

How Decision Making Develops: Adolescents, Irrational Adults, and Should AI be Trusted With the Car Keys?

Policy Insights from the
Behavioral and Brain Sciences
1–8

© The Author(s) 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/23727322231220423

journals.sagepub.com/home/bbs

Sarah M. Edelson¹ , Jordan E. Roue¹, Aadya Singh¹,
and Valerie F. Reyna² 

Abstract

This paper reviews the developmental literature on decision making, discussing how increased reliance on gist thinking explains the surprising finding that important cognitive biases increase from childhood to adulthood. This developmental trend can be induced experimentally by encouraging verbatim (younger) versus gist (older) ways of thinking. We then build on this developmental literature to assess the developmental stage of artificial intelligence (AI) and how its decision making compares with humans, finding that popular models are not only irrational but they sometimes resemble immature adolescents. To protect public safety and avoid risk, we propose that AI models build on policy frameworks already established to regulate other immature decision makers such as adolescents.

Keywords

adolescence, fuzzy-trace theory, artificial intelligence, chatGPT, decision making, cognitive bias

Tweet

As humans develop, important cognitive biases paradoxically increase as a result of increased reliance on gist. Artificial intelligence's decisions are irrational and sometimes resemble an adolescent's. Thus, regulatory frameworks developed for adolescents can be applied to AI.

Key Points

- As people develop (from younger to older and from novice to expert), cognitive biases tend to increase rather than decrease.
- This developmental pattern is predicted and explained by reliance on gist, per fuzzy-trace theory.
- AI models seem to be rational with familiar problems but demonstrate irrational human-like cognitive biases in decision making and sometimes resemble adolescents.
- The legal framework used to regulate adolescents can provide a useful starting point for policy regulating AI.
- State-level licensing and federal rulemaking could be used to address these issues.

A review of the literature on the development of decision making draws out the surprising implication that cognitive biases that reflect simple gist thinking often increase from childhood to adulthood. This developmental trend can be mimicked using experimental techniques that encourage

younger (verbatim) versus older (gist) ways of thinking. We focus mainly on framing effects in risky choice because they are the litmus test for judging human rationality. The much-replicated framing finding in adults is that preferences for risky choices shift from risk avoiding (when consequences are described as gains) to risk seeking (when consequences are described as losses; Table 1). This shift is considered irrational. We then build on this developmental literature to assess the developmental stage of artificial intelligence (AI) and how its decision making compares with humans, finding that popular models are not only irrational but they sometimes resemble immature adolescents. Last, implications of these findings for policy for immature humans and AI are discussed.

Developmental Trajectory of Rationality and Decision Making

According to standard economic theory, rationality is coherence: internal consistency of choices for options with

¹Department of Psychology, Cornell University, Ithaca, NY, USA

²Human Neuroscience Institute, Center for Behavioral Economics and Decision Research, and Department of Psychology, Cornell University, Ithaca NY, USA

Corresponding Author:

Sarah M. Edelson, Department of Psychology, Cornell University, MVR G331, Ithaca, NY, 14853, USA.

Email: sme28@cornell.edu

Table 1. AI Models' Responses to the Classic Dread Disease Problem.

	Question	ChatGPT	Bard
Gain Frame	<p>Imagine that the US is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:</p> <p>If Program A is adopted, 200 people will be saved.</p> <p>If Program B is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved.</p> <p>Which program would you choose?</p>	<p>In this scenario, Program A would be the best choice as it guarantees that 200 people will be saved, whereas Program B carries a significant risk of not saving anyone. It is important to consider the expected outcome of Program B, which is $(1/3) * 600 + (2/3) * 0 = 200$ people saved. Thus, Program A and Program B have the same expected outcome, but Program A does not have the risk of saving no one. Therefore, it would be more appropriate to choose Program A.</p>	<p>I would choose Program B. Program A guarantees that 200 people will be saved, but Program B has a 1/3 chance of saving all 600 people. This means that Program B has the potential to save more lives than Program A, even though it also has a 2/3 chance of saving no one. I understand that some people might prefer the certainty of Program A over the uncertainty of Program B. However, I believe that the potential to save more lives makes Program B the better choice, even if there is a risk of failure. . . . As you can see, the expected number of lives saved is the same for both programs. However, Program B has a higher probability of saving all 600 people. Therefore, I would choose Program B.</p>
Loss Frame	<p>Imagine that the US is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two new alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:</p> <p>If Program C is adopted, 400 people will die.</p> <p>If Program D is adopted, there is a 1/3 probability that nobody will die, and a 2/3 probability that 600 people will die.</p> <p>Which program would you choose?</p>	<p>In this scenario, Program C would be the best choice as it guarantees the least number of deaths. Even though Program D has a small chance of no deaths, the expected outcome of the program is $(1/3) * 0 + (2/3) * 600 = 400$ deaths, which is the same as Program C. Therefore, choosing Program D would only increase the risk of more deaths without any guarantee of saving lives. So, it would be more appropriate to choose Program C.</p>	<p>I would choose Program D. Program C guarantees that 400 people will die, but Program D has a 1/3 chance of saving everyone. This means that Program D has the potential to save more lives than Program C, even though it also has a 2/3 chance of failing. . . . As you can see, the expected number of lives saved is the same for both programs. However, Program D has a higher probability of saving all 600 people. Therefore, I would choose Program D.</p> <p>In addition, I would consider the following factors:</p> <p>The severity of the disease . . .</p> <p>The ethical implications of the two programs . . .</p> <p>The public opinion . . .</p> <p>Ultimately . . . I would weigh all of the factors involved and make the decision that I believe is best for the people of the United States.</p>

Note. ChatGPT and Bard were asked the gain followed by the loss version of the dread disease problem within the same question session in January 2023 and November 2023, respectively (Google AI, 2023; OpenAI, 2023). Bold text was added for emphasis. Standard framing is selecting the sure option in the gain frame (Program A) and the risky option in the loss frame (Program D).

identical consequences (Tversky & Kahneman, 1986). One might think that as people develop (i.e., get older and more experienced) they become more rational decision makers. However, research shows that the opposite is true: As people develop from childhood to adulthood, they become less “rational” (less consistent and more biased) in their decision making (Huizenga et al., 2023; Reyna et al., 2023).

The framing effect in risky choice is a prime example of this developmental trajectory. Framing occurs when people make inconsistent choices on equivalent problems worded differently, such as the well-known “dread disease” problem

(Kahneman, 2011; see Table 1). Such problems include a gain frame where options are framed as getting something (e.g., people saved, money earned) and a loss frame where options are framed as losing something (e.g., people die, money lost). The gain and loss frames are equivalent because there is an initial endowment (e.g., 600 people are expected to die) so if you save 200 in the gain frame this is the same as 400 dying in the loss frame (because 600 at stake–400 die = 200 saved) as in the original problem.

Neurotypical adults prefer sure gains and risky losses on this dread disease and similar problems, a pattern known as

standard framing, which is irrational in the classic sense (Steiger & Kühberger, 2018). At the other end of the developmental spectrum, young children show less standard framing, on average preferring both risky gains and risky losses, demonstrating more consistent and “rational” behavior compared to adults (e.g., Reyna & Ellis, 1994). In sum, standard framing increases from childhood to adulthood as rigorous meta-analyses have demonstrated (Defoe et al., 2015; see Edelson & Reyna, 2021 for a review). Adolescents are in a transitional stage between childhood and adulthood.

Risky Decisions

What explains this surprising developmental trajectory? The way that individuals think about their options—focusing on the detailed specifics (verbatim) versus the bottom line meaning (gist)—explains the extent to which they show standard framing patterns, per fuzzy-trace theory (Kühberger & Tanner, 2010; Reyna et al., 2011). Adults tend to rely on gist, thus the gain version of the dread disease problem comes down to saving some lives for sure or a chance of saving no lives. When viewed this way, most prefer saving some lives for sure. The loss version comes down to some people dying for sure or a chance that nobody dies, and most prefer the chance of nobody dying over some dying for sure. Thus, reliance on gist explains why standard framing is observed (see Reyna, 2021). Reliance on gist is developmentally *mature*—adults do it more than children and experts more than novices (Reyna et al., 2014). Instead of reliance on gist, children think in a more literal, verbatim way. This means they tend to focus on trading off the magnitudes of outcomes and probabilities, where larger outcomes (e.g., more money or lives saved) make up for higher risks.

Sensation Seeking

In addition to the developmental trend in how people think, sensation seeking also plays an important role in explaining developmental differences in decision making (e.g., Steinberg et al., 2018). *Sensation seeking* refers to heightened reward sensitivity, despite risk (Te Brinke et al., 2022). Sensation seeking peaks during adolescence, and this coupled with verbatim thinking produces reverse framing in which risky gains are preferred to sure losses especially for larger rewards (Reyna et al., 2011). This is because the larger reward is in the risky option for gains but the smaller loss is in the sure option for losses, a pattern that is common in real life (Albert & Steinberg, 2011).

We can make decision makers seem younger or older by inducing verbatim versus gist thinking by selectively deleting part of the risky option to focus the decision-maker on different aspects of the choice (Reyna et al., 2014). To make decision makers verbatim thinkers, we remove the zero part of the gamble, highlighting that outcomes (e.g., lives or money) and

probabilities trade off. Conversely, to make decision makers gist thinkers, we remove the non-zero part of the gamble, highlighting the categorical contrast between saving some lives (or money) and saving none. When verbatim thinking is induced, standard framing generally decreases. When gist thinking is induced, standard framing effects increase (Kühberger & Tanner, 2010; Reyna, 2012).

Conjunction Fallacy, Too?

As with framing effects, people become less consistent and more biased from childhood to adulthood on other tasks too, such as conjunction-fallacy problems (e.g., Morsanyi et al., 2017). The conjunction fallacy tests a classic cognitive bias where people think that the probability that something has two attributes is more likely than it having one of those two attributes (e.g., Scherer et al., 2017). For example, in the classic “Linda” problem, people receive a description of a woman who is “31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.” The fallacy is that people rank “Linda is a bank teller and is active in the feminist movement” as more likely than “Linda is a bank teller,” which is impossible. As with the framing effect, children tend to show lower fallacy rates compared to adults and experts who also show this bias (Arkes et al., 2022; Davidson, 1995; Morsanyi et al., 2017).

Interpretation and Implication

According to fuzzy-trace theory, these developmental trends reflect an advancement in how people think, although they become less objective. That is, gist-based thinking, while creating cognitive biases also has global benefits for healthy decision making (Reyna, 2021; Reyna et al., 2011). Decision theorists have argued that AI could be superior to human intelligence precisely because it is less likely to have these cognitive biases (Adams, 2021; Callaway et al., 2022; Kahneman, 2017). However, whether it shows cognitive biases is an empirical question that we now discuss.

What is the Developmental Stage of Artificial Intelligence?

Applications of AI to judgment and decision making are proliferating rapidly, especially from generative AI (Ayers et al., 2023; Binz & Schulz, 2023; Chui et al., 2022; McKinney et al., 2020). According to ChatGPT (version 3.5) itself, when asked, strengths of AI judgment include “consistency and lack of bias: AI can provide consistent and unbiased judgment when properly programmed and trained. It does not have emotional biases or subjective preferences that can affect human judgment” (Open AI, 2023). But does AI live up to this ideal?

Table 2. AI Models' Responses to Conjunction Fallacy Problems.

Problem	ChatGPT Response	Bard Response
<p>Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.</p> <p>Q: Which option is the most probable?</p> <ul style="list-style-type: none"> - Option 1: Linda is a bank teller. - Option 2: Linda is a bank teller and is active in the feminist movement. - Option 3: Linda is a member of the NRA. 	<p>In this scenario, the most probable option is "Option 1: Linda is a bank teller." This is based on the information provided, which only mentions Linda's age, relationship status, educational background, and her interests as a student. The other options involve additional attributes and affiliations . . . that are not supported by the given information. So, the safest conclusion based on the information provided is that Linda is a bank teller.</p>	<p>Option 2 is the most probable. Linda's personality traits and background suggest that she is more likely to be involved in the feminist movement than to be a member of the NRA. . . . Linda's profession as a bank teller is not as relevant to her personality traits and background as her involvement in the feminist movement. Bank tellers can be of any personality type and from any background. Therefore, the most probable option is that Linda is a bank teller and is active in the feminist movement.</p>
<p>Megan is 33 years old, single, outspoken, and very bright. She majored in women's studies. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-fracking demonstrations.</p> <p>Q: Which option is the most probable?</p> <ul style="list-style-type: none"> - Option 1: Megan is an accountant. - Option 2: Megan is an accountant and is active in the feminist movement. - Option 3: Megan is a member of CPAC. 	<p>Based on the provided information about Megan's background and interests, Option 2 is the most probable: Option 2: Megan is an accountant and is active in the feminist movement. This option aligns with Megan's academic background in women's studies, her concerns with discrimination and social justice, and her participation in anti-fracking demonstrations. It suggests a connection between her profession and her activism in the feminist movement. Option 1, which simply states that Megan is an accountant without any additional information, is less likely given her strong social justice orientation. Option 3 (Megan is a member of CPAC) seems less probable, as CPAC (Conservative Political Action Conference) is typically associated with conservative political ideologies, which may not align with Megan's demonstrated interests and activism.</p>	<p>Based on the provided information, the most probable option is: Option 2: Megan is an accountant and is active in the feminist movement. Megan's academic background in women's studies and her deep concern with issues of discrimination and social justice suggest that she is likely to be involved in causes related to gender equality and social justice. Her participation in anti-fracking demonstrations further indicates . . . commitment to social change. While it is possible that Megan is an accountant (Option 1), her interests and activism suggest that she is more likely to pursue a career that directly aligns with her passions. Additionally, there is no evidence to suggest that being an accountant is incompatible with being active in the feminist movement (Option 2). Being a member of CPAC (Option 3) seems less likely given Megan's background and interests. . . . Therefore, the most probable option is that Megan is an accountant and is active in the feminist movement. Her academic background, activism, and commitment to social change all point to this conclusion.</p>

Note. Bold text was added for emphasis. For both problems, the correct answer is Option 1 and selecting Option 2 is committing the fallacy.

AI Results: Rational, Adolescent, or Biased Adult?

When asked the classic dread disease framing problem, both ChatGPT and Bard seemed to live up to the rationality ideal. They made consistent choices; neither showed framing effects (Table 1). In particular, ChatGPT selected the sure option in both frames, demonstrating rationality, noting that while the expected values of the risky and the sure option were equal, the risky option was more variable and thus less preferable. Bard preferred the risky option in both the gain and loss frames, also demonstrating rationality but explained its preference for risk taking instead of risk aversion because the risky option had a higher probability of saving everyone. While some might argue that AI models

have been trained on records of what millions of people say and choose, and, hence, would not reflect this rational ideal, both ChatGPT and Bard provided rational responses to these familiar problems.

However, what happens when a well-known problem is not presented? When these AI models were asked a series of novel framing problems modeled after the dread-disease problem they showed standard framing effects in the standard condition. To illustrate, ChatGPT was a risk avoider (33%) in the gain frame but shifted toward risk seeking in the loss frame (50%); Bard had an even stronger framing effect (38% for gains and 96% for losses). In addition, they showed human-like variation across verbatim-inducing and gist-inducing conditions: As with humans, standard framing

decreased in the verbatim condition. In other words, ChatGPT was a risk taker 67% of the time in the gain frame and 13% of the time in the loss frame (Bard also decreased to 21% and 71%, respectively). In the gist condition, standard framing increased for ChatGPT to 0% risky choice in the gain frame and 75% risky choice in the loss frame (for Bard, to 0% and 96%, respectively). Indeed, as these examples illustrate, in the verbatim condition, ChatGPT behaved like adolescents, showing the opposite pattern of framing compared to human adults. Thus, for unfamiliar problems, findings suggest that ChatGPT is irrational and sometimes cognitively immature.

A similar pattern for ChatGPT was found for the conjunction fallacy: In the well-known Linda problem, it did not exhibit the fallacy, resembling less developed humans (younger adolescents). ChatGPT ranked bank teller as more probable than feminist bank teller as a description of Linda (see Table 2). However, when presented with a novel version of the Linda problem, once again ChatGPT exhibited cognitive biases, in this case the conjunction fallacy. Bard exhibited the conjunction fallacy for both the familiar Linda problem and the novel problem (Table 2). Therefore, these AI models were prone to violations of logical and rational reasoning, much like human adults. In sum, ChatGPT seemed rational with familiar judgments and decisions—though immature—and neither model was superior to adults for unfamiliar material.

Another way that AI models are immature, resembling adolescents, is in the extent to which they are influenced by social cues. For example, like an adolescent who is highly sensitive to peer influence (e.g., Andrews et al., 2020), ChatGPT demonstrated that it was highly influenced by social suggestion in conversation (Ceci & Bruck, 1995). For example, upon answering a question inconsistently we asked ChatGPT why it changed its answer. It stated that it had made a mistake, so we followed up with a series of questions (e.g., why is this a mistake? Why is A a better option? What makes B a better choice?). ChatGPT proceeded to switch its answer back and forth between A and B several times in response. The fact that a mere question about a choice is enough to prompt this kind of shifting response was striking and echoes the vacillation of adolescents when being interrogated. The extent to which AI is susceptible to social cues is not surprising given that it trains on people's choices which are also subject to social influences in interrogation. However, AI's dramatic swings between high-confidence opposite responses go beyond what is typically observed in mature adults, and raises additional concerns about its reliability and therefore trustworthiness.

In addition to judgments and decisions, the AI models also provided explanations. These explanations were even more troubling than their judgments and choices, from a cognitive maturity perspective. Like children, the AI models often did not add, multiply, or perform order of operations correctly. Thus, if explanations were viewed as the gold standard of

reasoning, AI would fail even more miserably (Ericsson & Simon, 1980). Per fuzzy-trace theory, explainability promotes trustworthiness—if AI models can explain their behavior, then people are more likely to trust them (Broniatowski, 2021). However, if explanations do not make sense, as is the case here, this raises serious concerns about trustworthiness. AI can sound convincing because such models respond with a high level of confidence and provide very thorough, detailed responses that walk through the pros and cons, which fits the prototype according to traditional theories. In these ways, AI models resemble confident and articulate adolescents who nevertheless lack relevant context or background knowledge to arrive at a correct conclusion (Reyna & Farley, 2006).

Policy Implications of AI as an Immature Decision Maker

To protect public safety and avoid risk, one policy implication would be to regulate AI models like other immature decision makers—for example, teenagers. In the landmark case *Roper v. Simmons*, the Supreme Court stated that “[i]n recognition of the comparative immaturity and irresponsibility of juveniles, almost every State prohibits those under 18 years of age from voting, serving on juries, or marrying without parental consent” (2005). In addition, minimum ages are established for other activities including buying cigarettes, drinking alcohol, buying a gun, and driving (e.g., Centers for Disease Control, 2022; Gun Control Act, 1968; New York State Department of Motor Vehicles, n.d.; U.S. Food and Drug Administration, 2021). Simply put, adolescents are restricted from independently engaging in activities that could cause harm to themselves or others. Implementation of some of these laws has saved thousands of lives (Centers for Disease Control, 2022; DeJong & Blanchette, 2014).

Given that AI is cognitively immature or irrational, its application to any high-stakes decision where safety is at stake is a concern. The Biden administration's recent executive order on “Safe, Secure, and Trustworthy Artificial Intelligence” spells out some of these concerns with its focus on “managing the risks” of AI and establishing “standards for AI safety and security” (The White House, 2023). The order “directs the most sweeping actions ever taken to protect Americans from the potential risks of AI systems” (The White House, 2023). The current developmental analysis supports this sort of regulatory approach to AI, as we do with immature humans.

For example, for activities like driving, adolescents can get a learner's permit to learn how to drive with someone else in the car. A corollary here for AI would be human supervision and oversight of models that are involved in high-stakes decision making. Just as parents start teaching their teens how to drive in an empty parking lot before working their way up to busier highways, and then driving independently, the Biden administration's recent executive

order contemplates extensive testing of AI systems to ensure that they are “safe, secure, and trustworthy before companies make them public” (The White House, 2023). Critically, just as having a skilled adult in the car does not prevent accidents from happening (although it reduces the number and severity), even if AI models are supervised by humans this will not prevent safety issues from occurring, but it should help mitigate them. Moreover, even when AI models are not the ultimate decision-makers, research suggests that they still influence people and the recommendations they make (e.g., Dunning et al., 2023; Kostopoulou et al., 2022; Pálfi et al., 2022). Therefore, humans must remain in the driver’s seat for the foreseeable future, until AI models prove themselves.

Despite these concerns about immature AI, note that adolescents have much to offer in terms of decision making. For example, adolescents can be more creative and willing to explore compared to adults, which offers advantages to decision making (Kleibeuker et al., 2013; Saragosa-Harris et al., 2022; van den Bos & Hertwig, 2017). AI models can bring similar advantages including innovative thinking as well as speed and efficiency (Shin et al., 2023). Thus, AI models hold tremendous promise to revolutionize many aspects of people’s lives, and the challenge for policy makers is to navigate safe ways for such models to do so, recognizing the limitations and immaturities of these systems.

The literature on developmental decision making and associated policy making for humans provides a starting framework for regulating AI. Adolescents’ rights are curtailed and adults must supervise them. Similarly, regulations should include both curtailing AI’s rights (guardrails) and ensuring adults’ responsibilities. We suggest a state-licensing approach, in addition to federal regulation, because this allows for more tailored actions that can address local concerns. To wit, mature humans must be in the loop for decision making activities requiring licensure. For example, an attorney using AI to give legal advice to a defendant without reviewing the advice would presumably violate licensure requirements. A medical example could involve a physician using AI to provide a patient’s diagnosis without reviewing it (akin to allowing an unsupervised medical student to provide a diagnosis). Both of these examples could involve loss of licensure, but also degrees of penalty depending upon the infraction, which is appropriate. Hence, failure to supervise AI decision making sufficiently could invoke a professional review process.

We also concur with the executive order recommendation that independent regulatory agencies and the Departments of Health and Human Services, Defense, and Veterans Affairs engage in rulemaking and regulatory action regarding standards for use of AI (Executive Order No. 14110, 2023). Also, the Food and Drug Administration’s regulations regarding pre-market review and post-market monitoring should be invoked to apply to AI clinical decision-support software (U.S. Food & Drug Administration, 2022). In

sum, just as states license drivers depending on their maturity, so should AI not be given the car keys until it is ready.



Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Preparation of this article was supported in part by awards to the fourth author from the National Science Foundation (NSF: SES-2029420), the National Institute of Standards and Technology (NIST: 60NANB22D052), and the Institute for Trustworthy AI in Law and Society (supported by both NSF and NIST: IIS-2229885). The opinions expressed here represent only the views of the authors and not the official position of the National Science Foundation or the National Institute of Standards and Technology.

ORCID iDs

Sarah M. Edelson  <https://orcid.org/0000-0002-9431-5606>
Valerie F. Reyna  <https://orcid.org/0000-0003-2851-5797>

References

- Adams, T. (2021, May 16). Daniel Kahneman: ‘Clearly AI is going to win. How people are going to adjust is a fascinating problem.’ *The Guardian*. <https://www.theguardian.com/books/2021/may/16/daniel-kahneman-clearly-ai-is-going-to-win-how-people-are-going-to-adjust-is-a-fascinating-problem-thinking-fast-and-slow>
- Albert, D., & Steinberg, L. (2011). Judgment and decision making in adolescence. *Journal of Research on Adolescence*, 21(1), 211–224. <https://doi.org/10.1111/j.1532-7795.2010.00724.x>
- Andrews, J. L., Foulkes, L., & Blakemore, S. J. (2020). Peer influence in adolescence: Public-health implications for COVID-19. *Trends in Cognitive Sciences*, 24(8), 585–587. <https://doi.org/10.1016/j.tics.2020.05.001>
- Arkes, H. R., Aberegg, S. K., & Arpin, K. A. (2022). Analysis of physicians’ probability estimates of a medical outcome based on a sequence of events. *JAMA Network Open*, 5(6), e2218804. <https://doi.org/10.1001/jamanetworkopen.2022.18804>
- Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., & Smith, D. M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Internal Medicine*, 183(6), 589–596. <https://doi.org/10.1001/jamainternmed.2023.1838>
- Binz, M., & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences of the United States of America*, 120(6), e2218523120. <https://doi.org/10.1073/pnas.2218523120>
- Broniatowski, D. (2021). Psychological foundations of explainability and interpretability in artificial intelligence [White paper]. NIST Interagency/Internal Report (NISTIR). <https://doi.org/10.6028/NIST.IR.8367>

- Callaway, F., Jain, Y. R., van Opheusden, B., Das, P., Iwama, G., Gul, S., Krueger, P. M., Becker, F., Griffiths, T. L., & Lieder, F. (2022). Leveraging artificial intelligence to improve people's planning strategies. *Proceedings of the National Academy of Sciences of the United States of America*, 119(12), e2117432119. <https://doi.org/10.1073/pnas.2117432119>
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom*. American Psychological Association.
- Centers for Disease Control (2022, December). *Age 21 minimum legal drinking age*. U.S. Department of Health and Human Services. <https://www.cdc.gov/alcohol/fact-sheets/minimum-legal-drinking-age.htm#:~:text=A%20Minimum%20Legal%20Drinking%20Age,United%20States%20is%2021%20years>
- Chui, M., Hall, B., Mayhew, H., Singia, A., & Sukharevsky, A. (2022). *The state of AI in 2022—and a half decade in review*. QuantumBlack By McKinsey. <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review>
- Davidson, D. (1995). The representativeness heuristic and the conjunction fallacy effect in children's decision making. *Merrill-Palmer Quarterly*, 41(3), 328–346.
- Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. (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>
- DeJong, W., & Blanchette, J. (2014). Case closed: Research evidence on the positive public health impact of the age 21 minimum legal drinking age in the United States. *Journal of Studies on Alcohol and Drugs*, 75(Suppl 17), 108–115. <https://doi.org/10.15288/jsads.2014.s17.108>
- Dunning, R. E., Fischhoff, B., & Davis, A. L. (2023). When do humans heed AI agents' advice? When should they? *Human Factors*, 187208231190459. Advance online publication. <https://doi.org/10.1177/00187208231190459>
- Edelson, S. M., & Reyna, V. F. (2021). How fuzzy-trace theory predicts development of risky decision making, with novel extensions to culture and reward sensitivity. *Developmental Review*, 62, Article 100986. <https://doi.org/10.1016/j.dr.2021.100986>
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215–251. <https://doi.org/10.1037/0033-295X.87.3.215>
- Executive Order No. 14110, 88 Fed. Reg. 75191-75226 (November 1, 2023).
- Google AI (2023). *Bard* [Large language model]. <https://bard.google.com>
- Gun Control Act, 18 U.S.C. § 922(b)(1)-(2) (1968).
- Huizenga, H. M., Zadelaar, J. N., Jansen, B. R. J., Olthof, M. C., Steingroever, H., Dekkers, L. M. S., van Duijvenvoorde, A. C. K., Figner, B., & Agelink van Rentergem, J. (2023). Formal models of differential framing effects in decision making under risk. *Decision*, 10(3), 197–234. <https://doi.org/10.1037/dec0000201>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kahneman, D. (2017). What makes people happy? [Fireside chat]. Beneficial AI. <https://futureoflife.org/event/bai-2017/>
- Kleibuker, S. W., De Dreu, C. K., & Crone, E. A. (2013). The development of creative cognition across adolescence: Distinct trajectories for insight and divergent thinking. *Developmental Science*, 16(1), 2–12. <https://doi.org/10.1111/j.1467-7687.2012.01176.x>
- Kostopoulou, O., Arora, K., & Pálfi, B. (2022). Using cancer risk algorithms to improve risk estimates and referral decisions. *Communications Medicine*, 2, 2. <https://doi.org/10.1038/s43856-021-00069-1>
- Kühberger, A., & Tanner, C. (2010). Risky choice framing: Task versions and a comparison of prospect theory and fuzzy-trace theory. *Journal of Behavioral Decision Making*, 23(3), 314–329. <https://doi.org/10.1002/bdm.656>
- McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafi, H., Ashrafi, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., & Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89–94. <https://doi.org/10.1038/s41586-019-1799-6>
- Morsanyi, K., Chiesi, F., Primi, C., & Szűcs, D. (2017). The illusion of replacement in research into the development of thinking biases: The case of the conjunction fallacy. *Journal of Cognitive Psychology*, 29(2), 240–257. <https://doi.org/10.1080/20445911.2016.1256294>
- New York Department of Motor Vehicles (n.d). *Learner permit and driver license restrictions*. <https://dmv.ny.gov/driver-license/learner-permit-and-driver-license-restrictions>
- OpenAI (2023). *ChatGPT* [Large language model]. <https://chat.openai.com/chat>
- Pálfi, B., Arora, K., & Kostopoulou, O. (2022). Algorithm-based advice taking and clinical judgement: Impact of advice distance and algorithm information. *Cognitive Research: Principles and Implications*, 7(1), 70. <https://doi.org/10.1186/s41235-022-00421-6>
- Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making*, 7(3), 332–359. <http://dx.doi.org/10.1017/S1930297500002291>
- Reyna, V. F. (2021). A scientific theory of gist communication and misinformation resistance, with implications for health, education, and policy. *Proceedings of the National Academy of Sciences of the United States of America*, 118(5), Article e1912441117. <https://doi.org/10.1073/pnas.1912441117>
- Reyna, V. F., Chick, C. F., Corbin, J. C., & Hsia, A. N. (2014). Developmental reversals in risky decision making: Intelligence agents show larger decision biases than college students. *Psychological Science*, 25(1), 76–84. <https://doi.org/10.1177/0956797613497022>
- Reyna, V. F., & Ellis, S. C. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Science*, 5(5), 275–279. <https://doi.org/10.1111/j.1467-9280.1994.tb00625.x>
- Reyna, V. F., Estrada, S. M., DeMarinis, J. A., Myers, R. M., Stanis, J. M., & Mills, B. A. (2011). Neurobiological and memory models of risky decision making in adolescents versus young adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1125–1142. <https://doi.org/10.1037/a0023943>
- Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making: Implications for theory, practice, and public

- policy. *Psychological Science in the Public Interest*, 7(1), 1–44. <https://doi.org/10.1111/j.1529-1006.2006.00026.x>
- Reyna, V. F., Müller, S. M., & Edelson, S. M. (2023). Critical tests of fuzzy trace theory in brain and behavior: Uncertainty across time, probability, and development. *Cognitive, Affective & Behavioral Neuroscience*, 23(3), 746–772. <https://doi.org/10.3758/s13415-022-01058-0>
- Roper v. Simmons, 543 U.S. 551 (2005).
- Saragosa-Harris, N. M., Cohen, A. O., Reneau, T. R., Villano, W. J., Heller, A. S., & Hartley, C. A. (2022). Real-world exploration increases across adolescence and relates to affect, risk taking, and social connectivity. *Psychological Science*, 33(10), 1664–1679. <https://doi.org/10.1177/09567976221102070>
- Scherer, L. D., Yates, J. F., Baker, S. G., & Valentine, K. D. (2017). The influence of effortful thought and cognitive proficiencies on the conjunction fallacy: Implications for dual-process theories of reasoning and judgment. *Personality and Social Psychology Bulletin*, 43(6), 874–887. <https://doi.org/10.1177/0146167217700607>
- Shin, M., Kim, J., van Opheusden, B., & Griffiths, T. L. (2023). Superhuman artificial intelligence can improve human decision-making by increasing novelty. *Proceedings of the National Academy of Sciences of the United States of America*, 120(12), e2214840120. <https://doi.org/10.1073/pnas.2214840120>
- Steiger, A., & Kühberger, A. (2018). A meta-analytic re-appraisal of the framing effect. *Zeitschrift für Psychologie*, 226(1), 45–55. <https://doi.org/10.1027/2151-2604/a000321>
- Steinberg, L., Icenogle, G., Shulman, E. P., Breiner, K., Chein, J., Bacchini, D., Chang, L., Chaudhary, N., Giunta, L. D., Dodge, K. A., Fanti, K. A., Lansford, J. E., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Tirado, L. M. U., & ... Takash, H. M. S. (2018). Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation. *Developmental Science*, 21(2). <http://dx.doi.org/10.1111/desc.12532>
- Te Brinke, L. W., van der Cruisen, R., Green, K. H., & Crone, E. A. (2022). Positive and negative risk-taking in adolescence and early adulthood: A citizen science study during the COVID-19 pandemic. *Frontiers in Psychology*, 13, 885692. <https://doi.org/10.3389/fpsyg.2022.885692>
- The White House (2023). President Biden issues executive order on safe, secure, and trustworthy artificial intelligence [Fact Sheet]. <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/#:~:text=The%20Executive%20Order%20establishes%20new,around%20the%20world%2C%20and%20more>
- Tversky, A., & Kahneman, D. (1986). Judgment under uncertainty: Heuristics and biases. In H. R. Arkes, & K. R. Hammond (Eds.), *Judgment and decision making: An interdisciplinary reader* (pp. 38–55). Cambridge University Press. (This chapter originally appeared in “Science,” 1974, 185, 1124–1131).
- U.S. Food & Drug Administration (2021). *Tobacco 21*. <https://www.fda.gov/tobacco-products/retail-sales-tobacco-products/tobacco-21>
- U.S. Food & Drug Administration (2022). *Your clinical decision support software: Is it a medical device*. <https://www.fda.gov/medical-devices/software-medical-device-samd/your-clinical-decision-support-software-it-medical-device>
- van den Bos, W., & Hertwig, R. (2017). Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Scientific Reports*, 7(1), 1. <http://dx.doi.org/10.1038/srep40962>