision makers are sometimes reluctant to apply LEX, even when it is highly adaptive (e.g., Bröder, 2003; Glöckner & Betsch, 2008). Especially children younger than ten years of age struggle with LEX (e.g., Betsch et al., 2013, 2014, 2016; Mata, et al., 2011). In a second study, we investigated whether children learned a more holistic strategy, the Equal Weight-strategy, better than LEX (Payne et al., 1988). This decision strategy is also quite simple: it requires to count positive values for each option on all dimensions and choose the option with the maximum positive values. In contrast to LEX, it does not require to focus on one dimension and ignore other, potentially salient, dimensions, and thus, might accommodate children (Mata et al., 2011). When dispersion of probabilities is low, that is, when dimensions are nearly equally important, this strategy can be very successful (Payne et al., 1988).

Study 1

The first study's goal is to demonstrate that 7-year-old and 9-year-old children learn to apply LEX from probabilistic decision feedback. We assessed children's decision strategies, based on the outcomes of their choices, under two different feedback conditions: The lenient feedback condition maintained a priori stated differences in probability. LEX performed best in this environment, but competing strategies performed also better than chance. The strict feedback condition further dispersed differences in probability. LEX performed best, and competing strategies' performances dropped to or below chance level. In a control condition, we offered no decision feedback, but participants could use a priori stated probabilities to select an adaptive strategy.

Method

The most prominent paradigm for studying decision strategies is the probabilistic multi-cue environment (e.g., Payne et al.,1988; Gigerenzer & Goldstein, 1991, Lagnado et al., 2006, Glöckner & Betsch, 2008). A child-friendly adaption allows to assess whether and what decision strategies children apply (Betsch & Lang, 2013; Betsch et al., 2014; 2016; Lang & Betsch, 2018a and b, Betsch et al., 2018, Mata et al., 2011). Probabilistic information is provided a priori to decisions in terms of validities of different cues. Cues predict outcomes of different options, but the probabilistic relation between each cue and outcomes—the cue validity—determines how accurately the cue predicts outcomes and how its predictions should be weighted in decisions.

We applied an environment with three cues and two options. Different animals represented cues and were assigned to specific cue validities, that is, $p = .57$ for the first and second cue and $p = .88$ for the third cue. In the decision phase, participants repeatedly chose between two options (i.e, houses) in order to achieve a positive outcome (i.e, treasure). For each decision, they could inspect cue values, which were either positive or negative (i.e., treasure, spider, Figure 1).

Decision strategies. In order to solve this decision problem, participants can apply several decision strategies. LEX is adaptive to the environment and maximizes decision outcomes. It requires to consider the predictions of the most important dimension, that is, the high valid cue, exclusively. However, child participants might also apply strategies that are non-adaptive to the decision environment: The Equal Weight-strategy ignores the cue validities and prescribes to choose the option with the highest number of positive predictions. When cue validities are highly dispersed, ignoring differences in cue validities is non-adaptive. Further non-adaptive strategies used by children are Take-your-Favorite—following the favored but low valid cue, and Option-Alternation—systematically switching between the two options (Lang & Betsch, 2018a). Finally, a Random Choice model guesses between the two options and captures unsystematic decisions.

Feedback structure. In a control condition, no immediate feedback was provided. Two feedback conditions offered feedback immediately after each decision. Most importantly, feedback was uninformative on the two options (i.e., options on the right yielded positive outcomes equally often as options on the left); thus, option learning was non-adaptive and unsuccessful. Decision strategies in contrast differed in success. In the lenient feedback condition, LEX was the most successful strategy (in 80%), other strategies performed worse but not much worse (Equal Weight, 62%, Take-your-favorite, 60%, Option Alternation, 62%, Random, 66%). Like in previous studies, a priori provided cue validities (p_1= .57, p_2 =.57, p_3 = .86) were maintained (Betsch & Lang, 2013; Betsch et al., 2014; 2016; Lang & Betsch, 2018a and b). In the strict feedback environment, LEX likewise was most successful, but other strategies were yielding more negative feedback (Success rates, Equal Weight, 37%; Take-your-favorite, 27%; Option Alternation, 50%; Random, 50%). Note that, consequently, the validities of the other cues dropped (from the observed validity prior to choices to p =.29 for the first and p =.34 for the second cue at the end of decision phase), while the validity of the high valid cue was constant across conditions.

Prediction patterns. Cue values varied and were visible during the decision to facilitate encoding (Glöckner & Betsch, 2008). We used two types of cue value patterns in 60 decision trials, which were organized in three blocks and separated by short breaks (Figure 1B-C). In the first pattern type, presented 12 times in each block, both low valid cues contradicted the high valid cue's value. In the second pattern type, presented six times, one of the low valid cues contradicted the high valid cue's value and the other was indifferent. In addition, we used filler items without contradiction between cues (relevant for identifying decision strategies).

Design and sample. The sample consisted of 110 6 to 7-year-old children in grade one (46% female, M = 79.84 months SD = 5.55) 107 8 to 9-year-old children in grade three (48% female, M = 106.31 months SD = 5.34) and 99 21-year-old adult controls (67% female, M = 257.42 months SD = 35.83); all living in Germany. Excluding participants who could not correctly identify the high valid cue (one 7-year-old) or did not complete the study (two adults) left 109 7-year-olds (46% female, M = 79.78 months SD = 5.55), 107 9-year-olds and 97 adults (66% female, M = 256.51 months SD = 34.97). The study applied a 3 (Age Group: 7-year-olds vs. 9-year-olds vs. adults) × 3 (Feedback Environment: No Feedback, Lenient Feedback vs. Strict Feedback) between design.

Procedure . The procedure was identical to previous research using the same paradigm (e.g., Lang & Betsch, 2018a). A trained experimenter supervised each participant in a separate cubicle of the lab using a desktop computer. The experimenter explained that the game's purpose was to find as many treasures as possible, which paid for presents afterwards. She started the game with introducing the cues as follows: "You don't have to find the treasures on your own; animals will be helping you. Do you see the animals on the screen? You are allowed to choose three of them to help you". The first chosen animal was assigned to the top row of the screen, the last chosen on the bottom row to ensure that participant's liking of cues was not corresponding with validity.

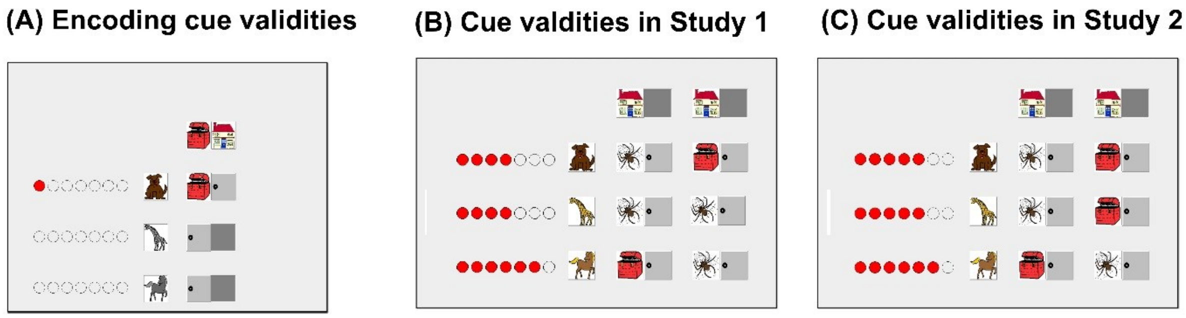


Figure 1. The Treasure Hunt Game. Participants first encoded each cue's validity (A): The most valid cue correctly predicted outcomes in six out of seven times, resulting in six smart points (p = .86). In Study 1, both low valid cues gained four (p = .57, B), in Study 2, they gained five smart points (p = .71, C). Participants could inspect cue values before choosing between two options. Cue values changed for each choice resulting in different cue value patterns (e.g., Type 2 in B and Type 1 in C). In feedback conditions, participants revealed the option's outcome immediately and earned a treasure point in case of a positive outcome.

Encoding cue validities prior to choices. To ensure that even preschool-aged children were able to grasp the concept of cue validity, we framed it as smartness of animals: The experimenter explained that the house at the top of the board would hide a treasure or a spider and that the animals would predict hidden outcomes but differed in smartness (Figure 1A). Each animal would earn smart points for correct predictions. The experimenter started with the animal placed on the top row, opened the door next to the cue and explained the meaning of its prediction (e.g., "The dog says there is a treasure in the house"). She then opened the house, revealed and explained the actual outcome, and granted a smart point to the cue (e.g., "There is a treasure in the house. The dog was smart and knew it. It gets a smart point"). In the following five trials, children identified correct predictions and granted smart points to the cue. For each smart point, the experimenter clicked on one of the point next to the cue that turned red. After seven trials, the experimenter summarized the cue's validity by counting smart points together with the child and stating that it had received four out of seven smart points. Participants then observed predictions of the second and third cue and corresponding outcomes. To keep cue validities salient, we displayed the magnitude of smart points (i.e., four and six) next to each cue throughout the game (Figure 1). Participants identified the cue highest in validity as a manipulation check before decision making.

Decisions. The experimenter explained that the participant should inspect the animals' predictions, would make several decisions between two houses, and would receive a treasure point for each treasure found. Two training trials ensured that the children understood the procedure and the information−board matrix. Participants then made 60 decisions between two options, organized in three blocks, with two short breaks after the 20th and 40th decision. Participants were either immediately informed about the outcome of their decisions in feedback conditions, or not informed until the end of the game in conditions without feedback. After the final decision, the manipulation check was assessed a second time and we assessed several judgments concerning the cues, which we will not further address in this paper.

Overall the study took about 30 minutes to complete. All children were highly engaged until the end of the game and debriefed afterwards in accordance with their age. Participants were rewarded according to their choice performance with 2 to 5 prizes for children and 5-11 € for adults. The procedure for adults was similar to children's, but adults were informed that they served as a control group in a children's study in advance.

Results

We expected children in both age groups to profit from strict feedback and learning the simple Lexicographic strategy (LEX, i.e., decisions following the most valid cue's predictions) over time. When feedback is more lenient, feedback should be less helpful for improving decision strategies. To detect change, we analyzed choice data on group level and on individual level. Collapsing choice data across individuals can disguise inter-individual differences in strategies (e.g., Bröder & Schiffer, 2003). But it can be more sensitive for detection of small shifts in strategy use. Individual data analysis shows inter-individual differences, but detecting change requires participants to coherently apply their decision strategy. It can thus be less sensitive for detecting small shifts.

Decisions in line with LEX. We assessed choices in line with LEX in the first, the second and the third block of decisions. Figure 2 shows the proportion of decisions in line with LEX over time for each feedback condition and age group. In both feedback conditions, decisions in all age groups aligned to LEX. Without feedback, only adults' decisions were in line with LEX above chance level. Nine-year-olds performed around and 7-year-olds performed below chance level.

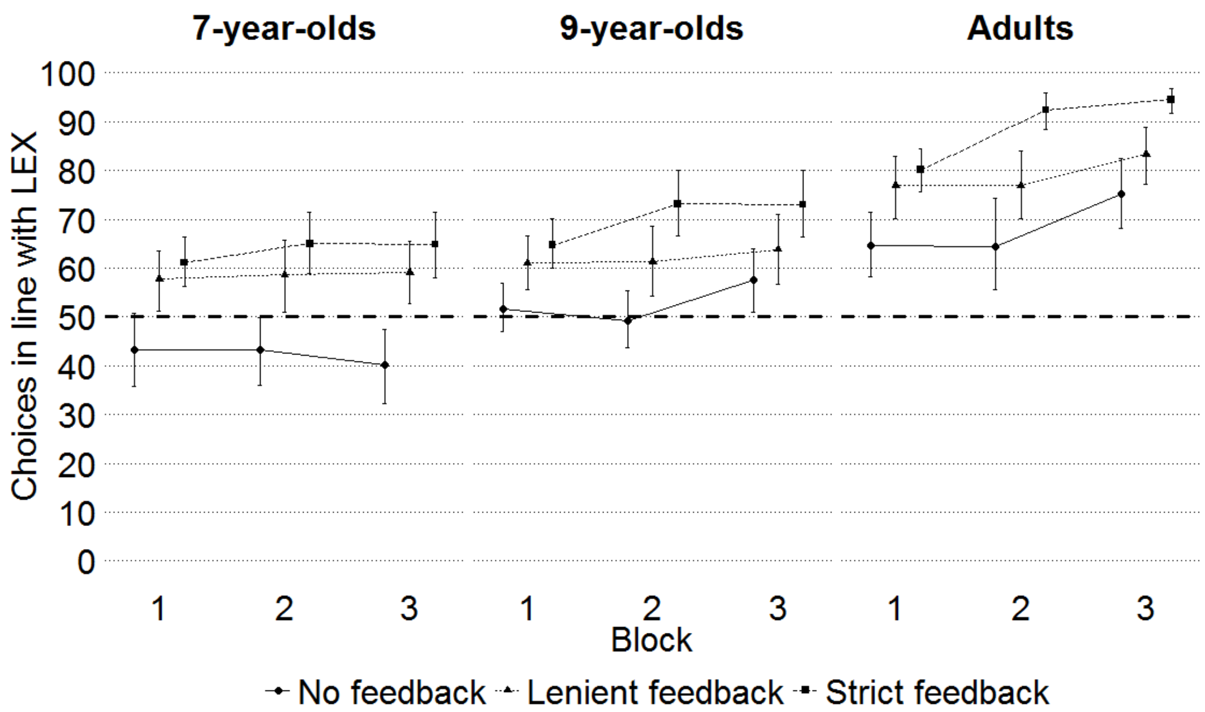


Figure 2. Percentage of choices in line with the Lexicographic strategy in Study 1. Bars represent 95% bootstrap confidence intervals. Dashed line represents chance level.

We conducted a GLM ANOVA with Age Group and Feedback Condition as between and Decision Block (1-3) as within factor for child groups. A small main effect of age, $F(1, 210) = 8.72$, $p = .004$ $\eta^2_p = .04$, indicated that older children generally chose more often in accordance with LEX; a larger main effect of feedback, $F(2, 210) = 23.61$, $p < .001$ $\eta^2_p = .18$, indicated that both child groups benefitted from feedback as expected, all other *p*s ≥.05. Older and younger children's choices aligned to LEX equally good and very quickly, noticeable already in the first decision block (Figure 1). In line with expectations, both child groups profited slightly more from strict feedback when compared to lenient feedback; $F(1, 145) = 5.85$, $p = .017$, $\eta^2_p = .04$, block, Huynh-Feldt-$F(1{,}739; 260.022) = 3.61$, $p = .033$, $\eta^2_p = .02$; all other *p*s ≥.21.

Decision strategies. We further analyzed children's individual decision strategies applied in the last block where feedback effects should be most pronounced[1]. We assigned participants to one of the considered decision strategies if their decisions in the last block fitted strategy predictions flawlessly. If this was not the case for any strategy, maximum likelihood classification determined the likelihood of the observed choices pattern when

---

[1] We also performed the analysis for the last 40 decisions for each participant. This produces more reliable results because the number of decisions in which models predictions differ increases and allows to better distinguish between models. At the downside, adaption might occur later and would remain undetected. Therefore, we decided to rather limit the analyzed decisions to the last block.

applying each strategy with a constant error maximizing the likelihood (Bröder & Schiffer, 2003). Each participant was thus assigned to the strategy which most likely produced the observed choice pattern.

Table 1 provides the results. Without feedback, children mainly relied on non-adaptive strategies, such as Take-your-Favorite, Option Alternation and Equal Weight (69% in younger and 45% in older children). With feedback, the prevalence of the adaptive LEX-strategy increases within child groups. A binary regression for child groups tested whether LEX was more often applied with lenient feedback (41%) compared to without feedback (11%) but less often compared with strict feedback (46%). The analysis shows that children applied LEX more often compared with conditions without feedback, $B = -1.62$, $p = .02$, $OR = .20$, but not less often compared with strict feedback, $B = -0.54$, $p = .74$; Wald- $\chi^2(2) = 5.67$, $p = .059$; Constant, $B = -0.61$; all other ps $\geq .31$; Nagelkerkes $R^2 = .21$. Accordingly, the small benefit of strict feedback on group level does not replicate in individual data analysis.

Note, that older children were not generally more likely to apply LEX than younger children (40% vs. 27%). Although older children performed surprisingly well with strict feedback peaking in 62% of LEX-users, we obtained no interaction between age and feedback condition. Unexpectedly, adults were not exclusively using LEX without feedback, and likewise benefitted from feedback, χ^2 (2, 97) = 20.14, $p \leq .001$, Cramer's $\varphi = .45$.

Surprisingly, non-adaptive strategies were not reduced in feedback conditions, all ps $\geq$ .06. Even in strict feedback conditions, 41% of younger and 24% of older children relied on non-adaptive strategies such as Option Alternation, Take-your-Favorite and Equal Weight.

Table 1

Decision strategies in Study 1

	No feedback		Lenient feedback		Strict feedback		Overall		
7-year-olds	n = 31		n =37		n = 41		n = 109		
	n	%	n	%	n	%	n	%	Error
Lexicographic	3	10	13	35	13	31	29	27	.17
Equal Weight	3	10	5	14	5	12	13	12	.35
Take-your-Favorite	3	10	1	3	1	2	5	5	.26
Option Alternation	15	49	15	40	16	39	46	42	.19
Random	3	10	3	8	6	14	12	11	
Unclassified	4	13	—	—	—	—	4	4	
9-year-olds	n = 36		n = 34		n = 37		n = 107		
	n	%	n	%	n	%	n	%	Error
Lexicographic	4	11	16	47	23	62	43	40	.17
Equal Weight	2	6	1	3	1	3	4	4	.23
Take-your-Favorite	2	6	4	12	2	5	8	8	.35
Option Alternation	12	33	10	29	7	19	29	27	.27
Random	15	42	3	9	4	11	22	25	
Unclassified	1	3	—	—	—	—	1	1	
Adults	n = 32		n = 31		n = 34		n = 97		
	n	%	n	%	n	%	n	%	Error
Lexicographic	17	53	23	74	34	100	74	76	.08
Equal Weight	4	13	—	—	—	—	4	4	.14
Take-your-Favorite	—	—	—	—	—	—	—	—	
Option Alternation	5	16	4	13	—	—	9	9	.34
Random	6	19	4	13	—	—	10	10	
Unclassified	—	—	—	—	—	—			

Note. Error = Mean error rate = the averaged proportion of strategy-incongruent choices. For example, adults diverge from the Lexicographic strategy in 8%, that is, in less than two out of 20 choices.

Evaluations of cues after decisions. 91% (*n* = 99) of younger, 90% (*n* = 96) of older children and 98% (*n* = 95) of adults correctly selected the last as the most valid cue after they made their decisions, χ^2 (2, 313) = 5.867, *p* = .053, Cramer's φ = .05. Thus, failure to implement LEX cannot be explained by a failure in memorizing the dispersion of cue validities.

Discussion

7-year-olds and 9-year-olds learned to use the simple and adaptive Lexicographic strategy (LEX) in a highly dispersed probabilistic environment with decision feedback. Decisions on group level and individual decision strategies aligned to LEX when feedback reinforced its use. Unexpectedly, the benefit of a strict feedback environment was limited. Children learned LEX nearly equally well in lenient feedback conditions, where though LEX performed best, other strategies performed better than chance. But in contrast to previous unsuccessful attempts in young children's strategy learning, a priori stated probabilities were more dispersed. Consequently, differences in strategy performance were already more pronounced and feedback favored LEX more vehemently in comparison to other decision strategies.

We conclude, that children as young as seven can learn adaptive decision strategies from feedback under certain circumstances (see General Discussion). Nevertheless, feedback's benefits were limited: Even with strict feedback, children's decisions still diverged considerably from the strategies' predictions on group level (cf. Figure 2); and many children still did not apply the adaptive LEX-strategy, and consistently held on to non-adaptive decision strategies (cf. Table 1). Those limitations might, however, specifically apply to LEX. Highly selective strategies, such as LEX, might be hard to learn for children, because inhibition of information is very challenging for younger children (Mata et al., 2011). Thus, they might learn decision strategies better that require to holistically integrate all dimensions. The Equal Weight-strategy integrates all cue value, yet is simple: it considers all cues' values equally, regardless of their validity. When cue validities are low in dispersion, ignoring small differences in cue validities is adaptive and Equal Weight performs well. In the second study, we investigated, whether children learned the holistic Equal Weight-strategy better than the selective LEX-strategy.

Study 2

We established a decision environment with low dispersion of cues before decisions ($p_{1,2}$ = .71; p_3 = .86). Crucially, we wanted to create an environment where a priori stated cue validities suggest that both LEX and Equal Weight can be applied successfully (based on the high validity of Cue 3, LEX can be expected to perform well: Since both other cues are also high in validity, it is also reasonable to neglect cue validities and apply Equal Weight). We further established two strict feedback conditions, one in which LEX was reinforced and one in which Equal Weight was reinforced[2]. In a control condition, we offered no feedback. The procedure and further material was identical to Study 1.

Method

The sample consisted of 92 7-year-old children in grade one (46% female, *M* = 81.83 months *SD* = 6.73) 90 9-year-old children in grade three (37% female, M = 110.21 months SD = 7.65) and 77 22-year-old adults (67% female, *M* = 272.64 months *SD* = 30.01). We excluded participants that failed the manipulation check (four 7-year-olds, one 9-year-old and four adults) or did not complete the whole study (one 7-year-old). Thus, the final sample included 86 7-year-olds (43% female, *M* = 82.00 months *SD* = 6.79), 89 9-year-olds (40% female, *M* = 110.31 months *SD* = 7.63) and 72 adults (68% female, *M* = 274.14 months *SD* = 29.64).

The study applied a 3(Age group: 7-year-olds vs.9-year-olds vs. adults) × 3 (Feedback environment: No feedback vs. Feedback for LEX vs. Feedback for Equal Weight) between design. In both feedback conditions, feedback favored the most adaptive strategy quite vehemently (LEX 80% overall success rate, Equal Weight 75% success rate). Competing strategies performed below chance (Equal Weight, 37%; LEX 41%; Option Alternation 50%; Take-your-Favorite, 27%, 51%)[3]. The patterns and their ordering was identical to Study 1.

Results

We again analyzed decisions on group level. Figure 3 show decisions in line with LEX when it contradicted Equal Weight (Figure 1C). As expected, decisions adapt quickly to feedback resulting in more decision in line with LEX or with Equal Weight (i.e., choices in line with LEX decrease) in dependence on feedback. Interestingly, younger children's decisions without feedback tended to align to Equal Weight. Older children and adults in

---

[2] We did not apply a feedback environment which only maintains cue validities as provided. Such feedback would be rather uninformative and we would not expect different decision behavior than without feedback. Note, that consequently, cue validities get more dispersed over time.

[3] Take-your-favorite was more accurate when feedback favored Equal Weight due to pattern constraints. But this works against our hypothesis, and should impede learning Equal Weight.

contrast, decided equally often in line with LEX and Equal Weight. This is expectable, because both strategies can be expected to perform well based on the a priori stated cue validities.

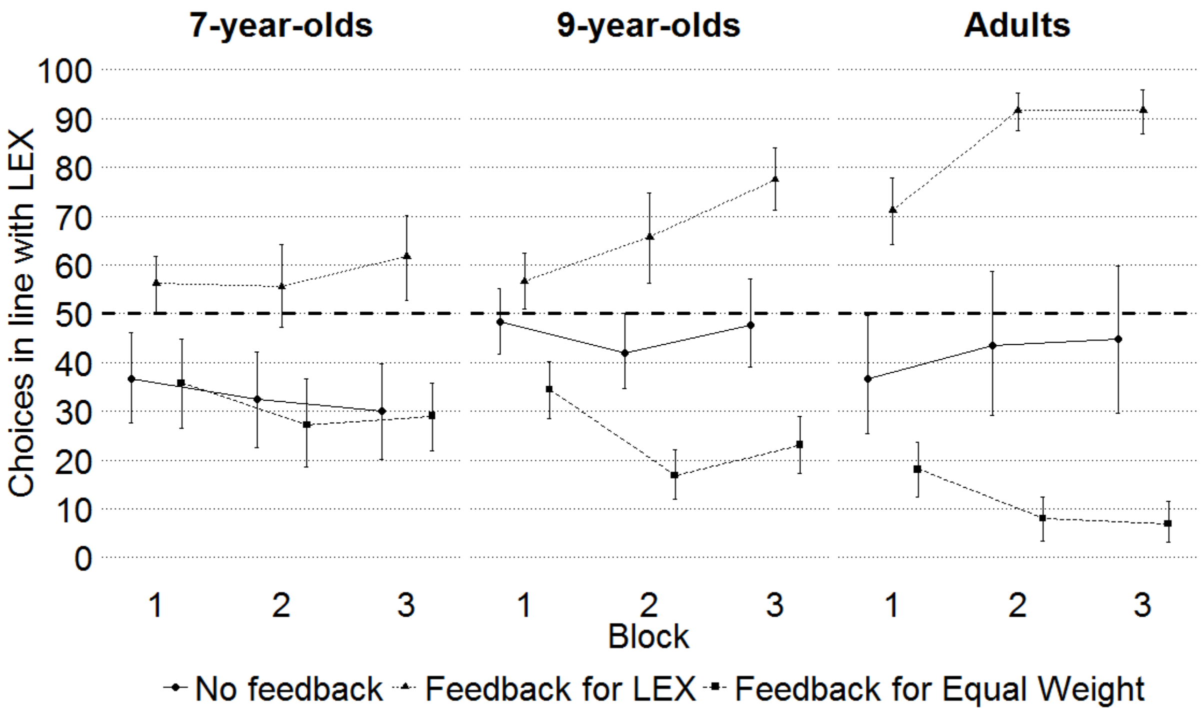


Figure 3. Percentage of choices in line with the Lexicographic strategy (i.e., against the Equal Weight-strategy) in Study 2' first prediction pattern. Bars represent 95% bootstrap confidence intervals. Dashed line represents chance level.

Adaptive decisions. For ease of comparison, we then analyzed adaptive decisions in feedback conditions. Adaptive decisions are in line with LEX when LEX is reinforced, and in line with Equal Weight when Equal Weight is reinforced. We expected adaption to be quicker and more efficient for Equal Weight. Figure 4 shows that all age groups adapted to feedback in both feedback conditions, but differ in children.

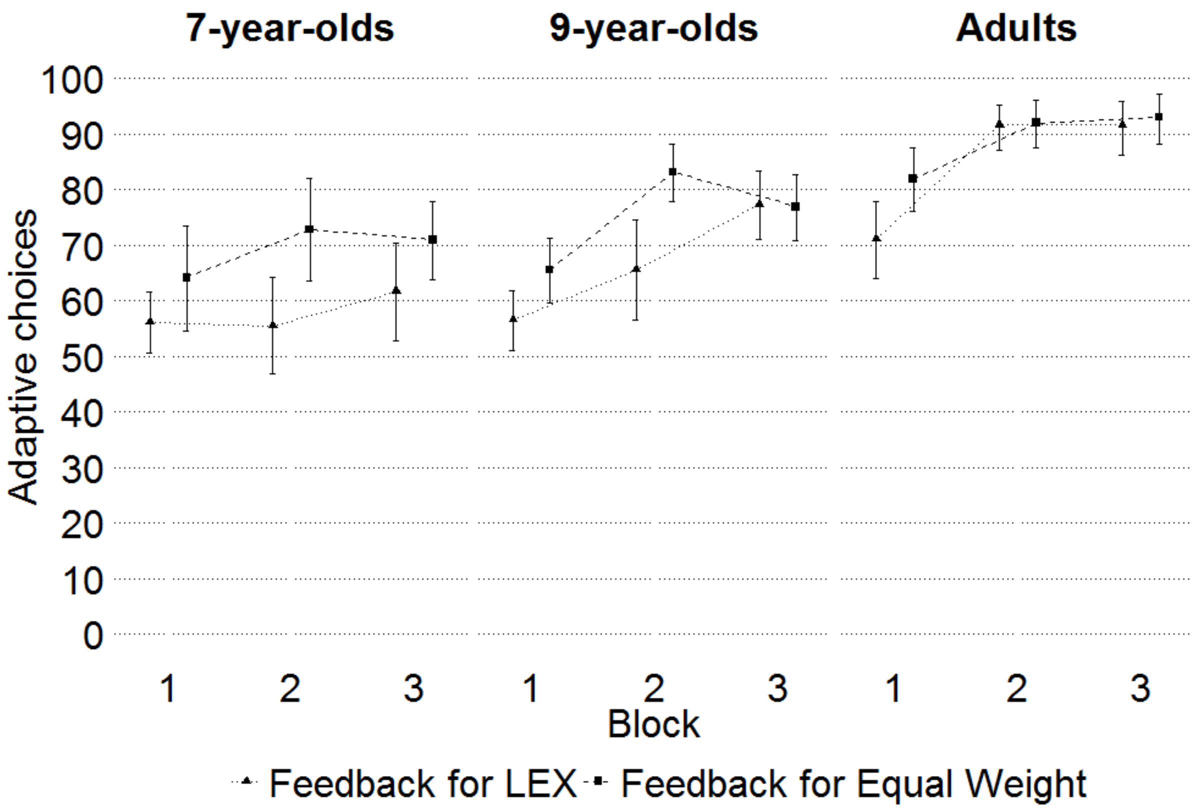


Figure 4. Percentage of adaptive choices in feedback conditions, that is, LEX or Equal Weight in the first prediction patterns. Bars represent 95% bootstrap confidence intervals.

An ANOVA on adaptive choices results in an effect of age, $F(2, 161) = 36.96$, $p < .001$, $\eta^2_p = .32$ with adults being more adaptive than older children, $p < .001$, but also older children being more adaptive than younger children, $p = .005$. Further main effects of block, and several interactions indicate that adaption was faster and more efficient in different feedback conditions.

Separate analysis for each age group with ANOVAs on adaptive choices and Feedback and Block as factors provide a clearer picture of adaptivity: 7-year-olds adapted decisions stronger when feedback favored Equal Weight as supported by a main effect for Feedback Condition, $F(1, 52) = 5.95$, $p = .018$, $\eta^2_p = .04$, but not quicker as the lack of a Block-Feedback interaction indicates, and no improvement beyond the first block; all other $ps \geq .15$.

9-year-olds also adapted stronger to feedback favoring Equal Weight, $F(1, 61) = 5.81$, $p = .019$, $\eta^2_p = .09$, their decisions continued to improve over blocks in both conditions, Huynh-Feldt-$F(1.98, 130.75) = 21.89$, $p < .001$, $\eta^2_p = .26$; and improved more rapidly when Equal Weight was favored, Block × Feedback Condition, Huynh-Feldt-$F(1.98, 120.78) = 6.02$, $p = .003$, $\eta^2_p = .09$.

Adults adapted equally well to both feedback environments, we obtained only an effect of block indicating that decisions in both conditions improved over blocks ; Block, Huynh-Feldt-$F(1.66, 79.53) = 3.21$, $p < .001$, $\eta^2_p = .40$, all other *p*s $\geq .055$.

Decision strategies. We conducted the same strategy classification method as in Study 1[4]. Table 2 provides the results. Without feedback, adults relied on either LEX or Equal Weight. Children used LEX and Equal Weight as well, but other decision strategies more frequently.

In all age groups, the prevalence of the most adaptive strategy—LEX when feedback reinforced LEX, and Equal Weight when feedback reinforced Equal Weight—increased with feedback. Nearly all adults but only 43% of younger and 51% of older children applied the favored strategy. Comparing prevalence of adaptive strategies in feedback conditions shows that older children took up Equal Weight and LEX equally well, χ^2 (1, 63) = 0.67, $p = .797$; but younger children used Equal Weight more often, χ^2 (1, 54) = 6.135, $p = .027$, Cramer's $V = .01$.

[4] In this decision environment, it is further possible to distinguish Weighted Additive (Payne et al, 1988) from a Lexicographic Strategy and Equal Weight-strategy. We did not do that for two reasons: First, our main focus was to distinguish between simple decision strategies, therefore we focused on the most promising candidate from these two categories in children. Previously, children very rarely applied Weighted Additive even it was reinforced (Betsch et al., 2018). Second, the cue validities in feedback conditions change over the course of choices. If decision makers updated cue validities, Weighted Additive-predictions would start to converge to LEX predictions in "Feedback for LEX"-conditions, and to Equal Weight predictions in "Feedback for Equal Weight"-conditions.

Table 2

Decision strategies in Study 2

	No feedback		Feedback for LEX		Feedback for EQW		Overall		
7-year-olds	$n = 32$		$n = 27$		$n = 27$		$n = 86$		
	n	%	*n*	%	*n*	%	*n*	%	Error
Lexicographic	2	6	7	26	—	—	9	11	.13
Equal Weight	8	25	2	7	16	60	26	30	.19
Take-your-Favorite	4	13	3	11	4	15	11	13	.25
Option Alternation	8	25	8	30	5	19	21	24	.21
Random	2	6	7	26	2	7	11	12	
Unclassified	8	25	—	—	—	—	8	9	
9-year-olds	$n = 26$		$n = 24$		$n = 39$		$n = 89$		
	n	%	*n*	%	*n*	%	*n*	%	Error
Lexicographic	5	19	13	54	2	5	20	26	.15
Equal Weight	4	15	—	—	19	49	23	26	.14
Take-your-Favorite	2	8	1	4	8	21	11	12	.27
Option Alternation	12	46	9	38	6	15	27	30	.26
Random	2	8	1	4	3	8	6	7	
Unclassified	1	4	—	—	1	3	2	2	
Adults	$n = 22$		$n = 26$		$n = 24$		$n = 78$		
	n	%	*n*	%	*n*	%	*n*	%	Error
Lexicographic	9	41	25	96	—	—	34	47	.08
Equal Weight	10	46	—	—	21	88	31	34	.06
Take-your-Favorite	—	—	—	—	—	—	—	—	
Option Alternation	3	14	1	4	2	8	6	8	.37
Random	—	—	—	—	1	4	1	1	
Unclassified	—	—	—	—	—	—			

Note. Error = Mean error rate = the averaged proportion of strategy-incongruent choices. For example, adults diverge from the Lexicographic strategy in 8%, that is, in less than two out of 20 choices.

Importantly, younger children used Equal Weight frequently without feedback (25/32), and even more frequently with feedback, χ^2 (1, 59) = 7.12, p = .008, Cramer's V = .35, OR = 2.4. Use of LEX was low without feedback and increased but remained low with feedback (7/27), χ^2 (1, 59) = 4.39, p = .036, Cramer's V = .27, OR = 6. In contrast, older children, use LEX and Equal Weight equally often without feedback and increase use of both with feedback; LEX, χ^2 (1, 50) = 6.61, p = .010 Cramer's V = .36, OR = 2.8; Equal Weight, χ^2 (1, 65) = 7.581, p = .006, Cramer's V = .34, OR = 3.2.

Similar to Study 1, non-adaptive strategies were still frequently used by children and not reduced in feedback conditions, all ps ≥ .47.

Evaluations of cues after decisions. Interestingly, feedback did not only initiate shifts in strategy use but also affected children's evaluation of the cues and their validity: When

Equal Weight was favored, and accordingly the feedback schedule dropped the initially high validity of the last cue, only 52% of younger children chose the last cue as the most valid one. In comparison, 84% selected the last cue without feedback, and 85% when LEX was favored, χ^2 (2, 86) = 10.53, *p* = .005, Cramer's *V* = .35. The same pattern is found in older children, when feedback favored Equal Weigh, 36%; and LEX, 83%; without feedback, 85%; χ^2 (2, 89) = 14.30, *p* = .001, Cramer's V = 40; and adults, when feedback favored Equal Weight, 58%; and LEX, 100%; without feedback, 95%; χ^2 (2, 72) = 19.56, $p \leq .001$, Cramer's V = .52.

Discussion

In the second study, we tested whether children learned the Equal Weight-strategy better than the Lexicographic strategy (LEX). And indeed, when feedback favored Equal Weight, 7 and 9-year-old children adapted decisions group level better: On individual level, 7-year-olds applied Equal Weight more often with and without feedback compared to LEX, but 9-year olds did not. Thus, it might be easier for younger children to make good decisions in an environment in which holistic strategies yield better results. This is not so clear-cut for the older children in our sample: they were more adaptive as a group, when Equal Weight was favored, but individually learned both strategies with equal success.

General Discussion

We investigated 7 and 9-year-olds' adaption of decision strategies to probabilistic feedback. We found that 7-year-olds, 9-year-olds learned simple, but adaptive decision strategies from feedback. However, age-dependent deficiencies remained even under ideal circumstances: Even when adaptive decision strategies are not overly complex and feedback reinforces them strictly, only adults showed very good to perfect uptake of the most adaptive decision strategy, a considerable proportion of children still clung to inferior, non-adaptive strategies.

Nevertheless, both studies clearly demonstrate that children learn decision strategies from probabilistic feedback. Evidence for successful learning of decision strategies had been limited to older children (Mata et al., 2011). Previous studies with children younger than nine showed no uptake of adaptive strategies in line with feedback (Betsch & Lang, 2013, Betsch et al., 2014; 2016; Betsch el al., 2018, Lang & Betsch, 2018a and b). Increased length of the decision phase cannot merely account for our findings because feedback benefits emerged early (e.g., compare LEX-decisions in Study 1's block 1, see also Betsch el al., 2018, Lang & Betsch, 2018b, for similar lengths).

We initially assumed that younger children only learn decision strategies in very strict feedback environments, which vehemently reinforce one simple strategy and simultaneously discourage other strategies, especially non-adaptive strategies children otherwise rely on, (Lang & Betsch, 2018b). Thus, we expected children to predominantly shift strategies with strict feedback, when strategy performance quickly falls short of children's, erroreous, high expectations. Increased discrepancies between strategies' performances should, in line with reinforcement learning mechanisms (Rescorla & Wagner, 1972; Sutton & Barto, 1998), facilitate strategy learning.

Strict feedback was indeed slightly more beneficial (Study 1), but children learned adaptive strategies also with more lenient feedback. Importantly in comparison to previous research, lenient feedback, which maintained stated differences in probability, produced more pronounced differences in strategy performance (since a priori stated probabilities was more dispersed, other strategies failed more often than previously).

One additional factor might further contribute to better learning: We increased the prevalence of one cue value pattern (Figure 1C). Cue values patterns varied from decision to decision, but this specific pattern was presented in 60% of choices. Previous studies, presented different cue value patterns equally often (e.g. in 33%; Lang & Betsch, 2018b; 16% of decisions, Betsch et al., 2018). Encountering one cue value pattern more frequently, might facilitate monitoring of strategy success and failure and lead to quicker shifts to adaptive strategies.

Moreover, successful strategy learning depends on the properties of the reinforced strategy. We focused on two simple strategies, the Lexicographic strategy (LEX) and the Equal Weigh strategy. Younger, but not older children learned Equal Weight-strategy more frequently than LEX (see "Why Equal Weight is easier than LEX"). In line with previous research (Mata et al., 2011), this suggest that not all decision strategies are learned equally well, respectively at all, by children. Learning more complex decision strategies, such as Weighted Additive might be doomed to fail in young age groups because of the more complex cognitive operations (Betsch et al., 2018, but see Glöckner & Betsch, Lindow, Lang & Betsch, 2017, for effortless application of a Weighted Additive like process in adults and children).

Why Equal Weight is Easier Than LEX

Seven-year-olds but not nine-year-olds learned an Equal Weight-strategy better than LEX and also used it more often without feedback. This is in contrast to previous research showing that 9-11-year-olds still learned Equal Weight better (Mata et al., 2011). Several reasons might account for Equal Weight's advantage: LEX requires to ignore all but the most important dimension. Children might fail allocating their attention to this one dimension and shielding themselves from all others. Strategies that do not require to focus on and ignore information might thus benefit children (Mata et al., 2011). Our results do not support the notion that children cannot selectively attend to information. For example, in Study 1, 59% of children focused on their favorite cue's prediction or on the previous trial's decision and successfully ignored all other information (see also Lang & Betsch, 2018a). Thus, we think it unlikely, that children are incapable of selectively considering information for decisions.

Rather, we think that two other factors might create an advantage for Equal Weight in comparison to LEX: First, LEX is a probabilistic strategy; its application requires to order dimensions according to their importance, here that is, the probabilistic relation between cue and outcome, the cue validity. Thus, its application at least requires to acknowledge the importance of probability for subsequent decisions. Recent research suggests, that although children are well aware of distribution of cue validities, they are not aware of its relevance for decision outcomes (Lang, 2018). Equal Weight in contrast, allows to ignore probabilistic relations between cues and outcomes.

Second, children might already enter the decision problems with a stronger preference for Equal Weight than for LEX. When a strategy is already expected to perform well, it will initially be used more often and consistent uptake in feedback environments is more likely (Rieskamp & Otto, 2006). In line with this, younger children applied Equal Weight more often than LEX without feedback. Potentially, younger children might prefer majority rules because they are fast and frugal in everyday decisions. However, research on the question shows mixed findings (e.g., Burdett, Lucas, Buchsbaum, McGuigan, Wood, &Whiten, 2016).

Attribution of Decision Feedback

Feedback mutably affects and sometimes even decreases cognitive performance (Bangerts-Drowns, Kulik, Kulil, & Morgan, 1991, Fyfe & Rittle-Johnson, 2015, Hattie & Timperley, 2007; Harvey & Fischer, 2005; Karelaia & Hogarth, 1998, Kluger & deNisi, 1996; Narciss & Huth, 2004, Shute, 2008,). Crucially, how individuals engage with feedback from the environment determines what they learn. One important factor for learning is feedback attribution. Especially, in multi-cue decisions, feedback can be attributed to different aspects of the decision environment: options, strategies and cues. For example, children could either attribute feedback to correct option-choice ("I found a treasure because I've chosen the right house.") or to correct selection of strategy ("I found a treasure because I've followed the majority/ the smartest animal"). Research in adults suggests that feedback attribution is volatile (Bröder, Glöckner, Betsch & Link, 2013): In probabilistic multi-cue decisions, adults attributed decision feedback to options and to strategies depending on subtle changes in feedback presentation. Importantly, in our studies, feedback format did not specifically invite attribution to decision strategies. Previous research with children suggested that children aged 5-6 attribute feedback to options (Lang & Betsch, 2018a). Children's success in strategy learning suggests that they attribute feedback to their decision strategy as well. Potentially, variation of feedback format further improves strategy learning, for example when children are rewarded for correct strategy application regardless of the actual outcome or elaborate on feedback themselves (Luwel, Foustana, Papadatos & Verschaffel; 2001; Siegler, 2007).

Children's evaluation of cues suggests that they attributed feedback to cues' performances, at least on a rudimentary level. When asked to indicate the most valid cue, children's selected the initially high valid cue less often when feedback decreased its validity (Study 2). Altogether, this provides first evidence for children attributing decision feedback in various ways.

Mechanisms of Strategy Adaption

The idea of a repertoire of decisions strategies evokes the problem of adaptive strategy selection. This adaption can work in a top-down fashion (Beach & Mitchell, 1979; Payne et al., 1988): Decision maker can consider the specifics of the decision environment (e.g., the distribution of probabilities or the complexity, e.g., the number of options) and weight costs (e.g., time and cognitive efforts) and benefits (expected performance) of available strategies. It can likewise work in a bottom-up fashion: decision makers gradually update strategy expectations in line with feedback and eventually shift to better strategies (Rieskamp & Otto, 2006; Shaffrer & Siegler, 1998).

Although, children aged seven successfully adapted strategies in a bottom-up fashion, we observed very little to no adaption in a top-down fashion. For illustration, consider conditions without feedback. Distribution of probabilities was established before the first decision, thus weighting of costs and benefits is normatively possible. In line with this, half of adults used LEX without feedback. Children in contrast, seldom used a strategy adaptive to stated probabilities without feedback. Overall, children benefitted more from feedback than from being informed about cue validities a priori (see Hogarth et al., 2011 for an overview in adults).

Conclusion

The goal of this paper was to demonstrate that adaption of decision strategies works in children under nine. Children aged seven and nine successfully learned adaptive decision strategy when feedback vehemently reinforced simple decision strategies (LEX, Equal Weight). It illustrates, that under ideal circumstances, that is, in highly dispersed probabilistic environments and when adaptive strategies are not too complex to perform, feeding back decision outcomes improves children’s, otherwise deficient, decision making, even in complex probabilistic environments.

References

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61,* 213–238. https://doi.org/10.3102/00346543061002213

Beach, L. R. & Mitchell, T.R. (1978). A Contigency Model for the Selection of Decision Strategies. T*he Academy of Management Review, 3* (3), 439–449. https://doi.org/10.5465/amr.1978.4305717

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology*, *115*, 108–125. https://doi.org/10.1016/j.jecp.2012.11.003

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica*, *153*, 74–86. https://doi.org/10.1016/j.actpsy.2014.09.008

Bröder, A., Glöckner, A., Betsch, T., Link, D., & Ettlin, F. (2013). Do people learn option or strategy routines in multi-attribute decisions? The answer depends on subtle factors. *Acta Psychologica, 143*, 200–209. https://doi.org/10.1016/j.actpsy.2013.03.005

Bröder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making*, *16*(3), 193–213. https://doi.org/10.1002/bdm.442

Bunch, K. M., Andrews, G., & Halford, G. S. (2007). Complexity effects on the children's gambling task. *Cognitive Development*, *22*(3), 376–383. https://doi.org/10.1016/j.cogdev.2007.01.004

Burdett, E. R. R, Lucas A. J., Buchsbaum, D., McGuigan, N., Wood, L. A., & Whiten, A. (2016) Do children copy an expert or a majority? Examining selective learning in instrumental and normative contexts. *PLoS ONE 11*(10*).* https://doi.org/0.1371/journal.pone.0164698.

Cassotti, M., Aïte, A., Osmont, A, Houdé, O., & Borst, G. (2014). What we learned about the processes involved in the Iowa Gambling Task. *Frontiers in Psychology*, *5*, 915. https://doi.org/ 10.3389/fpsyg.2014.00915

Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. G. (2014). A meta-analysis on age differences in risky decision making: Adolescents versus children and adults. *Psychological Bulletin, 141*(1), 48–84. https://doi.org/10.1037/a0038088

Fishburn, P. C. (1974). Lexicographic orders, utilities and decision strategies: A survey. *Management Science, 20,* 1442–1471. https://doi.org/10.1287/mnsc.20.11.1442

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology,108*(1)*,* 82–97. https://doi.org/10.1037/edu0000053

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650–669. https://doi.org/10.1037/0033-295X.103.4.650

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

Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1055–1075. https://doi.org/10.1037/0278-7393.34.5.1055

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Harvey, N. & Fischer, I. (2005). Development of experience-based judgment and decision making: The role of outcome feedback. In T Betsch & S. Haberstroh (Eds.). *The routines of decision making*. Mahwah; NJ: Lawrence Erlbaum Associates.

Hogarth, R. (2001). *Educating intuition.* Chicago: University of Chicago Press.

Hogarth, R. M., & Soyer, E. (2011). Sequentially simulated outcomes: Kind experience vs. non-transparent description. *Journal of Experimental Psychology: General, 140*(3), 434–463. https://doi.org/10.1037/a0023265

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254–284. http://dx.https://doi.org/0.1037/0033-2909.119.2.254

Lang, A. & Betsch, T. (2018a). Children's Neglect of Probabilities in Decision Making with and without Feedback. *Frontiers in Psychology, 9,* 191. https://doi.org/10.3389/fpsyg.2018.00191

Lang, A. & Betsch, T (2018b). Children use but can overcome non-adaptive decision strategies. *Manuscript submitted for publication.*

Lang, A. & Betsch, T (2018c). Utilization of Probabilistic Cues in Children's Judgments. *Manuscript submitted for publication.*

Lindow, S., Lang, A., & Betsch, T. (2017) Holistic information integration in child decision making. *Journal of Behavioral Decision Making*, *30*(5), 1131–1146. https://doi.org/10.1002/bdm.2029

Luwel, K., Foustana, A., Papadatos, Y., & Verschaffel, L. (2011). The role of intelligence and feedback in children’s strategy competence. *Journal of Experimental Child Psychology*, *108*(1), 61–76. http://doi.org/10.1016/j.jecp.2010.06.001

Mata, R., von Helversen, B., & Rieskamp, J. (2011), When easy comes hard: The development of adaptive strategy selection. *Child Development, 82*, 687–700. https://doi.org/10.1111/j.1467-8624.2010.01535.x

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner, & R. Brunken (Eds.), *Instructional design for multimedia learning* (pp. 181–195). Munster, NY: Waxmann.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. E. Prokasy (Eds.). *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*(2), 207–236. https://doi.org/10.1037/0096-3445.135.2.207

Shrager, J. & Siegler, R.S. (1998). A model of children’s strategy choices and strategy discoveries. *Psychological Science, 9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076

Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. https://doi.org/10.1111/j.1467-7687.2007.00571.x

Siegler, R.S. (2006). Microgenetic analyses of learning. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology volume two: Cognition, perception, and language*. Hoboken, NJ: John Wiley and Sons.

Shute, Valerie. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Article 4

Utilization of Probabilistic Cues in Children's Judgments

Utilization of Probabilistic Cues in Children's Judgments

Anna Lang
University of Erfurt, Germany

Author Note

Anna Lang, Department of Psychology, University of Erfurt, Germany

This research was supported by a research grant from the Faculty of Education at the University of Erfurt to the author. I am thankful to Tilmann Betsch, Niels Haase, Anne Lehmann and to the members of the Center for Empirical Research in Economics and Behavioral Sciences (CEREB) at the University of Erfurt for helpful comments and conversations.

Correspondence concerning this article should be addressed to Anna Lang, Department of Psychology, University of Erfurt, Nordhaeuser Strasse 63, 99089 Erfurt, Germany. E-mail: anna.lang@uni-erfurt.de

Declarations of interest: none

Word count: 9371

Abstract

Three studies investigate whether children at the age of six and nine utilize probabilities assigned to cues in judgments. They show that probability utilization in judgments emerges late and follows the same developmental trajectory as in decisions: Probability utilization is absent in 6-year-olds and emerging in 9-year-olds. Children's judgments show that their expectations of outcomes are not informed by probability until late elementary school age. Results remain consistent across studies and are unaffected by experience of participants or scale formats used to assess judgments. The findings contradict the notion that children utilize probabilities in judgments earlier than in decisions and highlight that deficits in probabilistic judgments persist until the age of nine.

Keywords: Probability learning, probabilistic inference, judgment and decision-making

Utilization of Probabilistic Cues in Children's Judgments

Utilizing probabilistic information correctly in judgment and decision making is a developmental challenge for children. In decisions, children do not exploit probabilistic information during the preschool period; this ability only emerges at elementary school age and continues to develop until late school age (Betsch & Lang, 2013; Betsch, Lang, Lehmann, & Axmann, 2014; Betsch, Lehmann, Lindow, Lang, & Schoemann, 2016; Cassotti, Aïte, Osmont, Houde, & Borst, 2014; Defoe, Dubas, Figner, & van Aken, 2014; Levin & Hart, 2003; Levin, Weller, Pederson, & Harshman, 2007; Mata, van Helversen, & Rieskamp, 2011; cf. Levin, Hart, Weller & Harshman, 2007; Girotto & Gonzalez, 2008). In judgments, however, children seem to consider probability at an earlier age (Acredolo, O'Connor, Banks, & Horobin, 1989; Anderson & Schlottmann, 1991; Schlottmann & Tring, 2005). For example, when 5-year-olds judge expected values of gambles their expectations are informed by the gamble's winning probability and they even integrate probability and value normatively correct (Schlottmann, 2001).

Decision tasks are often complex: They display multiple options and require children to successfully perform a sequence of processes in this environment (Betsch & Haberstroh, 2005). Decision makers must evaluate multiple options, compare those evaluations, and form behavioral intentions based on this comparison: for example, to choose the option highest in expected value, and reject others. They must then implement these intentions successfully while shielding all these processes from internal and external noise, such as intrusion of salient but irrelevant information. Thus, decisions may suffer from children's deficient working memory and executive functioning (Schlottmann & Wilkening, 2012; cf. Brand, Recknor, Grabenhorst, & Bechara, 2007 for executive function and decisions). This might explain why children often fall back on complexity-reducing strategies in decisions and ignore probability (Lang & Betsch, 2018a; Mata et al., 2011; Reyna & Brainerd, 1994).

Judgments tasks, on the other hand, are less challenging: They require one evaluation at a time, no comparison of evaluations, no deduction, and no implementation of behavioral intentions. Performance in judgments can thus differ considerably from performance in decision making tasks (Lichtenstein & Slovic, 1971); and children might utilize probabilities in judgments while not yet utilizing them in decisions (Anderson, 1991; Reyna & Brainerd, 1994; Schlottmann & Wilkening, 2012).

This paper investigates whether children consider probability for judgments in probabilistic inference tasks. In this paradigm, children consistently neglect probability in decisions until the age of nine (Betsch & Lang, 2013; Betsch et al., 2014, 2016; Lang &

Betsch, 2018a). Thus, it is well suited to investigate whether children utilize probability in judgments earlier than in decisions.

Probabilistic inference tasks mimic real-world decision problems in which available pieces of information—the cues—are probabilistically related to future decision outcomes (Brunswik, 1956). For example, whether a dog will be friendly and can be approached or will be dangerous and should be avoided cannot be predicted with certainty, but cues such as a wagging tail, perked-up ears and a relaxed mouth are probabilistically related to future dog behavior. These cues differ in their validity; that is, some cues predict outcomes better than others: A wagging tail in dogs can also indicate aggression and might thus not be a high valid cue. Normatively, available cues should be used to evaluate choice options and to optimize decisions (von Neumann & Morgenstern, 1944).

Adults master these tasks in varying contexts: They integrate multiple cues in accordance with the cues' validities or systematically rely on the most informative cue when appropriate (e.g., Bröder, 2000). Children below the age of nine, however, do not utilize cue validities for decisions. At the age of five to six, all children neglect cue validities in decisions; at the age of eight to nine, the majority of children neglect cue validities; and even older children's behavior diverges from adults' (Betsch & Lang, 2013; Betsch et al., 2014, 2016; Lang & Betsch, 2018a; Mata, von Helversen, & Rieskamp, 2011).

If children consider cue validities in judgments at an early age, this would support the notion of different developmental trajectories for judgments and decision making and might narrow down causes for children's deficient decisions. Based on previously reported superior performance in judgment tasks, children are expected to utilize cue validities in judgments earlier than in choices. Probability utilization should characterize judgments of 6- and 9-year-olds and adults, but judgments should become more accurate with age. Three studies use a child-friendly probabilistic decision paradigm, in which judgments were integrated. Decisions are not analyzed in detail but serve as a standard of comparison (see Lang & Betsch, 2018a, 2018b, for detailed analysis of decision data).

Method

Research Paradigm

The task is an adaption of the classic MouseLab Paradigm in adult decision making, in which three probabilistic cues predict binary outcomes of two options (Payne, Bettman, & Johnson, 1988; Figure 1C). The cues' binary predictive values can be used to infer options' expected value and maximize choice outcomes. Cues differ in their predictive validity: The first cue predicts outcomes correctly in three, the second cue in four and the third cue in five out of six cases (corresponding to p = .50, .66 and .83; Phase 1 in Procedure; Figure 1A). Participants encoded these differences in validity by observing the relative frequencies of correct predictions of each cue. Frequency formats are easier to understand than stated probabilities (Gigerenzer & Hoffrage, 1995). This allows even very young children to encode probabilistic information. To further facilitate recall, the probabilistic information was depicted graphically next to each cue as a magnitude of *smart points*; one for each correct prediction (Figure 1).

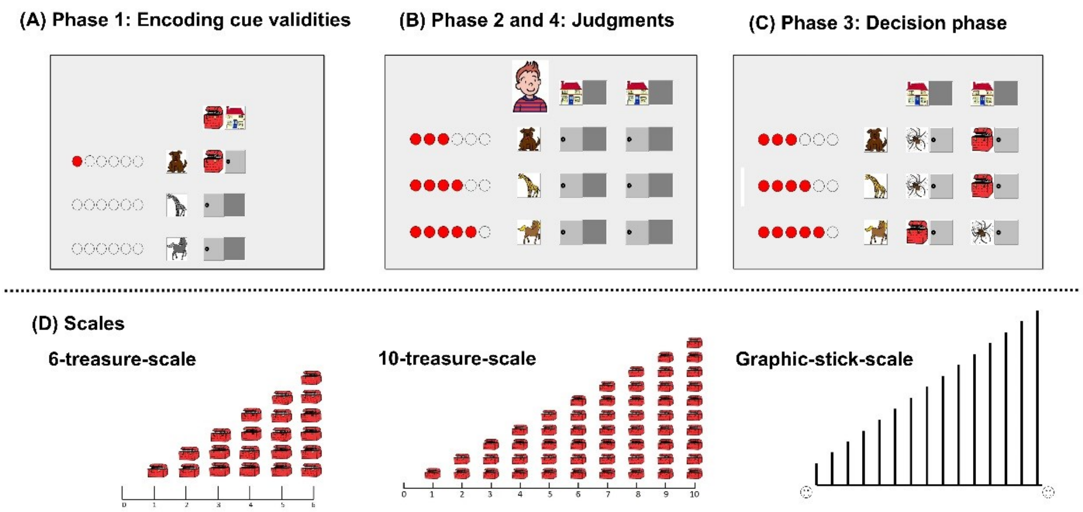


Figure 1. The research paradigm Figure shows stimulus material of the studies. During the encoding of cue validities (A), each cue predicts an outcome six times. If the prediction is correct, participants grant a smart point to the cue. Smart points subsequently represent each cue's validity (B, C). Participants predict outcomes for three fictional decision makers that play the game following one specific cue (B). Judgments are assessed a second time after the decision phase (C) and assessed on different scales (D). In the decision phase (C), participants repeatedly choose between two options and can inspect cue predictions before each choice. Cue predictions vary from choice to choice. Choice outcomes are available in feedback conditions.

Participants were then familiarized with the probabilistic decision game: the Treasure Hunt. It required repeatedly choosing between two options (i.e., houses), which could yield positive (i.e., treasures) or negative outcomes (i.e., spiders). For each decision, cue predictions could be inspected to optimize decisions. The probabilistic structure prescribed to follow the most valid cue's prediction to maximize outcomes.

Judgment Measurement

To assess children's utilization of probability in judgments, researchers test whether their expectations of outcomes are informed by available probabilistic information. For example, children may encounter a fictional player of probabilistic games with varying winning probabilities. They then predict outcomes or a related variable (e.g., happiness of the player) for different games (Anderson, 1991; Schlottmann, 2001; Schlottmann & Tring, 2005; Kalish, 2010; cf. Lagattuta & Saffran, 2013). We adapted this approach for the probabilistic inference task: Children encountered fictional players of the Treasure Hunt who played the game by relying either on the first, the second or the third cue (see Lagnado, Newell, Kahan, & Shanks, 2007, for similar procedures in adults). Children then predicted how many treasures each player would find. Normatively, predictive judgments should be informed by the cues' validities; that is, participants should expect better outcomes when the player relied on a more valid cue that predicted outcomes more accurately.

Young children are not yet able to provide stated probability judgments, but they can use nonverbal, context-adapted rating scales to express their predictions as graded expectancy (Acredolo et al., 1989; Anderson & Schlottmann, 1991; Ebersbach, 2009; Kalish, 2010; Schlottmann, 2001; Schlottmann & Tring, 2005). On two graphic scales visually similar to a bar chart, participants predicted outcomes as the frequency of treasures for each fictional player. The sample size was varied additionally: Participants judged how many treasures a player would find when looking for a treasure six times on the 6-treasure scale, ranging from 0 to 6 treasures, and when looking for a treasure ten times on the 10-treasure scale, ranging from 0 to 10 treasures (Figure 1D). Normatively, frequency judgments should follow expected value calculations (EV = cue validity × number of times the game is played, e.g., $.5 \times 6 = 3$). Correct predictions on the 6-treasure scale corresponded to correct predictions of the respective cue and thus to the magnitude of smart points associated with it (i.e., three, four, five). Relying only on smart point magnitudes can yield correct judgments on the 6- but not on the 10-treasure scale. Comparing judgments on both scales allows to test whether children only predict the magnitude they associate with each cue (e.g., only predict three treasures for the cue with three smart points, magnitude estimation hypothesis; Hoemann & Roess, 1982; Schlottmann & Wilkening, 2012).

Experience

Making correct predictions requires not only utilization of cue validities but also building up a valid mental model of the task using non-probabilistic information (e.g., knowing that only one option can be chosen in each trial). Incorrect understanding of the task can disguise correct utilization of probabilistic information (Anderson, 1991). Thus, predictive judgments were assessed before and after participants repeatedly played the Treasure Hunt themselves.

The availability of decision outcomes during the game varied: Without feedback, participants encoded all the relevant probabilistic information —the cue validities— in advance. For correct judgments, they must figure out logically how it affects the chance of outcomes. With feedback about decision outcomes, children can also actively engage with the probabilistic environment and learn how accurately cues predict outcomes through their own actions in the game. This should facilitate correct predictive judgments.

Study 1

Study 1 applied a 3 (age group: 6-year-olds vs. 9-year-olds vs. adults) × 2 (time of measurement: before vs. after choices) × 2 (feedback: no feedback vs. feedback) design. Time of measurement was varied as a within-subjects factor. Despite age-dependent differences in predictive judgments, all age groups should be sensitive to cue validities. Further, experiencing choices, in particular when feedback was included, should improve predictive judgments in child age groups.

Procedure

A trained experimenter supervised each participant and explained that the game's purpose was to find as many treasures as possible, which could be traded for presents afterwards. The experimenter started the computer game by introducing animals that served as cues as follows: "You don't have to find the treasures on your own; somebody will help you. Look, do you see the animals on the screen? You are allowed to choose three of them to help you play the treasure hunt." Participants then selected three out of the eight animals. The first chosen animal was placed on the top row of the screen, the last chosen on the bottom row to ensure that individual preference was never positively correlated with cue validity (Figure 1A).

Phase 1: Encoding cue validities. The concept of cue validity was framed during the game as the smartness of the animals. Accordingly, the experimenter explained that animals would predict hidden outcomes of a house but differed in smartness. Supervised by the experimenter, children encoded differences in cue validities: The experimenter started with the animal placed on the top row; she opened the door next to the cue to uncover the cue's prediction (e.g., "The dog says there is a treasure in the house."). Then she revealed the actual outcome by opening the house (e.g., "There is a treasure in the house."), and explained that for this correct prediction the animal was granted a smart point. In the following five trials, the participant was asked to indicate whether the animal should be granted a smart point. No child had difficulties in identifying correct predictions and granting smart points accordingly. For each smart point, the experimenter clicked on a point next to the cue which turned red

(Figure 1). After six trials, the experimenter summarized the cue's validity by counting smart points together with the child and stating that it had received three out of six smart points. Similarly, predictions of the second and third cue and corresponding outcomes were observed. Each cue's validity was displayed during the whole game by the corresponding magnitude of smart points (i.e., three, four and five, Figures 1B, C). Participants had to identify the most valid cue as a manipulation check before choices and were excluded if they failed to do so.

Phase 2: Predictive judgments before choices. The experimenter first explained the choice procedure in detail. She then presented the first of three fictional players following the predictions of the first animal (Figure 1B): "This is Tim. He plays the treasure hunt game just like you. He plays with the same animals, the dog, the giraffe, and the horse. Tim is only interested in what the dog says. The other animals do not interest him at all. Tim only relies on the dog." While explaining, the experimenter pointed to the boy's picture and the corresponding cue (Figure 1B). She continued to explain that the boy looked for a treasure six times—to make this clear to the child the experimenter slowly counted with her fingers from one to six—and placed the 6-treasure scale in front of the child (Figure 1B). The experimenter explained the meaning of each scale value. Children then made their predictive judgments by pointing to the corresponding treasure bar. Next, the experimenter presented the 10-treasure scale in a similar fashion. The procedure was then repeated for other fictional players following the medium and high valid cue (named Max and Hannes; order was counterbalanced; all names are common German names for boys).

Phase 3: Choices. Experimenters ensured in two training trials that the child understood the game. Participants then played the game 24 times, each time choosing between two options. They were either immediately informed about the outcome of the chosen option in feedback conditions, or not informed in conditions without feedback. The feedback structure matched the encoded cue validities; for example, positive predictions by the high valid cue yielded successful choices in 83% (n = 20) of trials.

Phase 4: Judgments after choices. After having played the game, participants provided the same predictive judgments as before. Additionally, participants' recall of cue validities was assessed. First, they pointed out the high and medium valid cue and then judged the cue validities on a Graphic-stick scale, which is visually similar to the treasure scales used for predictive judgments (adapted from the Wooden-stick scale, Schlottmann, 2001; adapted by Lindow, 2014; Figure 1). Experimenters again carefully explained the scale and participants pointed to scale values to indicate their judgment of each cue's validity.

Finally, participants were rewarded with prizes contingent on their choice performance. The whole study took about 25 minutes.

Procedure for adults. Adults were informed in advance that they served as a control group in a children's study and would receive money instead of prizes. Otherwise, the procedure was identical.

Sample

Eighty 6-year-olds (49 female, *M* = 69.2 months, *SD* = 5.3), 62 9-year-olds (28 female, *M* = 104.4 months, *SD* = 4.7), and 53 adults (39 female, *M* = 258.59 months, *SD* = 31.96) participated. Two 6-year-olds, two 9-year-olds, and three adults were excluded because they did not complete the study or data were missing. Seven 6-year-olds and one adult failed the manipulation check. The final sample included 71 6-year-olds (46 female, *M* = 69.1 months, *SD* = 5.3), 60 9-year-olds (27 female, *M* = 104.5 months, *SD* = 4.7), and 49 adults (36 female, *M* = 267.5 months, *SD* = 31.3). Child participants were recruited from German schools or kindergartens with varying socioeconomic background. Parents consented to children's participation. Adult participants were students of different majors.

Results

Mean predictive judgments. Figure 2A displays mean judgments on each scale as a function of cue validity. Visual inspection reveals that only adults' judgments closely matched the normative values on both scales. Small confidence intervals demonstrate the low variance in adults' judgments. Nine-year-olds' predictive judgments discriminated between cues—that is, they expected better outcomes for more valid cues—but diverged considerably from normative values and displayed more variance. Six-year-olds' judgments did not systematically discriminate between cues. Their judgments consistently exceeded older children's and adults' judgments; that is, 6-year-olds seemed to be over-confident and generally expected better outcomes. However, the judgments on the 10-treasure scales show that all age groups adapted their judgments to the corresponding sample size; that is, children did not only stick to mental magnitudes associated with cues.

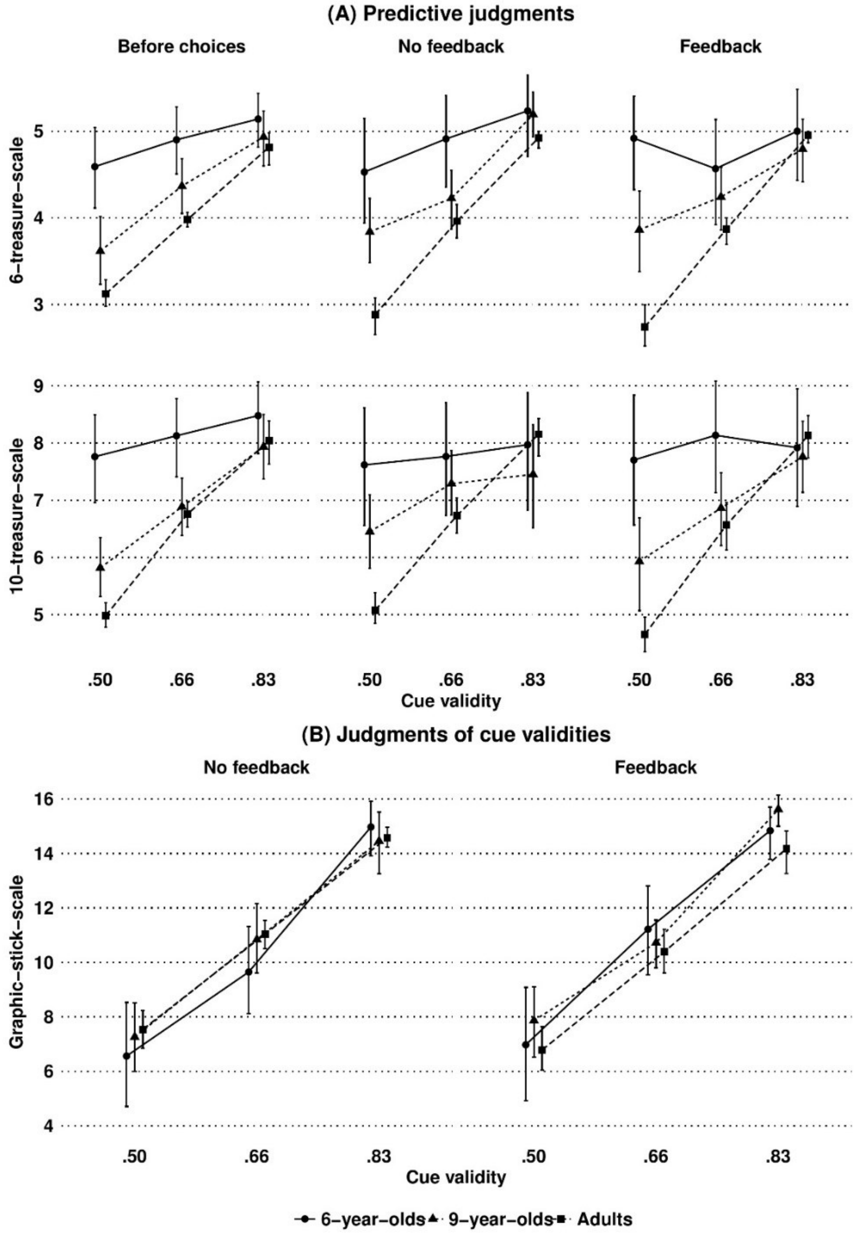


Figure 2. Judgments in Study 1 as a function of cue validity. Figure shows predictive judgments on the 6-treasure scale (Range: 0–6) and the 10-treasure scale (Range: 0–10) before choices for all participants; and judgments on both scales after choices for participants in each feedback condition. Bars represent 95% bootstrap confidence intervals.

Sensitivity and accuracy of predictive judgments. Whether judgments were sensitive to cue validities was assessed by calculating the Pearson correlation coefficient between judgments and normative expectations for each age group. High correlations indicate that participants in this age group relied on cue validities normatively; medium correlations indicate that participants considered cue validities but their judgments diverged from normative expectations. If participants did not rely on cue validities, no correlation is expected. The root-mean-square deviations (*RMSD)* from normative expectations served as a measure of accuracy. Lower values indicate higher accuracy.

Table 1 displays sensitivity and accuracy for each age group. Adults were highly sensitive to cue validities and accurate in their predictive judgments, 9-year-olds were sensitive to cue validities but less accurate than adults, and younger children were insensitive to cue validities and the least accurate. Two ANOVAs revealed that age strongly affected accuracy on the 6-treasure scale, $F(2, 174) = 102.64$, $p < .001$, $\eta^2_G = .47$, as well as on the 10-treasure scale, $F(2, 174) = 174.58$, $p < .001$, $\eta^2_G = .57$.

Irrelevance of experience. The ANOVAs further tested whether the accuracy in children's judgments improved with decision experience and with or without outcome feedback. The 6-treasure scale showed a marginally significant effect of Time, $F(1, 174) = 3.64$, $p = .058$, $\eta^2_G = .005$; and two interaction effects, Time × Feedback, $F(1, 174) = 3.24$, $p = .074$, $\eta^2_G = .01$; and Time × Feedback × Age, $F(2, 174) = 2.37$, $p = .097$, $\eta^2_G = .01$, all other *p*s ≥ .49. Separate analyses for each group revealed that 6-year-olds and adults were not affected by feedback on either scale, all *p*s ≥ .30. Nine-year-olds' judgments on the 6-treasure scale were more accurate without feedback, $M_{\text{No feedback}} = 0.79$, 95% Bootstrap[1] CI [0.54, 1.05]; $M_{\text{Feedback}} = 1.05$, [0.84, 1.30]; Time*,* $F(1, 58) = 3.23$, $p = .078$, $\eta^2_G = .01$; Time × Feedback*,* $F(1, 58) = 5.82$, $p = .019$, $\eta^2_G = .23$; and more sensitive, $M_{\text{No Feedback}} = 0.51$, [0.34, 0.67], $M_{\text{Feedback}} = 0.33$, [0.12, 0.51]. Sensitivity and accuracy on the 10-treasure scale was unaffected by feedback condition, all *p*s ≥ .26.

Discrimination between cues. Younger children's failure to consider cue validities in predictive judgments was not due to memory deficits. Most individuals successfully recalled how cues differed in validity. 78% of 6-year-olds ($n = 55$), 80% of 9-year-olds ($n = 48$) and 96% of adults ($n = 47$) correctly identified the high, medium and low valid cue, $\chi^2(2, N = 180) = 7.83$, $p < .02$, Cramer's $V = .21$. Figure 2B shows mean judgments of cue validities on the Graphic-stick scale and Table 1 presents respective sensitivity indices. All age groups discriminated between and were sensitive to different cue validities in both eedback conditions, though children less so than adults.

[1] All CIs are based on 10000 bootstrap samples.

Table 1

Sensitivity and accuracy of predictive judgments in Studies 1 and 2.

Study	Scale	Time	6-year-olds				9-year-olds				Adults			
			Sens.	95% CI	Acc.	95% CI	Sens.	95% CI	Acc.	95% CI	Sens.	95% CI	Acc.	95% CI
1	6-treasure	Before	.13	[.01, .25]	1.90	[1.72, 2.06]	.37	[.24, .51]	1.10	[0.89, 1.32]	.79	[.64, .92]	0.25	[0.13, 0.40]
		After	.09	[-.03, .22]	1.83	[1.63, 2.03]	.42	[.29, .54]	0.91	[0.74, 1.09]	.89	[.85, 94]	0.21	[0.11, 0.33]
	10-treasure	Before	.10	[-.04, .23]	3.30	[3.08, 3.51]	.38	[.23, .51]	1.88	[1.65, 2.12]	.76	[.62, .87]	0.80	[0.62, 1.10]
		After	.04	[-.09, .18]	3.29	[3.05, 3.54]	.28	[.13, .42]	1.94	[1.71, 2.17]	.83	[.77, .89]	0.73	[0.61, 0.88]
	Graphic-stick	After	.56	[.46, .65]	—		.68	[.58, .77]	—		.86	[.81, .90]	—	
2	6-treasure	Before	.09	[-.04, .23]	1.91	[1.73, 2.08]	.38	[.23, .49]	1.02	[0.84, 1.18]	.89	[.82, .96]	0.13	[0.05, 0.21]
		After	.05	[-.08, .19]	2.04	[1.83, 2.22]	.49	[.39, .59]	0.99	[0.83, 1.17]	.92	[.89, .95]	0.14	[0.07, 0.22]
	10-treasure	Before	.19	[.05, .31]	3.02	[2.78, 3.25]	.33	[.22, .44]	2.00	[1.82, 2.24]	.78	[.71, .85]	0.82	[0.68, 0.99]
		After	.07	[-.06, .21]	3.34	[3.07, 3.60]	.35	[.23, .46]	1.96	[1.73, 2.22]	.83	[.78, .87]	0.76	[0.66, 0.87]
	Graphic-stick	After	.35	[.23, .48]	—		.70	[.62, .77]	—		.81	[.77, .85]	—	

Note. Before = Before choices; After = After choices; Sens. = Sensitivity = Pearson Correlation Coefficients for each age group; Acc. = Accuracy = Mean *RMSD* for each age group. Confidence intervals are bootstrapped with 10000 samples.

Discussion

Children were exopected to utilize cue validities for predictive judgments earlier than in decision making. However, while all age groups discriminated between cues based on encoded validities, whether and how accurately participants used this information to predict outcomes varied according to age. As expected, adults' judgments were highly sensitive, accurate and close to normative expectations; 9-year-olds' predictions were also sensitive to cue validities but less accurate. Six-year-olds did not consider cue validities at all in their judgments, but were generally over-confident; that is, they expected better outcomes than older children and adults.

Experience did not increase children's utilization of cue validities in predictive judgments. Even after participants had made their own decisions and experienced actual outcomes they did not rely on this information when predicting outcomes. However, learning opportunities were limited: 24 trials may not provide enough experience and monitoring the relationship between cue predictions and choice outcomes was restricted to their own decision outcomes. Thus, the feedback was fragmentary and possibly not suited to experience differences in cues' accuracies correctly (Fiedler, 2008). A second study, therefore increased the number of choices and fully informed participants about outcomes of chosen and non-chosen options.

Study 2: Experience

The 24 trials of the choice phase were replicated three times in a randomized order. In addition, a full feedback condition was added resulting in a 3 (age group: 6-year-olds vs. 9-year-olds vs. adults) × 3 (feedback: no feedback vs. selective feedback vs. full feedback) × 2 (time of measurement: before vs. after choices) design. In full feedback conditions, participants inspected the chosen and the foregone option's outcome immediately after each choice. Otherwise, the procedure and data analysis were identical to Study 1. Age-dependent differences in sensitivity and accuracy of judgments should replicate, but children's judgments should be more sensitive and accurate when they made 72 decisions, each displaying cue predictions and full outcomes.

Sample

The sample consisted of 81 6-year-olds (34 female, $M = 68.61$ months, $SD = 5.15$), 91 9-year-olds (55 female, $M = 107.25$ months, $SD = 8.64$), and 85 adults (66 female, $M = 253.20$ months, $SD = 38.52$) from the same subject pool. One 6-year-old and one 9-year-old were excluded because they did not complete the study. Eight 6-year-olds and three 9-year-old failed the manipulation check and were excluded. The final sample consisted of 72

6-year-olds (31 female, M = 68.89 months, SD = 5.27), 87 9-year-olds (52 female, M = 106.91 months, SD = 8.45) and 85 adults.

Results

Mean predictive judgments. Figure 3A displays predictive judgments. The results were very similar to Study 1. Adults' judgments closely matched normative values. Adults and 9-year-olds discriminated between cues; 6-year-olds did not but, again, displayed over-confidence in all judgments. Predictions on the 10-treasure scale were higher in all age groups, illustrating again that children did not simply rely on the mental magnitudes of smart points. Obviously, children's judgments varied between feedback conditions, but importantly, 6-year-olds did not systematically discriminate between cues in any feedback condition, while 9-year-olds did so in every feedback condition.

Sensitivity and accuracy of predictive judgments. Age-dependent differences in sensitivity and accuracy were replicated (Table 1). ANOVAs on accuracy yielded equal effect sizes for age on both scales; 6-treasure scale, $F(2, 232) = 186.31, p < .001, \eta^2_G = .55$, all other ps $\geq .31$; 10-treasure scale, $F(2, 233) = 156.44, p < .001, \eta^2_G = .51$.[1]

Irrelevance of experience. In line with visual inspection of Figure 3A, sensitivity and accuracy after decisions were not improved and did not differ between feedback conditions in any age group, all ps $\geq .20$. So contrary to expectations, even after an increased number of decisions in which all possible outcomes were experienced, judgments did not reflect more consideration of cue validities.

Discrimination between cues. 67% of 6-year-olds (n = 48), 78% of 9-year-olds (n = 68) and 95% of adults (n = 81) correctly identified the high, medium, and low valid cue at the end of the game, $\chi^2(2, N = 180) = 7.83, p < .02$, Cramer's $V = .21$. Figure 3B shows mean judgments of cue validities on the Graphic-stick scale. All groups' judgments were sensitive to differences in cue validities. Again, this makes it unlikely that younger children's memory deficits can fully explain the insensitivity in predictive judgments, even though their performance was inferior to adults with 6-year-olds worse than 9-year-olds (Table 1). In general, the recall of cue validities was worse than in Study 1, which might be due to extended duration of the second experiment.

[1] In addition to Age, an interaction between Age and Time, $F(2, 232) = 3.58, p = .029, \eta^2_G = .007$ was obtained, all other ps $\geq .20$. It indicated that accuracy varied before and after the choices as a function of age group. Specifically, 6-year-olds' generally inaccurate judgments were even more inaccurate after they made choices themselves, $t(72) = -2.16, p = .034$.

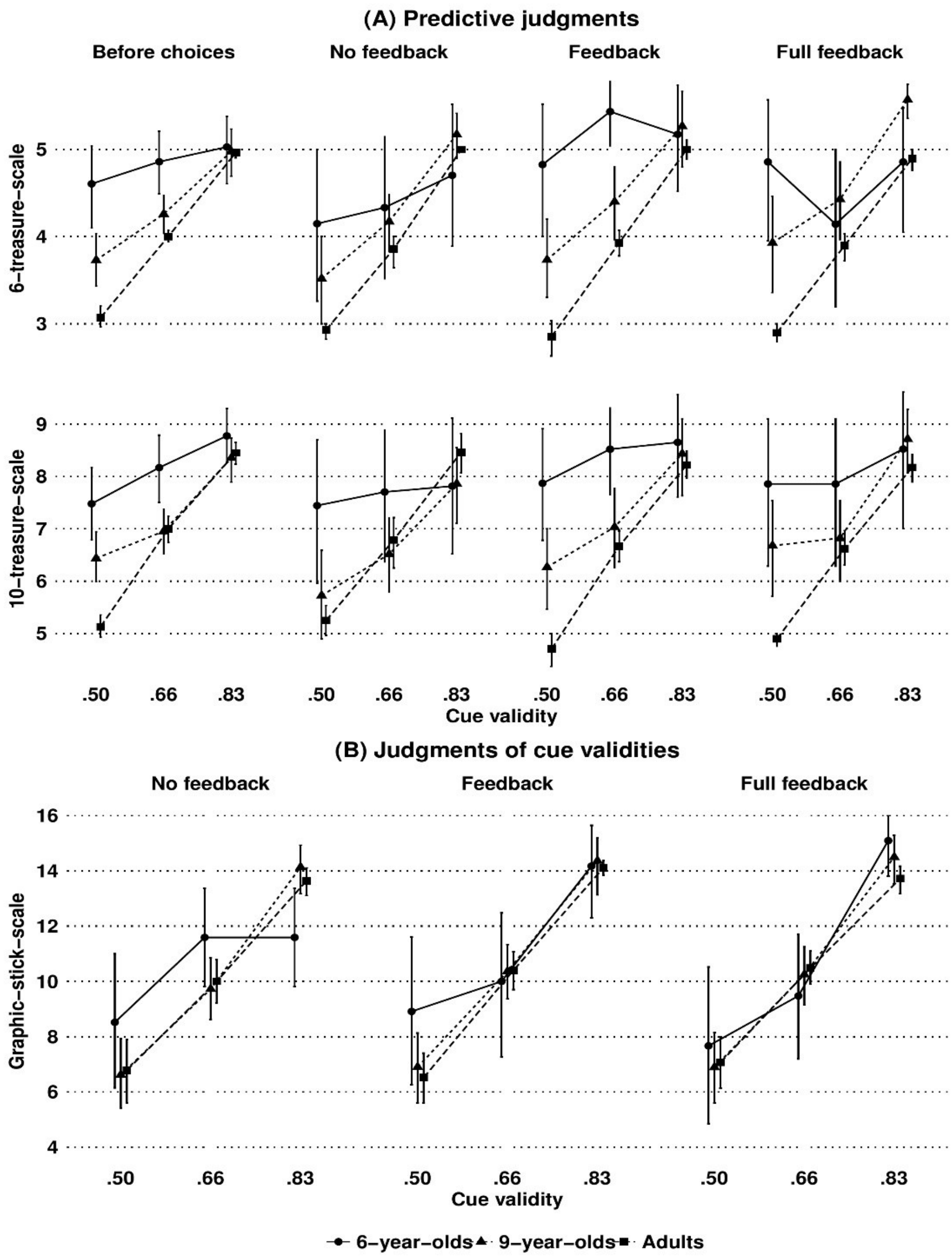


Figure 3. Judgments in Study 2 as a function of cue validity. Predictive judgments as a function of cue validity for each scale (A). Displayed are judgments for all participants before choices and judgments after choices for participants in each feedback condition. Below, judgments for cue validities are shown as a function of cue validity for each feedback condition (B). Bars represent 95% bootstrap confidence intervals.

Discussion

The results of Study 1 were replicated; increased learning opportunities with more decision trials and full display of all outcomes did not alter these findings. The final study ensured that the poor performance was not due to the scales. The treasure scales displayed bars of treasures but also included numeric values to facilitate normatively correct responses (Figure 1D). Numeric values, however, might have impeded intuitive processes and induced analytical thinking (Anderson, 1991; Reyna & Brainerd, 1994; Windschitl & Wells, 1996). Thus, it cannot be ruled out that children are be able to utilize cue validities for predictive judgments intuitively (Schlottmann & Wilkening, 2012), but not to map these judgments on numeric scales. Study 3 therefore varied the scale format for predictive judgments using an intuitive and analytical scale. If children can intuitively utilize cue validities, predictive judgments should improve on the intuitive scale format.

Study 3: Scale Format

The study replicated the selective feedback condition of Study 1, but predictive judgments were assessed only once after choices. Analytical judgments used the same treasure scales as in previous studies. Intuitive judgments employed the non-numeric Graphic-stick scale for predictive judgments (Figure 1D). The study thus implemented a 2 (age group: 6-year-olds vs. 9-year-olds) × 2 (scale format: Graphic-stick vs. Treasure scales) between-subjects design.

Sample

Forty-two 6-year-olds (20 female, M = 70.46 months, SD = 4.92) and 41 9-year-olds (28 female, M = 110.93 months, SD = 6.57) from the same subject pool participated in the study. Six 6-year-olds and one 9-year-old were excluded because they failed the manipulation check. The final sample consisted of 36 6-year-olds (16 female, M = 70.54 months, SD = 4.82) and 40 9-year-olds (28 female, M = 110.90 months, SD = 6.66).

Procedure

The procedure for analytical judgments on the treasure scales was identical to Study 2's feedback condition. For intuitive judgment condition, children used the Graphic-stick scale for predictive judgments and judgments of cue validities. The scale minimum for the former was introduced as *finding no treasures*, the scale maximum was introduced as *finding as many treasures as possible*, the scale values between as *each means finding a few more treasures*. Children were discouraged from counting the sticks. Beyond this, instructions were identical in both conditions. Children predicted success when following one specific cue in six choices and in ten choices, respectively.

Results

Mean predictive judgments. Figure 5 shows mean predictive judgments for the two conditions. Nine-year-olds, but not 6-year-olds, discriminated between cues in predictive judgments regardless of the scale. Six-year-olds' judgments displayed over-confidence on both scales. Interestingly, 9-year-olds' judgments on the Graphic-stick scale closely matched predictions derived from analytically calculating the expected value as a product of the cue validity and sample size (i.e., mean judgments ranged between the scale values of six and ten for predicted outcomes in ten decisions).

Sensitivity of predictive judgments. Six-year-olds' judgments were insensitive to cue validities on all scales, $r_{\text{Treasure scale, 6 choices}} = .05$, 95% Bootstrap CI [-.28, .28], $r_{\text{Graphic-stick scale, 6 choices}} = -.18$, [-.42, .37], $r_{\text{Treasure scale, 10 choices}} = .14$, [-.12, 37], $r_{\text{Graphic-stick scale, 10 choices}} = .07$, [-.19, .32]. Nine-year-olds' judgments were sensitive to cue validities on all scales, $r_{\text{Treasure scale, 6 choices}} = .50$, [.28, .70], $r_{\text{Graphic-stick scale, 6 choices}} = .54$, [.33, .73], $r_{\text{Treasure scale, 10 choices}} = .43$, [.19, 62], $r_{\text{Graphic-stick scale, 10 choices}} = .51$, [.33, .69].

Discrimination between cues. 75% of 6-year-olds and 93% of 9-year-olds could point out the high, medium and low valid cue at the end of the game, $\chi^2(2, N = 76) = 9.78$, $p = .008$, Cramer's $V = .36$. Mean judgments of cue validities on the Graphic-stick scale reflected the encoded cue validities in both age groups, $r_{\text{6-year-olds}} = .41$, 95% Bootstrap CI [.23, .56], $r_{\text{9-year-olds}} = .66$, [.54, .77], but older children outperformed younger children.

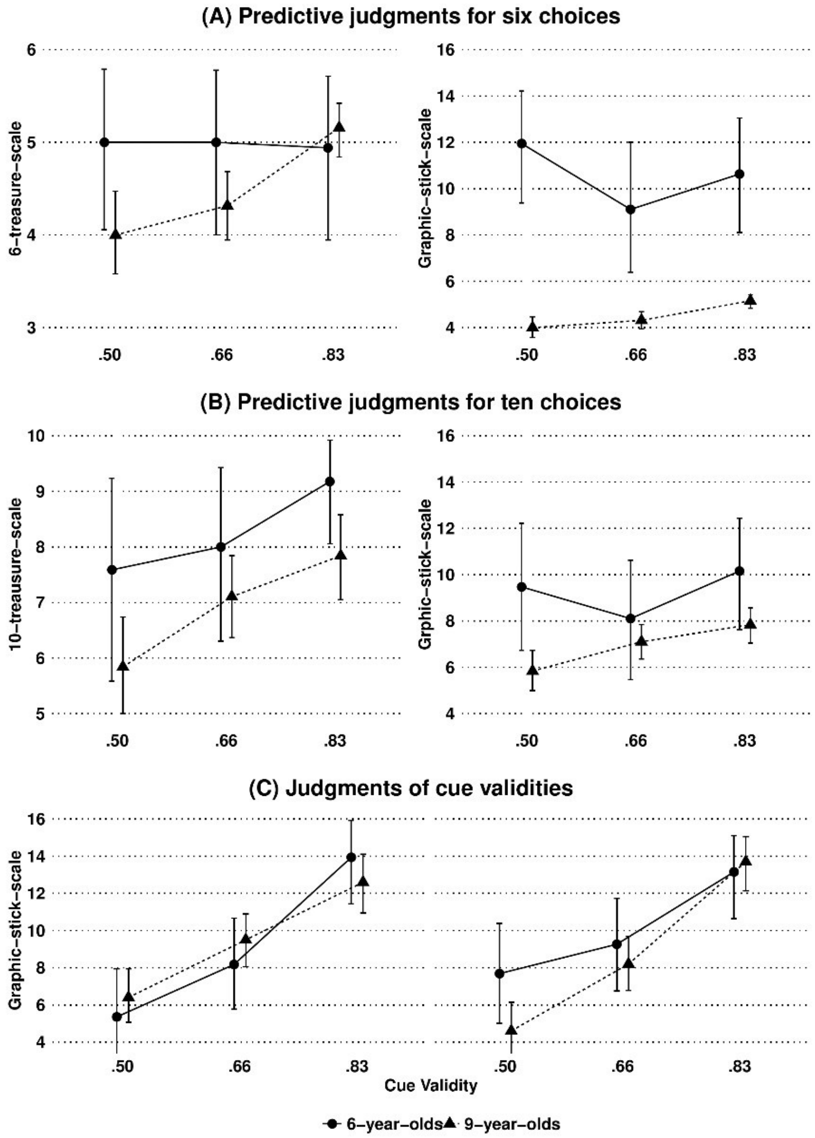


Figure 4. Judgments in Study 3 as a function of cue validity.Predictive judgments for six choices (A) on the 6-treasure scale in the analytical-judgment condition and on the Graphic-stick scale in the intuitive-judgment condition. Predictive judgments for ten choices on the 10-treasure scale and Graphic-stick scale (B). Cue validity judgments are displayed as a function of cue validity (C) for the analytical-judgment condition on the left and intuitive-judgment condition on the right.

Discussion

In Study 3, predictive judgments were either assessed on the partly numeric treasure scales or on a more intuitive scale, the Graphic-stick scale. On both scales the results from Studies 1 and 2 were replicated: 6-year-olds' judgments were insensitive to cue validities while 9-year-olds' were sensitive. Thus, the poor performance of younger children cannot be attributed to an unsuitable scale that prohibits the expression of correct intuitive judgments. Intriguingly, 9-year-olds' predictions seemed to be derived from analytical processes rather than from intuition, even when intuitive predictions were encouraged by the procedure. Both age groups recalled the differences in cue validities successfully, even though older children's recall is more accurate.

Summary of Results

Summarizing findings in a small-scale meta-analysis provides better estiamations of the observed effects in four largely similar studies (*N* = 664; Cummings, 2012). The additional study is not reported here, but was identical to Study 1 with the exception that during the choice phase, the cue predictions were hidden and had to be uncovered by participants.[2] All analyses were computed using the metafor Package in R (Viechtbauer, 2010).

Age-Dependent Differences in Cue Utilization

Adults are highly sensitive for cue validities in their predictive judgments, 9-year-olds are less sensitive, and 6-year-olds are insensitive. The data consistently showed large correlation coefficients for adults, medium-sized correlations for 9-year-olds and no correlation for 6-year-olds (Table 1). We additionally analyzed sensitivity for subsamples of 6-year-olds, excluding individuals with invariant judgments or individuals with false recall of cue validities in direct questions or on the Graphic-stick scale. However, neither of those subsamples showed sensitivity in predictive judgments.

A meta-analysis estimated the proportion of individuals in each age group that consistently utilized cue validities normatively in predictive judgments, that is, individuals whose judgments were sensitive to cue validities on the 6- and 10-treasure scales, and adapted to sample size.[3] 2% of 6-year-olds, 31% of 9-year-olds, and 83% of adults utilized cue validities in this way (see Table 2). Thus, although group-level analyses showed sensitivity in 9-year-olds, only a minority of children that age were truly consistent in cue utilization like the vast majority of adults.

---

[2] Study 4's sample, 56 preschoolers (23 female, *M* = 68.0 months, *SD* = 7.3), 53 elementary schoolers (24 female, *M* = 107.6 months, *SD* = 5.8) and 55 adults (45 female, *M* = 282.00 months, *SD* = 44.72).

[3] That is, more treasures are predicted when the game is played ten times. On the non-numeric Graphic-stick, no adaption to sample size is expected normatively, thus, all subjects with sensitive judgments were included.

Table 2

Results of random effects meta-analyses of judgment data

	Proportions of normative cue utilization					
	6-year-olds		9-year-olds		Adults	
Study	%	95% CI	%	95% CI	%	95% CI
1	1	[0, 4]	28	[17, 40]	86	[76, 95]
2	4	[0, 9]	32	[22, 42]	82	[74, 90]
3	0	[0, 5]	34	[26, 43]	—	
4	5	[0,11]	28	[16, 40]	82	[71,92]
RE-Model	**2**	**[0, 4]**	**31**	**[26, 37]**	**83**	**[78, 89]**
I^2	16%		0%		0%	

	Experience effect on cue utilization					
	6-year-olds		9-year-olds		Adults	
Study, Scale	*q*	95% CI	*q*	95% CI	*q*	95% CI
1, 6-treasure	-.16	[-.43, .12]	-.22	[-.52, .08]	.14	[-.19, .47]
1, 10-treasure	-.01	[-.28, .26]	.16	[-.13, .46]	-.01	[-.34, .32]
2, 6-treasure	-.01	[-.33, .32]	-.22	[-.52, .08]	-.14	[-.44, .17]
2, 10-treasure	.09	[-.23, .42]	-.01	[-.31, .29]	.50	[.19, .81]
4, 6-treasure	.07	[-.24, .37]	-.14	[-.45, .18]	.10	[-.21, .41]
4, 10-treasure	.16	[-.15, .47]	-.44	[-.76, -.13]	.10	[-.21, .41]
RE-Model	**.01**	**[-.11, -.14]**	**-.14**	**[-0.30, .03]**	**.11**	**[.06, .29]**
I^2	0%		43%		45%	

Note. RE-Model = Random Effects Models, are in boldface; I^2 indicates the percentage of variation due to heterogeneity; CI = Confidence Interval; *q* = Cohen's *q* = difference of Fisher-z-transformed sensitivity after decisions with and without feedback.

Utilization in Judgments vs. in Decisions

Several criteria can be considered to determine whether participants utilized probabilistic cues in decisions (Betsch & Lang, 2013; Betsch et al., 2014, 2016, 2018; Lang & Betsch, 2018a; see these articles for detailed descriptions of the respective evaluative criteria). The first criterion, performance above chance level, prescribes (in the established non-compensatory decision environment) to choose in line with the most valid cue's predictions in more than half of the decisions. Only 6-year-olds did not achieve this (Table 3).

A stricter criterion requires participants to systematically utilize probabilistic cues in decisions, which only 7% of 6-year-olds, but 35% of 9-year-olds and 97% of adults did. An even stricter criterion requires participants to use probability-based decision strategies that consider cue validities, and are described in the decision literature (Lexicographic and Weighted Additive Rule, e.g., Payne et al., 1988). These strategies are very rarely found in 6-year-olds, but in 24% of 9-year-olds and 60% of adults (Table 3).

Table 3

Results of random effects meta-analyses of decision data

Above chance performance

	6-year-olds		9-year-olds		Adults	
Study	*d*	95% CI	*d*	95% CI	*d*	95% CI
1	-0.35	[-0.69, -0.02]	0.37	[0.01, 0.73]	2.18	[1.68, 2.68]
2	-0.55	[-0.88, -0.22]	0.48	[0.18, 0.78]	2.61	[2.20, 2.02]
3	-0.70	[-1.17, -0.22]	0.42	[-0.02, 0.86]	—	
4	-0.35	[-0.72, 0.03]	0	[-0.37, 0.37]	1.5	[1.41, 1.59]
RE-Model	**-0.46**	**[-0.65, -0.28]**	**0.32**	**[0.11, 0.54]**	**2.10**	**[1.45, 2.74]**
I^2	95%		30%		84%	

Proportions of systematic cue utilization

	6-year-olds		9-year-olds		Adults	
Study	%	95% CI	%	95% CI	%	95% CI
1	17	[8, 26]	42	[29, 55]	98	[94, 100]
2	10	[3, 17]	42	[37, 47]	99	[97, 100]
3	0	[0, 5]	38	[33, 43]	—	
4	5	[0, 11]	21	[17, 25]	89	[81, 97]
RE-Model	**7**	**[0, 14]**	**35**	**[24, 46]**	**97**	**[92, 100]**
I^2	73%		72%		67%	

Proportions of probability-based strategies

	6-year-olds		9-year-olds		Adults	
Study	%	95% CI	%	95% CI	%	95% CI
1	0	[0,3]	15	[5, 25]	47	[32, 62]
2	1	[0, 4]	40	[29, 51]	71	[61, 80]
3	6	[0, 11]	28	[14, 41]	—	
4	0	[0,3]	13	[4, 22]	60	[47, 73]
RE-Model	**3**	**[0,8]**	**24**	**[11, 36]**	**60**	**[47, 74]**
I^2	63%		82%		73%	

Note. RE-Model = Random Effects Models, are in boldface; I^2 indicates the percentage of variation due to heterogeneity; CI = Confidence Interval.

Arguably, a comparison of utilization in judgments and decisions lacks the same criterion. Additionally, utilization in decisions varies considerably between studies whereas utilization of judgments varies less (see I^2 in Table 2 and 3). But whether the criterion for decisions is above-chance or systematic utilization of cue validities, or the use of probability-based strategies, probability utilization is absent in 6-year-olds' and emerging in 9-year-olds' decisions. Thus, and most importantly, the same developmental pattern characterizes the utilization of probabilistic cues in judgments and in decisions: absence in 6-year-olds and emergence in 9-year-olds.

Irrelevance of Experience

We suspected that children's utilization of probabilistic information in judgments would improve if they experienced decision outcomes themselves. However, providing outcomes did not benefit judgments. A comparison of Fisher-*z*-transformed sensitivity in conditions with selective feedback with conditions without feedback obtained overall no effect of experiencing outcomes (Table 2). Even when children gained experience with the probabilistic game through their own actions and encountered feedback that reproduced provided cue validities, their expectations of outcomes did not become more accurate.

Overconfidence in 6-Year-Olds

Unexpectedly, 6-year-olds' judgments exceed those of older children and adults indicating that younger children overall expected better outcomes (Figures 2–4). This age-dependent over-confidence was consistent and stable across studies: 6-year-olds compared to 9-year-olds, RE $\text{Model}_{\text{6-treasure scale}}$, $d = 0.36$, CI [0.18, 0.56]; $I^2 = 0\%$; RE $\text{Model}_{\text{10-treasure scale}}$, $d = 0.67$, [-0.07, 1.40], $I^2 = 74\%$; compared to adults, RE $\text{Model}_{\text{6-treasure scale}}$, $d = 0.79$, [0.61, 0.97], $I^2 = 21\%$; RE $\text{Model}_{\text{10-treasure scale}}$, $d = 1.08$, [0.57, 1.59], $I^2 = 59\%$.

Analysis of individual data patterns provides further explanations: 21% of 6-year-olds indifferently selected the maximum scale value in all judgments, RE Model, 21%, [15, 27], $I^2 = 16\%$. When these individuals were excluded, 6-year-olds' judgments no longer exceeded those of 9-year-olds and only slightly those of adults. Younger children's over-confidence at the group level is thus mainly caused by a subsample of 6-year-olds that chose scale maxima in all judgments.

General Discussion

In three independent studies, 6-year-olds, 9-year-olds and adult controls were confronted with a probabilistic inference game with three cues differing in validity ($p = .5$, .66, .83). Participants encoded cue validities in advance and then judged the expected outcomes when playing the game relying on each of the cues. We expected children's judgments to reflect probability utilization already at the age of six and to only become more accurate with age. Results did not support this hypothesis: 6-year-olds fully and 9-year-olds partly neglect probability in judgments. A small-scale meta-analysis demonstrates that the observed developmental pattern is very stable and not affected by decision experience. Additionally, it was independent of the scale format used for judgments.

The results not only contradict the notion that children utilize probabilities in judgments earlier than in choices—in fact, they match those observed in decision making in the same environment very closely (Betsch & Lang, 2013; Betsch et al., 2014, 2016; Lang &

Betsch, 2018a)—they also contradict previous findings showing that children at the age of five rely on winning probabilities for evaluating gambles (Acredolo et al., 1989; Anderson & Schlottmann, 1991; Schlottmann, 2001; Schlottmann & Tring, 2005). On the other hand, they are in line with research showing that, although children encode co-variation in variables, they do not automatically utilize probabilistic relations for predictions (Kalish, 2010).

Probabilistic information was available in different formats. Children observed a sequence of cue predictions that were either correct or false, and assigned smart points for each correct prediction. Probabilistic information was thus learned in a supervised fashion and represented by the magnitude of smart points. This resembles previous research where probabilistic information is demonstrated and then offered in a summarized and salient format to use it in judgments (e.g., Schlottmann, 2001). In addition, children were further able to actively engage with the probabilistic environment and experience outcomes of their own actions. The feedback structure additionally entailed the probabilistic information and reflected it during choices children made themselves. We expected children to benefit from this additional experience but this was not the case. Regardless of whether children received only a summarized format of probability or could acquire experienced-based formats through their own actions, they did not utilize this information until the age of nine. The probability format did not affect the developmental trajectory from age six to nine.

Challenging Probabilistic Environments

The reported studies demonstrate that demands of the decision process (e.g., forming and implementing choice intentions, consistently applying choice strategies) cannot account for previously observed probability neglect as it also characterizes judgments. Consequently, eligible causes must account for deficient performance in judgments *and* in decisions. Two specific aspects of the environment—high complexity and low contiguity—are potential candidates (Schlottmann, 2011; Wohlwill, 1968).

The paradigm creates high complexity in two ways. First, imagine, for example, a simple gamble. There, the probabilistic information is embedded in the frequency distribution of a single variable; that is, the outcome of said gamble. In probabilistic inferences, on the other hand, the probabilistic information is embedded in the relation of two variables (cue predictions and outcomes) and only the contingency of both is informative. While children had no difficulties encoding probabilistic relations between the two variables, utilizing them for predictions requires an understanding of the conditional probability *p* (positive outcome | positive cue prediction). In line with this, Kalish (2012) observed in a similarly complex paradigm that 7-year-olds were able to encode the probabilistic relationship between a bird's diet and the color of its eggs but not to use this information consistently to predict egg colors.

Second, the environment provides a highly complex stimulus field to navigate through: Three cues each with an assigned probability, two options, and additional irrelevant information, such as the position of cues or the animals representing the cues. Valid judgments require prioritizing of relevant information, which is harder for younger children (Lang & Betsch, 2013). Performance may be generally better in perceptually poor environments that offer only relevant information, for example, when only one option with a probability and a value is displayed (e.g., Schlottmann, 2001).

The environment further offers low contiguity between probability and the variable to be judged (see Wohlwill, 1968 for the role of contiguity in children's inferences). That is, probabilities are assigned to cues, but must be used to infer outcomes of choices between options. Low contiguity is not inherent to probabilistic inference tasks. Recall the attempt to predict dog behavior from its body language, such as a wagging tail, perked-up ears, and a relaxed mouth. These cues are part the dog itself and thus a high contiguity between the probability assigned to those cues and future dog behavior exists. Other cues might also be helpful but more distal; for example, whether the dog is behind a fence or the behavior of its owner. In experimental settings, high contiguity can be established by the presentation format (options as cue-compounds, e.g., von Helversen, Mata, & Olsen, 2010) and by the experimental procedure. For example, when probabilities are used as anchor points, it creates proximity between the probabilistic information and the variable to be judged, in this case an expected value (e.g., Schlottmann, 2001). As spatial and temporal contiguity facilitates associative learning and children's inferences about causal structures (Renaux, Riviere, Craddoc, & Miller, 2017; Bühner, 2005) it may likewise facilitate probability utilization in judgments and decision making. Challenging environments that are low in contiguity, on the other hand, require to mentally construct proximity between probability and the variable to be judged and may impede probability utilization.

The probabilistic inference environment is both high in complexity and low in contiguity and might therefore hinder probability utilization in judgments as well as in decisions. Future research must validate this assumption by demonstrating that varying complexity and contiguity affects probability utilization.

Over-Confidence and Invariant Judgments in 6-Year-Olds

The analysis of individual subjects revealed that the over-confidence observed in judgment by 6-year-olds at the group level was mainly caused by a subsample of children that constantly predicted the maximum number of favorable outcomes. This behavior has also been found in other probabilistic judgment tasks and is even quite common when children rate emotion or motivation (e.g., Kalish, 2010; Schlottmann, 2001; Chambers & Johnston, 2002). In order to make a judgment, information must be integrated to form a value representation which then needs to be mapped to a behavioral response; for example, to an available scale value (e.g., Birnbaum, 1978). Thus, invariant responses can be due to invariant value representations or to a failure in mapping. As children made outcome predictions and judged cue validities on the same scale (Study 3), but showed invariant responses only in the former case. Presumably, these 6-year-olds really represented maximum values invariantly; that is, they expected only positive outcomes. This suggests that they might have been unable to consider the whole sample space of outcomes (positive treasures & negative spiders). Realizing that different outcomes are possible is the first challenge for children when dealing with probabilistic environments (Bryant & Nunes, 2012; Piaget & Inhelder, 1975) and might be especially difficult when the response format highlights one possible outcome but requires consideration of both (we specifically asked to predict treasures and did not mention spiders). Thus, children that did show variant responses might at least acknowledge the possibility of different outcomes, even when they could not quantify their probability yet.

Another factor that might have contributed to the large proportion of invariant responses (21% over all studies) is that children were not trained to correctly use the scales. Consider, for example, Anderson & Schlottmann's procedure (1990; Appendix A, Schlottmann, 2001). Children were provided with anchors (p = 0; .5; 1) and trained to choose the corresponding scale values (minimum, middle, maximum) when judging happiness of a player. Such training may greatly reduce invariant responses, encourages consideration of all possible outcomes, and even boosts children's utilization of probability.

Conclusion

Until late elementary school age, children neglect probabilities in decisions when they are assigned to cues (Betsch & Lang, 2013; Betsch et al., 2014, 2016, 2018; Lang & Betsch, 2018a; Mata et al., 2011). This paper shows that children neglect probability likewise in judgments. Reasons for probability neglect are therefore not due to specific demands of the decision process, but may affect both judgments and decisions similarly. Children's judgments show that their expectations of outcomes are not informed by probability assigned to cues at the age of six, but that starts to emerge at the age of nine. In similar areas, probabilistic judgments are mastered by younger children and earlier than probabilistic decisions (cf. Schlottmann & Wilkening, 2012). When probability is assigned to cues, its utilization emerges later, and judgments and decisions are equally deficient until the age of nine.

References

Anderson, N. H. (1991). Probability development. In N. H. Anderson (Ed.), *Contributions to information integration theory: Vol. III. Developmental* (pp. 83–108). Hillsdale, NJ: Lawrence Erlbaum.

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

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*(4), 933–945. https://doi.org/10.2307/1131034

Betsch, T. & Haberstroh, S. (2005). Current research in routine decision making: Advances and prospects. In T. Betsch & S. Haberstroh (Eds.), *The Routines of Decision Making* (pp. 359–377). Mahwah, NJ: Lawrence Erlbaum.

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology, 115*, 108–125. https://doi.org/10.1016/j.jecp.2012.11.003

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica, 153*, 74–86. https://doi.org/10.1016/j.actpsy.2014.09.008

Betsch, T., Lehmann, A., Lindow, S., Lang, A., & Schoemann, M. (2016). Lost in Search: (Mal-)adaption to probabilistic decision environments in children and adults. *Developmental Psychology, 52*(2), 311–325. https://doi.org/10.1037/dev0000077

Birnbaum, M. H. (1978). Differences and ratios in psychological measurement. In N. J. Castellan & F. Restle, (Eds.), *Cognitive Theory, Vol. 3*. (pp. 33–74). Hillsdale, NJ: Lawrence Erlbaum.

Brand, M., Recknor, E. C., Grabenhorst, F., & Bechara, A. (2007). Decisions under ambiguity and decisions under risk: Correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *Journal of Clinical and Experimental Neuropsychology, 29*(1), 86–99. https://doi.org/10.1080/13803390500507196

Bröder, A. (2000). Assessing the empirical validity of the "Take-the-best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, 26(5), 1332–1346. https://doi.org/10.1037/0278-7393.26.5.1332

Bryant, P., & Nunes, T. (2012). *Children's understanding of probability. A literature review (full report).* London: Nuffield Foundation. Retrieved from https://www.nuffieldfoundation.org/sites/default/files/files/Nuffield_CuP_FULL_REPORTv_FINAL.pdf

Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago, IL: University of Chicago Press.

Bühner, M. (2005). Contiguity and covariation in human causal inference. *Learning & Behavior, 33*(2), 230–238. https://doi.org/10.3758/BF03196065

Cassotti, M., Aïte, A., Osmont, A., Houdé, O., & Borst, G. (2014). What we learned about the processes involved in the Iowa Gambling Task. *Frontiers in Psychology, 5*, 915. https://doi.org/10.3389/fpsyg.2014.00915

Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology, 27*(1), 27–36. https://doi.org/10.1093/jpepsy/27.1.27

Cummings, G. (2012). *Understanding the new statistics. Effect sizes, confidence intervals, and meta-analysis*. Hove, UK: Routledge.

Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. G. (2014). A meta-analysis on age differences in risky decision making: Adolescents versus children and adults. *Psychological Bulletin, 141*(1), 48–84. https://doi.org/10.1037/a0038088

Ebersbach, M. (2009). Achieving a new dimension: Children integrate three stimulus dimensions in volume estimations. *Developmental Psychology, 45*(3), 877–883. https://doi.org/10.1037/a0014616

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 186–203. https://doi.org/ 10.1037/0278-7393.34.1.186

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684-704. https://doi.org/10.1037/0033-295X.102.4.684

Girotto, V., & Gonzalez, M. (2008). Children's understanding of posterior probability. *Cognition, 106*(1), 325–344. https://doi.org/10.1016/j.cognition.2007.02.0

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13*(12), 517–523. https://doi.org/ 10.1016/j.tics.2009.09.004

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, NY: Springer

Kalish, C.W. (2010). How children use examples to make conditional predictions. *Cognition, 116*(1), 1-14. https://doi.org/10.1016/j.cognition.2010.03.008

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General, 135*(2), 162–183. https://doi.org/10.1037/0096-3445.135.2.162

Lang, A., & Betsch, T. (2018a). Children's neglect of probabilities in decision making with and without feedback. *Frontiers in Psychology, 9*, 191. https://doi.org/10.3389/fpsyg.2018.00191

Lang, A., & Betsch, T. (2018b). Children use but can overcome non-adaptive decision strategies. *Manuscript under review.*

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*(5), 397–413. https://doi.org/10.1002/bdm.453

Levin, I. P., Hart, S. S., Weller, J. A., & Harshman, L. A. (2007). Stability of choices in a risky decision making task: A 3-year longitudinal study with children and adults. *Journal of Behavioral Decision Making, 20*(3)*,* 241–252. https://doi.org/10.1002/bdm.552

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. https://doi.org/10.1080/03601277.2014.900263

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*(1), 46–55. https://doi.org/10.1037/h0031207

Lindow, S. (2014). *Entscheidungen bei Kindern. Wie entwickeln sich Entscheidungsstrategien?* [*Decisions in Children: How do decision strategies develop?*] Göttingen, Germany: Optimus.

Mata, R., von Helversen, B., & Rieskamp, J. (2011), When easy comes hard: The development of adaptive strategy selection. *Child Development, 82*, 687–700. https://doi.org/10.1111/j.1467-8624.2010.01535.x

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Piaget, J., & Inhelder, B. (1951). *The origin of the idea of chance in children*. London, UK: Routledge & Kegan Paul.

Renaux, C., Rivière, V., Craddock, P., & Miller, R. (2017). Role of spatial contiguity in sensory preconditioning with humans. *Behavioural Processes, 142*, 141–145. https://doi.org/10.1016/j.beproc.2017.07.005

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). Oxford: Wiley.

Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *Child Development, 72*(1), 103–122. https://doi.org/10.1111/1467-8624.00268

Schlottmann, A., & Tring, J. (2005). How children reason about gains and losses: Framing effects in judgment and choice. *Swiss Journal of Psychology, 64*(3), 153–171. https://doi.org/10.1024/1421-0185.64.3.153

Schlottmann, A., & Wilkening, F. (2012). Judgment and decision making in young children. In M. K. Dhami, A. Schlottmann & M. R. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development and evolution* (pp. 55–83). Cambridge: University Press.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

von Helversen, B., Mata, R., & Olsson, H. (2010). Do children profit from looking beyond looks? From similarity based to cue abstraction processes in multiple-cue judgment. *Developmental Psychology, 46*(1), 220 –229. https://doi.org/10.1037/a0016690

Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton, NJ: University Press.

Windschitl, P. D., & Wells, G. L. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied, 2*(4), 343–364. https://doi.org/10.1037//1076-898X.2.4.343

Wohlwill, J. F. (1968): From perception to inference: A dimension of cognitive development. In I.E. Sigel & F.H. Hooper (Eds.), *Logical thinking in children* (pp. 472–494). New York, NY: Holt, Rinehard & Winston.

General Discussion

In the following sections, I report results of meta-analyses of all studies included in this thesis to summarize and discuss the main findings. All analyses were calculated using the metafor package in R (Viechtbauer, 2010). I first overview effects of feedback and age on children's decisions, decision strategies and judgments, and then address alternative accounts for age or feedback effects. Preschoolers of Studies 1-3, and 6, and 1st graders of Studies 4 and 5 were summarized to younger children (aged 5-7), and compared with older children (aged 8-10) and adults. Note that children attended the first grade not longer than three months when the data was collected.

Feedback Benefits Decision Making

Articles 1-3 investigate the effect of feedback on decisions. Participants either experienced the decision's outcome or it was delayed until the end of the task. Feedback mutually affects performance in decisions and other cognitive tasks (Bangerts-Drowns, Kulik, Kulil, & Morgan, 1991; Fyfe & Rittle-Johnson, 2015; Hattie & Timperley, 2007; Jessup, Bishara, & Busemeyer, 2008; Karelaia & Hogarth, 1998; Kluger & deNisi, 1996; Lejarraga & Gonzalez, 2011; Narciss & Huth, 2004; Newell & Rakow, 2007; Shute, 2008). Initially, we assumed that providing feedback would decrease children's utilization of probabilistic information, because feedback causes cognitive overload, distracts children from probabilities, or invites exploration. But this assumption must be rejected: Providing feedback did not decrease utilization of probabilistic information in any age group or study.

In contrast, feedback benefitted choices in all age groups. Figure 1 shows the effect of selective compared to no feedback on decisions in line with the high valid cue (i.e., this equals normative decision making when probabilities are highly dispersed). Weighted effect sizes for each age group (black diamonds in Figure 1) indicate that feedback increased decisions in line with the high valid cue. Age was included as a moderator in the analysis, but all groups benefitted equally from feedback, $Q(1) = 0.057$, $p = .811$.

Feedback format moderated the effect of selective feedback instead, $Q(1) = 26.93$, $p < .001$. I distinguished between feedback that maintained a priori stated differences in probabilities (in Studies 1-3, 4) and feedback that further increased the dispersion over time, and strictly reinforced the Lexicographic decision strategy (LEX, Studies 4-5). The latter was more effective. Note, that this should be interpreted cautiously. Besides feedback format, other factors varied between studies, such as the variation of cue value patterns (i.e., Studies 4 and 5 entailed less variation in cue value patterns compared with Studies 1-3) or salience of the most valid cue (i.e., in Studies 4 and 5, both other cues were equal in validity, thus, the differing validity of the high valid cue was more salient). In addition, a direct comparison of feedback that maintained stated probabilities and one that further dispersed them, only yielded a small benefit of the latter (Study 4, see Article 3 for discussion).

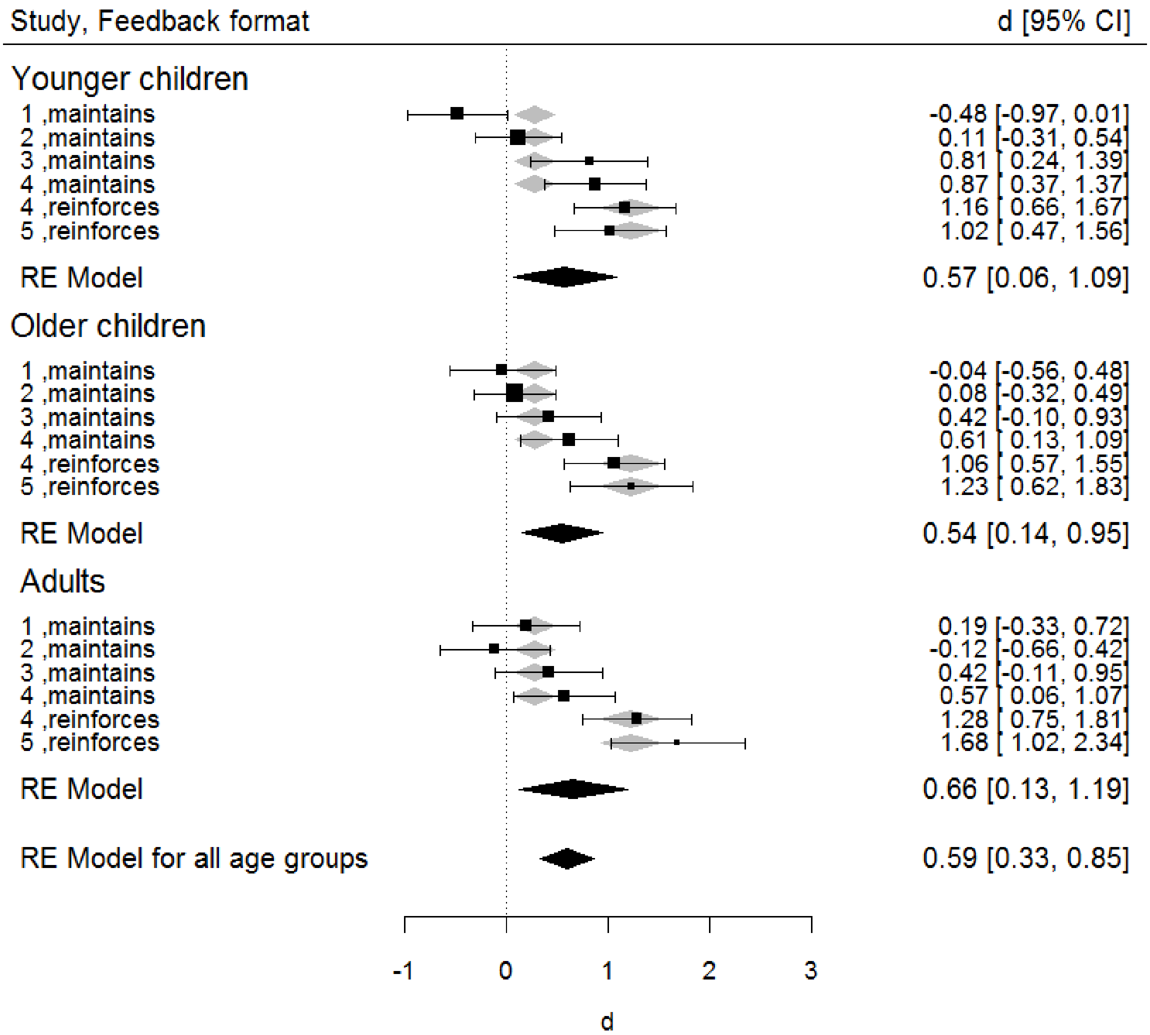


Figure 1. Forest plot for selective feedback effects on decisions. Figure shows Cohen's *d* for decisions with selective feedback compared with decisions without feedback. The dependent variable is decisions in line with the high valid cue. The figure shows weighted effect sizes for each age groups without including feedback format as a moderator (black diamonds) and with feedback format as a moderator (grey diamonds). It further shows the overall effect size for all age groups; maintains = maintains stated probabilities; reinforces = reinforces the Lexicographic strategy, RE Model = random effects model for the specified group.

In adults, the feedback effect was unexpected and unpredicted by decision theories. According to normative theory and is subjective variants, rational decisions makers should—regardless of feedback—maximize outcomes by calculating the weighed sum of cue values and choose the option with the highest expected value (Edwards, 1954; Glöckner & Betsch, 2008; Kahnemann & Tversky, 1979; Savage, 1954; von Neumann & Morgenstern, 1947).

From the perspective of adaptive decision making, the a priori established dispersion of probabilities suggested the use of a Lexicographic strategy (LEX, except in Study 5 where other strategies such as Equal Weight were also adaptive); adults should have applied this fast and frugal heuristic, and followed the predictions of the most informative cue. This mechanism does not rely on feedback (Gigerenzer & Goldstein, 1996).

Research shows that adults profit from feedback, when stated probabilities are difficult to process or when decision maker's cognitive abilities are constrained (Brand, Laier, Pawlikowski, & Markowitsch, 2009; Hogarth & Soyer, 2011). However, the stated probabilities established a quite simple cue hierarchy suggesting LEX as the most adaptive strategy. A closer look at adults' decisions with and without feedback reveals that adults especially deviated from LEX when both the other cues contradicted its prediction (e.g., Type 3 pattern in Studies 1-3) and no feedback was provided. This tendency was reduced with feedback. In line with prior research, this illustrates that adults neglect simple strategies and prefer more complex strategies such as weighted-additive under some circumstances (Bröder & Schiffer, 2003; Glöckner & Betsch, 2008). Importantly, this unexpected finding in adults highlights the significance of adult controls when evaluating children's decision making.

In children, both detrimental and beneficial effects of feedback on decisions are conceivable. Contrary to initial expectations, feedback did not cause deviations from normative or descriptive decision models. Instead, children neglected probabilities to a large degree in the presence and absence of feedback. Feedback moderately benefitted decisions in children, but the effect varied considerably between studies, leading to large heterogeneity in the meta–analysis. I^2-values indicate that the percentage of variability reflecting real differences between studies instead of sampling variability was large; without feedback format as moderator, $I^2_{\text{Younger}} = 84\%$; $I^2_{\text{Older}} = 74\%$; $I^2_{\text{Adults}} = 82\%$. The analysis of individual decision strategies provides further insight into how children profited from feedback.

Age and Feedback Effects on Decision Strategies

Analyzing decisions at the group level can disguise inter-individual differences in decision strategies (Bröder & Schiffer, 2003). In group analysis, this is reflected in large standard deviations when analyzing, for example, choices in line with LEX. Therefore, I used individual choice patterns to determine each individual's decision strategy (Bröder & Schiffer, 2003). The maximum likelihood approach allows to identify the strategy that most likely produced the choice pattern, assuming that the strategy application was flawed.

Importantly, I did not only search for decision strategies prominent in adult decision making but also for child-specific decisions strategies.

Article 1 shows first evidence for such child-specific and non-adaptive strategies. Children do not rely on probabilities but on irrelevant information, such as which cue was favored or which option has been chosen in the preceding decision. Article 2 shows that such non-adaptive strategies dominate at age 6, co-exist with adaptive strategies at age 9, but are rare in adults. Additionally, non-adaptive strategies are reduced when outcomes of decisions and counterfactual outcomes are displayed. Article 3 investigates whether, when and which feedback increases adaptive strategy use. Altogether, this investigation of decision strategies offers three new insights:

First, decision strategies are applied in all age groups, but the prevalence of strategies changes with age. One major contribution of our studies is showing consistently, that children do not make decisions without any systematic plan. Instead, they apply decision strategies that are non-adaptive to the probabilistic decision environment, and do not maximize outcomes. Preliminary evidence for children relying on normatively irrelevant information when making decisions has been reported previously (see Brainerd, 1981, for alternating options and van van Duijvenvoorde, Jansen, Bredman, & Huizenga, 2012, for suboptimal strategies in gambling tasks). But it should be recognized, that children systematically utilize this information to make decisions. In other areas of cognitive research, it is widely acknowledged that inferior strategies exist in children and affect the development of better ones (see Björklund, 2018, for an overview). Here, we show that the developmental trajectory from inferior to better strategies likewise applies to the domain of probabilistic decision making. Also in line with this research, non-adaptive and adaptive strategies co-exist in one age group resulting in large inter-individual differences in decision performance (Crowley, Shrager, & Siegler, 1997; Siegler, 2007).

Strategy use in preschool-aged children suggests that they are generally capable of rule-based decision making. When children do not use adaptive strategies, but can use non-adaptive ones, they are not merely failing in consistent strategy application. Other reasons for lack of adaptive strategies must be considered. The adaptive decision making approach would argue that either probabilistic strategies are not part of children's repertoire, or if they are, children fail in prioritizing them over others.

Second, as Article 3 demonstrates, the prevalence of adaptive decision strategies is increased by feedback. Table 3 reports proportion of adaptive strategies in each study and age group as well as estimations for proportions over all studies. It shows that this feedback effect varies considerably between studies. Feedback effects on performance depend on a variety of individual and environmental factors (Karelaia & Hogarth, 2008; Kluger & deNisi, 1998). In line with this, the observed feedback effects depended on the feedback format (selective vs. full, Study 3; strict vs. lenient, Study 4), and on the reinforced strategy (Lexicographic vs. Equal Weight, Study 5). Clearly, not all feedback is equally helpful.

In a follow-up study, we further tested whether adaptive strategy use can also be improved by instruction. We informed children that LEX would perform best in the task and explained in detail how the strategy worked (Lang & Betsch, 2018a). We found increased LEX-use in 9-year-olds, before the induction = 36%, 95% Bootstrap CI [23; 49]; after induction = 58%, CI [44; 72]. At the group level, decisions in line with LEX also increased, $t(58) = -5.33$, $p < .001$, 95% CI [0.35;1.10], $d = 0.71$. Accordingly, feedback learning and instruction affected strategy use, but only in a limited fashion.

Third, non-adaptive strategies are equally prevalent with and without feedback. Table 4 shows the proportions of non-adaptive strategies in each study and estimations across all studies. Non-adaptive strategies were only reduced when the decision's outcome and the counterfactual outcome were displayed (i.e., full feedback condition in Study 3). Since only one study tested the effect of full vs. selective feedback, results should be interpreted with caution.

This begs the question whether selective feedback is suited to reveal strategies as non-adaptive, at least in environments where experiencing one outcome does not allow to infer the other outcome (Fiedler, 2008). Selective feedback depends on the participant's choices: Like in many real-world situations, the decision's outcome is experienced, but the outcome of choice alternative remains unknown. However, based on assumptions of outcome distributions, this non-experienced outcome might be mentally constructed (Elwin, Juslin, Olssen, & Enkvist, 2007). For example, children might construct the non-experienced outcome as the expected value for the non-chosen option, which would be equal or less than the outcome of the chosen option (otherwise participants would choose the other option). Misbeliefs in strategy success could thus be easily maintained even when the strategy fails. Imagine, a child that believes that following the favored, but low valid, cue is an appropriate strategy. If the strategy produces a negative outcome, the child might assume that the choice alternative would have produced a negative outcome as well and keep up the strategy. This is no longer possible, when children experience both outcomes.

Clearly, so far we know very little about whether children engage in such inferences about non-experienced outcomes at all. We do, however, know that children compare chosen to non-chosen alternatives and devalue the latter when making choices (Egon, Santos, & Blum, 2007; McCormack, O'Connor, Beck, & Feeneya, 2016). Thus, a child might have some notions about non-experienced choice alternatives, which might affect their future decisions.

Table 1

Proportions of adaptive strategy users

Younger children						
Study	Control condition	%	95% CI	Feedback conditions	%	95% CI
1	No feedback	0	[0, 6]	selective, maintain	0	[0, 6]
2	No feedback	0	[0, 6]	selective, maintain	3	[0, 8]
3	No feedback	0	[0, 6]	selective, maintain	9	[0, 21]
				full, maintain	9	[0,22]
4	No feedback	10	[0, 20]	selective, maintain	35	[20, 50]
				selective, reinforce LEX	32	[18, 45]
5	No feedback	31	[15, 48]	selective, reinforce LEX	26	[10, 42]
				selective, reinforce EQW	60	[41, 78]
RE-Model		**6**	**[0, 15]**		**20**	**[7, 34]**
I²		80%			91%	

Older children						
Study	Control condition	%	95% CI	Feedback conditions	%	95% CI
1	No feedback	12	[0, 24]	selective, maintain	15	[1, 29]
2	No feedback	16	[4, 29]	selective, maintain	14	[2, 26]
3	No feedback	31	[14, 48]	selective, maintain	40	[23, 57]
				full, maintain	50	[31, 69]
4	No feedback	11	[1, 22]	selective, maintain	47	[30, 64]
				selective, reinforce LEX	62	[46, 78]
5	No feedback	35	[16, 53]	selective, reinforce LEX	54	[34, 74]
				selective, reinforce EQW	49	[45, 64]
RE-Model		**20**	**[10, 27]**		**41**	**[28, 54]**
I²		53%			80%	

Adults						
Study	Control condition	%	95% CI	Feedback conditions	%	95% CI
1	No feedback	61	[43, 79]	selective, maintain	59	[41, 77]
2	No feedback	40	[20, 57]	selective, maintain	57	[36, 77]
3	No feedback	71	[63, 89]	selective, maintain	82	[68, 96]
				full, maintain	59	[40, 77]
4	No feedback	53	[36, 71]	selective, maintain	74	[59, 90]
				selective, reinforce LEX	100	[92, 100]
5	No feedback	86	[72, 100]	selective, reinforce LEX	96	[88, 100]
				selective, reinforce EQW	88	[74, 100]
RE-Model		**63**	**[47, 79]**		**79**	**[63, 87]**
I²		77%			30%	

Note. Table shows proportions of adaptive strategy users in conditions without feedback on the left and in different feedback conditions on the right; maintain = feedback maintains stated probabilities, reinforce LEX, EQW = feedback reinforces the Lexicographic or Equal Weight strategy; RE = random effect model; studies with 0% and 100% are excluded from the model; *I²* measures heterogeneity between studies, lower values indicate less heterogeneity.

Table 2

Proportions of non-adaptive strategy users

Study	Control condition	%	95% CI	Feedback conditions	%	95% CI
Younger children						
1	No feedback	11	[0, 23]	selective, maintain	21	[6, 35]
2	No feedback	56	39, 73]	selective, maintain	41	[25, 56]
3	No feedback	60	[41, 78]	selective, maintain	57	[37, 76]
				full, maintain	18	[2, 36]
4	No feedback	58	[41, 76]	selective, maintain	43	[27, 59]
				selective, reinforce LEX	42	[27, 56]
5	No feedback	38	[38, 55]	selective, reinforce LEX	48	[30, 66]
				selective, reinforce EQW	33	[16, 51]
RE-Model		**44**	**[25, 63]**		**37**	**[28, 46]**
I²		86%			57%	

Study	Control condition	%	95% CI	Feedback conditions	%	95% CI
Older children						
1	No feedback	15	[2, 28]	selective, maintain	19	[4, 34]
2	No feedback	39	[22, 55]	selective, maintain	21	[6, 36]
3	No feedback	28	[11, 444]	selective, maintain	7	[0, 16]
				full, maintain	14	[1, 27]
4	No feedback	39	[23, 55]	selective, maintain	41	[25, 57]
				selective, reinforce LEX	24	[10,39]
5	No feedback	54	[35, 73]	selective, reinforce LEX	42	[22, 62]
				selective, reinforce EQW	41	[21, 61]
RE-Model		**34**	**[21, 48]**		**25**	**[15, 34]**
I²		69%			71%	

Study	Control condition	%	95% CI	Feedback conditions	%	95% CI
Adults						
1	No feedback	0	[0, 7]	selective, maintain	7	[0, 17]
2	No feedback	0	[0, 7]	selective, maintain	4	[0, 13]
3	No feedback	0	[0, 6]	selective, maintain	0	[0,6]
				full, maintain	14	[1, 28]
4	No feedback	16	[3, 28]	selective, maintain	13	[1, 25]
				selective, reinforce LEX	0	[0, 0]
5	No feedback	14	[0, 28]	selective, reinforce LEX	4	[0, 12]
				selective, reinforce EQW	8	[0, 20]
RE-Model		**15**	**[5, 24]**		**3**	**[0, 8]**
I²		15%			0%	

Note. Table shows proportions of non-adaptive strategy users; maintain = feedback maintains stated probabilities, reinforce LEX, EQW = feedback reinforces the Lexicographic or Equal Weight strategy; RE = random effect model; studies with 0% and 100% are excluded from the model; *I²* measures heterogeneity between studies, lower values indicate less heterogeneity.

Methodological Limits of Strategy Assessment

An outcome-based method allows to identify strategies non-verbally and undiluted by sequential search processes (see Betsch, Funke, & Plessner, 2011, for a discussion). Despite its benefits, it also entails some downsides of particular relevance with child participants: First, it requires a critical number of decisions for which predictions of investigated strategies differ (e.g., Bröder, 2002 for this separation problem). Still, child participants can only cope with a limited number of decisions even when breaks are included (e.g., Mata, van Helversen, & Rieskamp, 2011 for effects of fatigue). Here, the number of decisions was much lower compared with adult studies, which resulted in less reliable differentiation of strategy models.

Moreover, only strategies that were defined a priori and applied consistently over a defined period of decisions (e.g., in the last block) could be identified. It is very likely that some children used other strategies or switched between two or more strategies (see Scheibehenne, Rieskamp, & Wagenmakers, 2013, and Siegler, 2007, for application of multiple strategies in adults and children). Presumably, children's strategic decision making is thus still underestimated.

Further, outcome-based strategy assessment does not differentiate between deliberate and intuitive strategy application. However, this differentiation is crucial for dual-process frameworks, which assume very different developmental trajectories for intuitive and deliberate decision making (Jacobs & Klaczynski, 2002; Klaczynski, 2005; Schlottmann & Wilkening, 2012; Stanovich, Toplak, & West, 2010; cf. Reyna & Brainerd, 2011).

When strategies are applied in a deliberate fashion, children should be able to explicate their strategies (Evans & Stanovich, 2013; Walsh & Gluck, 2016). In a follow-up study (replication of Study 3's feedback condition, Lang & Betsch, 2018b), we asked 20 9-year-olds which decision strategies they used. Only one child indicated deciding "with feeling", all others clearly described one of the considered decision strategies: 35% indicated they followed the smartest animal (= LEX), 40% indicated they followed the smartest animal but not if the other animals contradicted it (= weighted-additive with uncorrected cue validities), 20% indicated that they switched from one option to the other (= option alternation). Further research must show, if this matches actual decision behavior and also applies to younger children, but preliminarily, it points to rather deliberate strategy use. Interestingly, this is in line with children's decisions in non-probabilistic environments, where most children could at least partially report strategies (Lindow, Lang & Betsch, 2017).

Age and Feedback Effects on Judgments

The fourth article exclusively investigates children's utilization of probabilities in judgments. Previous developmental research suggests that probability utilization in judgments is easier and therefore emerges earlier than in decisions (Schlottmann & Wilkening, 2012). However, the same developmental pattern—probability neglect at age six and emerging utilization at age nine—characterizes judgments in our studies.

In each study, we measured whether children utilized differences in cue validities when predicting outcomes for a decision maker following the first cue, the second, the third, respectively. This is an analogical adaption of previous research where preschool-aged children were able to predict successes for players of gambles with differing winning probabilities (Schlottmann, 2001). One of this analogy's shortcoming, however, might be that the game's complexity affects probability utilization (Wilkening & Anderson, 1991). Specifically, children might be able to anticipate playing a gamble with a specified probability but not following a cue with a specified validity.

In a follow-up study (Lang & Betsch, 2018b), we showed decisions in line with each cue to 9-year-olds before they predicted outcomes. For example, children observed the first decision makers choosing in line with the first cue several times, and then predicted outcomes for these choices. Previous results were replicated. 9-year-olds were sensitive to cue validities; their judgments were correlated with normative predictions, indicating that they expected better outcomes for decisions with more valid cues, $r_{\text{6-treasure-scale}} = .50$, 95% CI [.27, .69], $r_{\text{10-treasure-scale}} = .33$, CI [.05, .58]. But in individual analysis, only a minority (17%) used cue validities consistently in all judgments.

A second shortcoming might be that the stimulus field for judgments was more complex: For example, children were confronted with two options instead of one. This reflects that probability utilization per se is more complex in some situations than in others (see Article 4 for a discussion). But as Schlottmann & Wilkening (2012) point out, reducing the stimulus field so that probability is the only information left might be the most effective way to ensure its utilization in children.

Article 4 further shows that feedback did not improve children's judgments. Even after plenty of experience with the Treasure Hunt and with observing cue-outcome contingencies in feedback conditions, children did not expect better outcomes for decisions in line with more valid cues. Interestingly, predictive judgments were neither improved in Studies 4 and 5, despite the adaption of decision behavior. Consider for example, that in Study 4's strict feedback condition, 27% of younger, 60% of older children and 85% of adults consistently expected better outcomes for the high valid cue before decisions, in comparison with 24% of younger children, and 51% of older children, and 97% of adults after decisions.

In line with this, in complex tasks adults also sometimes adapt their decision behavior to the probabilistic structure, but not their judgments (Evans, Clibbens, Cattani, Gluck, Shohamy, & Myers, 2002; Franco-Watkins, Derks, & Dougherty, 2003; Harries, Evans, & Dennis, 2000, cf. Lagnado, Newell, Kahan & Shanks, 2006). This supports dual-process theories' claim that decision are guided by implicit and judgments by analytical system (e.g., Sloman, 1996).

Table 3

Proportion of normative judgments

Younger children						
	No feedback		**Feedback**		**Overall**	
Study	*%*	95% CI	*%*	95% CI	*%*	95% CI
1	4	[4, 11]	7	[0, 17]	5	[0,12]
2	0	[0, 6]	3	[0, 8]	1	[0, 4]
3	4	[0, 11]	4	[0, 10]	4	[0, 9]
4	19	[5, 34]	23	[15, 31]	23	[15, 31]
5	19	[5, 33]	37	[19, 55]	26	[17, 35]
6	—		—		0	[0,7]
RE-Model					**11**	**[15, 21]**
*I*²					94%	

Older children						
	No feedback		**Feedback**		**Overall**	
Study	*%*	95% CI	*%*	95% CI	*%*	95% CI
1	39	[19, 59]	19	[4, 33]	28	[16, 40]
2	29	[13, 45]	28	[11, 44]	28	[17, 40]
3	31	[14, 48]	33	[20, 45]	32	[22, 42]
4	50	[33, 67]	54	[45, 63]	54	[45, 63]
5	42	[23, 62]	50	[29, 71]	58	[48, 68]
6	—		—		34	[26, 43]
RE-Model					**39**	**[29, 50]**
*I*²					85%	

Adults						
	No feedback		**Feedback**		**Overall**	
Study	*%*	95% CI	*%*	95% CI	*%*	95% CI
1	77	[64, 94]	85	[72, 99]	82	[71, 92]
2	85	[71, 98]	87	[74, 100]	86	[76, 96]
3	82	[67, 97]	83	[73, 92]	82	[75, 90]
4	94	[86, 100]	96	[92, 100]	96	[92, 100]
5	86	[72, 100]	92	[82, 100]	76	[66, 86]
6	—		—		—	
RE-Model					**85**	**[78, 92]**
*I*²					62%	

Note. Table shows proportions of individuals whose utilization of cue validities in judgments were in line with normative expectations, that is, they expected better outcomes for decisions in line with more valid cues. Criteria differed between studies due to variation in stated cue validities; Studies 1-3, 6: $j_{cue1} > j_{cue2} > j_{cue3}$ & $j_{6\text{-treasure scale}} > j_{10\text{-treasure scale}}$; Studies 4, 5: $j_{cue1} > j_{cue3}$ & $j_{cue2} > j_{cue3}$ & $j_{6\text{-treasure scale}} > j_{10\text{-treasure scale}}$; RE = random effect model; atudies with 0% and 100% are excluded from the model; I^2 measures heterogeneity between studies, lower values indicate less heterogeneity.

The Relation Between Probability Utilization in Judgments and Decisions

Assessment of both utilization of probability in judgments and decisions allows to investigate whether normative integration of probability in judgments is associated with better decision making. For this analysis, I concentrated on older children because only this age group showed considerable variance in judgments and decisions: Approximately one third to one half of the children's judgments aligned to normative expectations and likewise around the same proportion used adaptive decision strategies.

We compared the proportion of adaptive decisions (i.e., in line with the high valid cue; we excluded one condition of Study 5, where other strategies were also successful) between children whose judgments followed normative standards (i.e., they expected better outcomes for decisions in line with more valid cues) and those whose judgments did not. Across all studies, children that used cue validities normatively for judgments, made slightly better decision than their peers (Figure 2). However, the overall effect was small, varied considerably between studies, $Q(5) = 19.15$, $p < .001$, $I^2 = 75\%$, and did not reach significance in single studies.

This is in line with scarce empirical findings showing that the association between judgment and decision may be rather lose in children (Schlottmann & Tring, 2005). However, future research with larger samples of older children should illuminate this issue further.

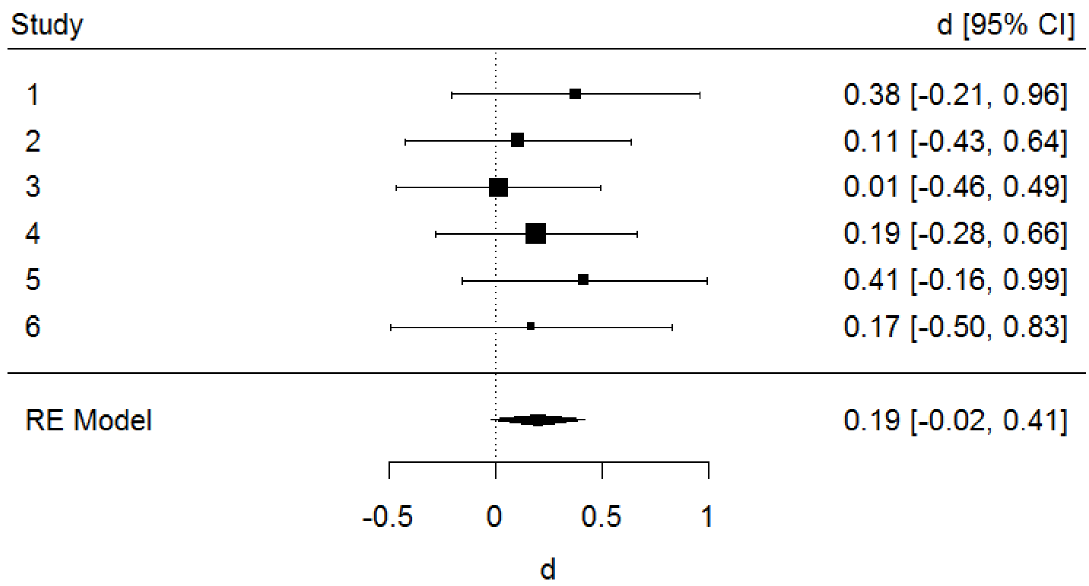


Figure 2. Forest plot for the effect of normative judgments on adaptive decisions in older children. Figure shows Cohen's *d* for 9-year-olds without normative judgments compared to 9-year-olds with normative judgments.

Stability of Age Effects in Judgments and Decisions

Overall, we found stable age effects in decisions and in judgments: no probability utilization in younger children and partial probability utilization in in a subgroup of older children. This pattern of results has been replicated in other studies that used the same paradigm in German samples (Betsch, Lang, Lehmann, & Axmann, 2014; Betsch, Lehmann, Lindow, Lang & Schoemann, 2016; Betsch, Lehmann, Jekel, Lindow, & Glöckner, 2018) as well as in a Japanese sample of 9-year-olds. Direct comparison between German and Japanese 3rd graders revealed no differences in decisions or judgments (Lang, Eisen, & Betsch, 2018c).

Probability utilization nevertheless varies with environmental factors. For example, in Studies 4 and 5 where the criterion (for judgments that were defined as accounting for probability differences) was less restrictive, probability utilization in judgments obviously improved (Table 3, see annotation). It is unclear though, whether increased salience of validity differences might have further facilitated probability utilization. Further, children's adaptive strategy use in decision increases considerably, when the non-probabilistic Equal Weight strategy was included as an adaptive strategy (Table 1, Study 5). Finally, manipulating environmental factors, such as demands of information search (Studies 1 & 2), dispersion of probability (Studies 4 & 5), and number of trials (Studies 2 & 3) clearly affected children's decisions and decision strategies. It did not, however, substantially increase probability utilization on any measure.

Alternative Accounts for Age-Dependent Differences

Here I discuss possible explanations for the observed performance differences between age groups or feedback conditions inherent to the research paradigm.

Information Encoding

For a meaningful comparison of different age groups, all children must be able to encode the relevant information, such as cue values. However, in information-board studies it is conceivable that children, unfamiliar with tables, might not be able to correctly encode the information in the matrix, and simply disregard it. We addressed this by using an information-board matrix that had only pictorial values and was not overly complex. Children at preschool age can handle information boards with up to four dimensions and three options (Lindow et al., 2017). Further, in each study, we assessed whether children understood the matrix.

Decoding the information-board. Before the first decision, the experimenter explained the matrix and the child indicated the meaning of each pictorial cue value in at least one training trial verbally and by pointing. Table 1 shows that only a negligible proportion of children could do this correctly.

Table 4

Percentage of children that could not interpret the information-board matrix correctly

Study	Age Group	
	Younger	Older
1	5	4
2	3	0
3	3	1
4	8	0
5	7	0
6	7	3

Encoding of differences in cue validities. A main benefit of the Treasure Hunt is the child-friendly presentation of probabilistic relations. Using a guided sampling procedure and graphical representations, we ensured that all children encoded probabilities correctly. Accordingly, children of all age groups were able to choose the most valid cue based on differences in cue validity (see low drop-out rates in all studies).

Note, that cue validity was framed as smartness of animals for children (i.e., we asked, *Which animal was the smartest?*; *Which was the second smartest?; How smart was this animal?*). Conceivably, children might have encoded and recalled only differences in smartness, without relating them specifically to differences in validity. However, children assigned smart points for correct predictions themselves and every single child correctly indicated when to assign a smart point. Thus, I think it unlikely that children did not grasp that animals' smartness in the game depended on correctly predicting outcomes. It is, however, possible that they forgot this during the game, and later only recalled differences in smartness rather than in prediction accuracy.

To explore this issue further, we coded children's answers when asked after the game why the last (i.e., high valid) cue was "the smartest". Most children referred to the number of smart points, younger: RE-Model, 73%, 95% CI [67, 79], I^2 = 37%; older: RE-Model, 63%, [56, 70], I^2 = 57%; adults: RE-Model, 65%, [56, 73], I^2 = 61%. However, a small proportion also spontaneously referred to the number of correct cue predictions, younger: RE-Model, 10%, [6,15], I^2 = 47%; older: RE-Model, 19%, [15, 23], I^2 = 0%; adults: RE-Model, 28%, [24, 33], I^2 = 0%.

In a follow-up study with older children (Lang & Betsch, 2018b), we specifically tested whether children were able to reproduce differences in correct predictions for each cue. Children indicated how often each cue's predictions were correct on the familiar graphic-stick-scale (see Article 4 for display of the scale). Mean estimates were correlated with normative values, Pearson's $r = .66$, 95% Bootstrap CI [64; 82], and 72% of children, [50; 89] correctly reproduced the cue hierarchy. I conclude, that most children remembered until after the game, which cue's predictions were more valid. Yet, this is to be replicated with younger children.

Altogether, I think, that age-dependent differences in encoding or recalling of differences in probabilities may contribute to, but cannot fully explain lack of utilization. Every age group successfully encoded and to a large extend recalled differences in cue validities in various manipulation checks.

Children's motivation

Another factor contributing to age or feedback effects might be children's motivation. To ensure, that all children were highly motivated to maximize outcomes, we repeatedly informed participants, that they would be rewarded contingent on their performance in decisions. For each treasure found, children received a treasure point and could trade them for prizes afterwards. Although, incentives might affect only older children's performance (Varghese & Nielsen, 2013), we rewarded all children in the same way. Additionally, we assessed extrinsic and intrinsic motivation.

Measuring children's motivation is tricky, especially with a simple and brief method (the studies took quite a long time even without further measurements). Children often indicate indifferently that they are highly motivated (see Chambers & Johnson, 2002, for an overview). Nevertheless, we attempted to measure children's motivation during the game by adapting the short version of the Intrinsic Motivation Inventory, which has been used in a German sample of children (Deci & Ryan, 2013; Wilde, Bätz, Kovaleva, & Urhahn, 2009). We assessed the factors enjoyment and perceived competence, each with three items. Children indicated their answer on the familiar Graphic-stick scale.

Cronbach's alpha indicated poor internal consistency especially in the younger age group for both factors and in all studies, $\alpha s \leq .74$. Presumably, some of the items were not meaningful to younger children (e.g., *Wie unterhaltsam war das Spiel?*, How entertaining was the game?). Thus, we report only results for the linguistically simplest item of each factor (*Wie viel Spaß hat dir das Spiel gemacht?* How much fun was the game?; *Wie gut warst du in dem Spiel?,* How good were you at playing the game?).

Further, we assessed how much children were looking forward to the prizes provided for performance as a proxy for extrinsic motivation (*Wie sehr freust du dich mit deinen Schatzpunkten Preise zu kaufen?* How much do you look forward to buying prizes with your treasure points?).

Motivation with and without feedback. Overall, children indicated to be highly motivated in all conditions. Importantly, whether decision feedback was available or not did not affect children's reported enjoyment, RE-Model, $d = -0.12$, CI [-0.43, 0.19], $I^2 = 74\%$; perceived competence, RE-Model, $d = 0.14$, CI [-0.09, 0.37], $I^2 = 74\%$; or extrinsic motivation, $d = 0.08$, CI [-0.09, 0.25], $I^2 = 1\%$. This suggests that feedback did not affect children's motivation.

Age-dependent differences in motivation. Additionally, we tested for an age effect on enjoyment, perceived competence and extrinsic motivation in children. Age only affected children's perceived competence: Younger children felt more competent than older children, despite their worse performance, RE-Model, $d = 0.77$, 95% CI [0.61, 0.93], $I^2 = 16\%$.

However, a significant proportion of children indifferently selected maximum scale values. This may indeed reflect high motivation but can also be a result of demand effects or inappropriate scale use in children and reduce the validity of motivation assessment (see Chambers & Johnston, 2002, and Article 4 for a discussion).

Irrespective of this restriction, the data suggests that children were highly motivated, intrinsically and extrinsically. Thus, I am confident, that all children tried to perform well in the decision game, and doubt that children's failure in probability utilization or strategy application can be explained by a lack of motivation.

Further evidence for this claim stems from an additional study with 9-year-old children: In this study (Lang & Betsch, 2018d), children first made decisions and gained treasure points for each correct decisions. In a second part they lost treasure points for false decisions. While children were highly motivated to avoid losing treasure points, their decisions resembled our findings here: Only a proportion of children relied on adaptive decision strategies and minimized losses.

The Role of Executive Functions

Executive functions, such as working memory, inhibitory control, and cognitive flexibility develop considerably during preschool and elementary school age and are an important predictor for decision making in developmental studies (Diamond, 2006; Miyake et al., 2006; Weller, Levine, Rose, & Bossard, 2012; see Schiebener, Zamarian, Delazer, & Brand, 2014, and Steinbeis & Crone, 2016, for adolescents and adults). However, one study using the Treasure Hunt with 9-year-olds showed no association between performance in multiple executive function tasks and probability utilization (Lehmann & Betsch, 2018). Thus, it stands to reason that all children at age nine have sufficient executive functioning to succeed in the game and that other factors must be responsible for inter-individual differences.

Conclusions

In a series of studies, I investigated children's judgment and decision making in a probabilistic inference paradigm. Focusing on two major points, the work in this thesis shows the following:

The Development of Probability Utilization

Children at the ages of 5-7 and 8-9 encode differences in probabilistic relations between cues and outcomes successfully. They can further utilize this probabilistic information to choose the most valid of several cues. But most children until the age of nine fail to utilize probabilistic relations to form informed expectations of outcomes and to choose options that maximize outcomes. This probability neglect is very robust, decreases considerably from age five to nine, and follows the same developmental trajectory in decisions and judgments. This finding contrasts research showing that children at preschool age and even infants utilize probability in various contexts (see Schlottmann & Wilkening, 2012, for an overview).

From a developmental perspective, dual-process frameworks account for heterogeneity by assuming that intuitive probability utilization is available early in life, while analytical utilization develops later. When environments are "merciful", they allow children's intuitive capabilities to flourish (Schlottmann & Wilkening, 2012, p. 77). So why might the environment created in our studies not be "merciful"? Presumably, two environmental factors—high complexity of the stimulus field and low contiguity between probability, which were assigned to cues, and options' outcomes, which should be predicted, might require an analytical approach for successful probability utilization (Wohlwill, 1968; see Betsch et al., 2018, and Article 3 for discussions). Preliminary evidence suggest that the older children that successfully utilized probability, applied an analytical processing mode (Article 4 and Lang & Betsch, 2018b). Further research must show, whether variations of these environmental factors affects probability utilization as expected.

From a decision making perspective, the results support models of adaptive decision making (Gigerenzer, Todd, & the ABC Research Group, 1999; Payne, Bettman & Johnson, 1988), which assume a repertoire of decision strategies. They add a developmental component by showing the trajectory from non-adaptive to adaptive strategies during elementary school age. However, the results also contradict the assumption that decision makers tend to reduce complexity (Gigerenzer et al., 1999, Simon, 1954). Though the dispersion of stated probabilities invited the use of complexity-reducing strategies such as the Lexicographic strategy, a substantial proportion of individuals did not do this. Adults and in part older children tended to integrate more information than necessary in a more complex weighted-additive fashion (Glöckner & Betsch, 2008).

The Complex Role of Decision Feedback

Feedback effects are small. Decision feedback benefits children's probabilistic decision under very specific circumstances and to a limited extend. When the outcome distribution vehemently reinforces simple but adaptive decision strategies, their prevalence is increased, but less adaptive strategies are not eradicated. Overall, three separate feedback effects were found in children:

First, we obtained the over-responsiveness to options. Younger children only reluctantly chose options again after failure. Further research must show whether this also is the case at the level of decision strategies (that is, whether children only reluctantly apply a decision strategy after failure). Second, we obtained strategy switches in relation to feedback. Children increasingly applied decision strategies that yielded better outcomes with feedback. Third, we obtained preliminary evidence of changes in perceived cue validity in dependence on feedback. When the feedback schedule decreased a cue validity (that is, the cue predictions became less accurate over time), children's perception of cue validities changed accordingly.

Children seem to attribute feedback to different elements of the decisions—options, strategies and cues. Large inter-individual differences suggest that these attributions may be volatile and moderated by attentional processes (Bröder, Glöckner, Betsch, Link, & Ettlin, 2013; Roelfsema, van Oooyen, & Watanabe, 2010). A model of feedback learning in complex decision, such as probabilistic inferences, should aim to separate them.

From a decision making perspective, the results support the assumption that individuals adapt strategies through decision feedback (Beach and Mitchel, 1978; Gigerenzer et al., 1999; Payne, et al., 1988; Rieskamp & Otto, 2006). But they illustrate likewise that feedback processing is more complex even at preschool age. Besides updating of strategies, other process, such as updating of expectations for options and cues must be accounted for (see Bröder et al., 2013, for a discussion).

From a developmental perspective, the research confirms sensitivity towards feedback at every age, but also developmental differences in feedback processing, such as over-responsiveness to negative outcomes (van Duijvenvoorde et al., 2012). Feedback is suited to improve children's decision making, but clearly, not all feedback is helpful for children. Further research must clarify what feedback can benefit decisions.

As always, more research is needed to replicate and extend the findings and tie up all the remaining loose ends. Still, the majority of parents and educators that supported our research over the years, asked for advice on children's decisions. Of course, evidence-based advice must be given cautiously and cannot be informed by a single study. However, I think that some preliminary pieces of advice might be justified and helpful:

First, children should make decisions. Even at preschool age children can handle decisions in a very systematic fashion. Their strategies are not always suited to the decision at hand, but that does not hinder the process of decision making—however, the outcomes might not be optimal. Strategies get better with age and can also be improved from outside, for example by feedback or instruction. Ideally, adults should let children make their own decisions (without too high expectations for results), but help the child to analyze decision situations, possible strategies, and outcomes.

Second, searching for information is hard. Children's decision quality is severely impaired when they have to search for the choice-relevant information, not only in probabilistic but also in deterministic environments (Betsch et al., 2016; Lindow & Betsch, 2018). In contrast, when relevant information is available beforehand, decisions get better in older children. Accordingly, relevant information should always be made available for children. Parents or educators should actively collect relevant information for or with the child. Further research on children's active information search suggests that encouraging to ask questions, might be a good method to engage children in information search (Legare, Mills, Souza, Plummer, & Yasskin, 2003).

Third, children do not necessarily use important information in decisions. Naturally, children might consider different information to be important for decisions, than adults. But knowing something and using it in decisions are two different things. Adults should explicitly state when a piece of information is relevant for a decision. For example, explain how it might affect decision outcomes.

This research on children's probabilistic judgment and decision making has revealed several deficits and some capabilities of young children. When evaluating this, we should keep in mind, that even many adults' probabilistic judgments and decisions are deficient. Similar to adults, acknowledging children's difficulties creates opportunities to create decision situations that children can handle successfully.

References

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61,* 213–238. https://doi.org/10.3102/00346543061002213

Beach, L. R. & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review, 3*, 439–449. https://doi.org/10.5465/amr.1978.4305717

Betsch, T., Funke, J., & Plessner, H. (2011). *Denken - Urteilen, Entscheiden, Problemlösen*. Berlin: Springer.

Betsch, T., Lang, A., Lehmann, A., & Axmann, J.M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica, 153,* 74–86. https://doi.org/10.1016/j.actpsy.2014.09.008

Betsch, T., Lehmann, A., Lindow, S., Lang, A., & Schoemann, M. (2016). Lost in search: (Mal-)adaptation to probabilistic decision environments in children and adults. *Developmental Psychology,* 52(2*),* 311–325. https://doi.org/10.1037/dev0000077

Betsch, T. Lehmann, A., Jekel, M., Lindow, S., & Glöckner, A. (2018). Children's application of decision strategies in a compensatory environment. *Paper under review.*

Björklund, D. F., & Causey, K. B. (2018). *Children's thinking: Cognitive development and individual differences.* Thousand Oaks, CA: Sage Publications.

Brainerd, C. J. (1981). Working memory and the developmental analysis of probability judgment. *Psychological Review, 88,* 463–502. https://doi.org/10.1037/0033-295X.88.6.463

Brand, M., Laier, C., Pawlikowski, M., & Markowitsch, H. J. (2009). Decision making with and without feedback: The role of intelligence, strategies, executive functions, and cognitive styles. *Journal of Clinical and Experimental Neuropsychology, 31(8)*, 984–998. https://doi.org/10.1080/13803390902776860

Bröder, A., Glöckner, A., Betsch, T., Link, D., & Ettlin, F. (2013). Do people learn option or strategy routines in multi-attribute decisions? The answer depends on subtle factors. *Acta Psychologica, 143*, 200–209. https://doi.org/10.1016/j.actpsy.2013.03.005

Bröder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making*, *16*(3), 193–213. https://doi.org/10.1002/bdm.442

Byrnes, J. P., Miller, D. C., & Reynolds, M. (1999), Learning to make good decisions: A self-regulation perspective. *Child Development, 70*, 1121–1140. https://doi.org/10.1111/1467-8624.00082

Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology, 27*(1), 27–36. https://doi.org/10.1093/jpepsy/27.1.27

Crowley, K. & Shrager, J. & Siegler, R. (1997). Strategy discovery as a competitive negotiation between metacognitive and associative Mechanisms. *Developmental Review*, *17*; 462–489. https://doi.org/10.1006/drev.1997.0442.

Deci, R. M., & Ryan, E. L. (2013). *Intrinsic motivation inventory*. Retrieved from http://selfdeterminationtheory.org/questionnaires/

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. http://doi.org/10.1146/annurev-psych-113011-143750

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*, 380–417. https://doi.org/10.1037/h0053870

Egan, L. C., Santos, L. R., & Bloom, P. (2007). The origins of cognitive dissonance: Evidence from children and monkeys. *Psychological Science, 18,* 978–983. https://doi.org/10.1111/j.1467-9280.2007.02012.x.

Elwin, E., Juslin, P., Olssen, H., & Enkvist, T. (2007). Constructivist coding: Learning from selective feedback. *Psychological Science, 18(2),* 105–110. https://doi.org/10.1111/j.1467-9280.2007.01856.x

Evans, J. S. B. T., Clibbens, J., Cattani, A., Harris, A., & Dennis, I. (2003). Explicit and implicit processes in multicue judgment. *Memory & Cognition, 31,* 608–618. https://doi.org/10.3758/BF03196101

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223–241. https://doi.org/10.1177/1745691612460685

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 34*(1), 186–203. https://doi.org/10.1037/0278-7393.34.1.186

Franco-Watkins, A., Derks, P., & Dougherty, M. (2003). Reasoning in the Monty Hall problem: Examining choice behaviour and probability judgements. *Thinking & Reasoning, 9*(1), 67–90. https://doi.org/10.1080/13546780244000114

Fyfe, E. R., & Rittle-Johnson, B. (2016). Feedback both helps and hinders learning: The causal role of prior knowledge. *Journal of Educational Psychology,108*(1)*,* 82–97. https://doi.org/10.1037/edu0000053

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650–669. https://doi.org/10.1037/0033-295X.103.4.650

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

Glöckner, A. & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5)*,* 1055–1075. https://doi.org/1055-1075

Gluck, M. A., Shohamy, D., & Myers, C. E. (2002). How do people solve the "weather prediction" task? Individual variability in strategies for probabilistic category learning. *Learning & Memory*, *9***,** 408–418. https://doi.org/10.1101/lm.45202

Hattie, J., & Timperley, H (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hogarth, R. M., & Soyer, E. (2011). Sequentially simulated outcomes: Kind experience vs. non-transparent description. *Journal of Experimental Psychology: General, 140*(3), 434–463. https://doi.org/10.1037/a0023265

Jacobs, J. E., & Klaczynski, P. A. (2002). The development of judgment and decision making during childhood and adolescence. *Current Directions in Psychological Science, 11*, 145–149. https://doi.org/10.1111/1467-8721.00188

Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychological Science, 19*, 1015–1021. doi:10.1111/j.1467-9280.2008.02193.x.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291. https://doi.org/10.2307/1914185.

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, *134*(3), 404–426. https://doi.org/10.1037/0033-2909.134.3.404

Klaczynski, P. A. (2005). Metacognition and cognitive variability: A two-process model of decision making and its development. In J. E. Jacobs & P. A. Klaczynski (Eds.), *The development of decision making in children and adolescents* (pp. 39–76). Mahwah, NJ: Erlbaum.

Kluger, A., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254.

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General, 135*(2), 162–183. https://dx.doi.org/10.1037/0096-3445.135.2.162.

Lang & Betsch (2018a). Instruction of simple, probabilistic decision strategies in children. *Manuscript in preparation.*

Lang & Betsch (2018b). Children's decision strategies in probabilistic environments with and without social models. *Manuscript in preparation.*

Lang & Betsch (2018c). Cross-cultural stability of children's decision making skills: Probabilistic utilization and decision strategies in Japanese and German children. *Manuscript in preparation.*

Lang & Betsch (2018d). Children's probability utilization for choices in gains and loss domains. *Manuscript in preparation.*

Legare, C., Mills, C., Souza, A. L. & Plummer, L., & Yasskin, R. (2013). The use of questions as problem-solving strategies in early childhood. *Journal of Experimental Child Psychology, 114*(1), 63–76. https://doi.org/10.1016/j.jecp.2012.07.002

Lehmann, A., & Betsch, T. (2018). The relation of executive functions and probabilistic decision making in children. *Manuscript in preparation.*

Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes, 116,* 286–295. https://doi.org/10.1016/j.obhdp.2011.05.001

Lindow, S., & Betsch, T. (2018). Child decision-making: On the burden of predecisional information search, J*ournal of Cognition and Development, 19*(2)*,* 137–164. https://doi.org/10.1080/15248372.2018.1436057

Lindow, S., Lang, A., & Betsch, T. (2017) Holistic information integration in child decision making. *Journal of Behavioral Decision Making*, *30*(5), 1131–1146. https://doi.org/10.1002/bdm.2029

Mata, R., von Helversen, B., & Rieskamp, J. (2011), When easy comes hard: The development of adaptive strategy selection. *Child Development, 82*, 687–700. https://doi.org/10.1111/j.1467-8624.2010.01535.x

McCormack, T., O'Connor, E., Beck, S. R., & Feeney, A. (2016). The development of regret and relief about the outcomes of risky decisions. *Journal of Experimental Child Psychology, 148*, 1–19. https://doi.org/10.1016/j.jecp.2016.02.008

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager,T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner, & R. Brunken (Eds.), *Instructional design for multimedia learning* (pp. 181–195). Munster, NY: Waxmann.

Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin & Review, 14*, 1133–1139. https://doi.org/10.3758/BF03193102

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135*(2), 207–236. https://doi.org/10.1037/0096-3445.135.2.207

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). Oxford: Wiley.

Roelfsema P. R., van Ooyen, A., & Watanabe T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences, 14*, 64–71. https://doi.org/10.1016/j.tics.2009.11.005

Savage, L. J. (1954). *The foundations of statistics*. Oxford, GB: Wiley.

Scheibehenne, B., Rieskamp, J., & Wagenmakers, E. (2013). Testing adaptive toolbox models: A Bayesian hierarchical approach. *Psychological Review, 120*, 39–64. https://doi.org/10.1037/a0030777

Schiebener, J., Zamarian, L., Delazer, M., & Brand, M. (2011). Executive functions, categorization of probabilities, and learning from feedback: What does really matter for decision making under explicit risk conditions? *Journal of Clinical and Experimental Neuropsychology, 33*(9), 1025–1039. https://doi.org/10.1080/13803395.2011.595702

Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *Child Development, 72*, 103–122. https://doi.org/10.1111/1467-8624.00268

Schlottmann, A., & Tring, J. (2005). How children reason about gains and losses: Framing effects in judgment and choice. *Swiss Journal of Psychology, 64*(3), 153–171. https://dx.doi.org/10.1024/1421-0185.64.3.153

Schlottmann, A., & Wilkening, F. (2012). Judgment and decision making in young children. In M. K. Dhami, A. Schlottmann, & M. R. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development and evolution* (pp. 55–83). Cambridge: University Press.

Shute, V. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. https://doi.org/10.1111/j.1467-7687.2007.00571.x

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69,* 99–118. Retrieved from http://www.jstor.org/stable/1884852

Sloman, S. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3-10. https://doi.org/1037/0033-2909.119.1.3.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2011). The complexity of developmental predictions from dual process models. *Developmental Review, 31*, 103–118. https://doi.org/10.1016/j.dr.2011.07.003

Steinbeis, N. & Crone, E. A. (2016). The link between cognitive control and decision-making across child and adolescent development. *Current Opinion in Behavioral Sciences, 10,* 28–32. https://doi.org/10.1016/j.cobeha.2016.04.009

van Duijvenvoorde, A. C. K, Jansen, B. R. J., Bredman, J. C., & Huizenga, H. M. (2012). Age-related changes in decision making: Comparing informed and noninformed situations. *Developmental Psychology*, *48*(1), 192–203. https://doi.org/10.1037/a0025601

Varghese, A.L. & Nilsen, E. (2013). Incentives improve the clarity of school-age children's referential statements, *Cognitive Development*, *28(4)*, 364–373. https://doi.org/10.1016/j.cogdev.2013.07.001.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03 von Neumann, J. & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Walsh, M. M., & Gluck, K. A. (2016) Verbalization of decision strategies in multiple-cue probabilistic inference. *Journal of Behavioral Decision Making*, *29*, 78–91. https://doi.org/10.1002/bdm.1878.

Weller, J. A., Levin, I. P., Rose, J. P., & Bossard, E. (2012), Assessment of decision-making competence in preadolescence. *Journal of Behavioral Decision Making*, *25(4)*, 414–426. https://doi.org/10.1002/bdm.744

Wilde, M., Bätz, K., Kovaleva, A. & Urhahn, D. (2009). Testing a short scale of intrinsic motivation. *Zeitschrift für Didaktik der Naturwissenschaften, 15*, 31–45. Retrieved from http://archiv.ipn.uni-kiel.de/zfdn/pdf/15_Wilde.pdf

Wilkening, F. & Anderson, A. (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.

Wohlwill, J. F. (1968): From perception to inference: A dimension of cognitive development. In I. E. Sigel & F. H. Hooper (Eds.), *Logical thinking in children* (pp. 472–494). New York, NY: Holt, Rinehard & Winston