International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Adolescent Reasoning on Probability Paradoxes: Cognitive Mechanisms Underlying Intuitive Success and Error

Varad Acharya[1], Harshit Mishra[2]

[1]Dehradun, Uttarakhand, India
Corresponding Author Email: *varad1981op[at]gmail.com*

[2]Dehradun, Uttarakhand, India
Corresponding Author Email: *harshmodule19[at]gmail.com*

Abstract: ***This study examined how 106 adolescents (ages 15–20, M=17.48, SD=1.67) approached three probability paradoxes: Monty Hall, Birthday Paradox, and Gambler's Fallacy. Performance varied dramatically across paradoxes, χ^2 (2, N=106) = 59.59, p<.001: Gambler's Fallacy (80.2% correct), Monty Hall (33.0%), and Birthday Paradox (6.6%). Advanced mathematics predicted Monty Hall success (OR=1.78) and especially Birthday Paradox success (OR=4.94). Female participants outperformed males on Gambler's Fallacy (OR=1.90), but had a reversed effect on Birthday Paradox (OR=0.5) an unexpected finding. Qualitative analysis of reasoning explanations identified systematic bias patterns: equiprobability thinking (Monty Hall), representativeness heuristic (Birthday), and sound independence reasoning (Gambler's Fallacy). Adolescents reported high confidence (M=7.84–8.27/10) despite 20–39% error rates, indicating severe metacognitive blindness on difficult tasks. These results demonstrate that adolescent probability reasoning is paradox-dependent: intuitive System 1 reasoning succeeds on independence tasks but fails catastrophically on conditional probability and combinatorics. Traditional instruction adequately teaches independence but neglects conditional probabilities and combinatorics misconceptions. Evidence-based interventions with explicit mechanism teaching, simulation-based learning, and confidence calibration training may bridge these gaps while preserving accurate intuitions on independence reasoning.***

Keywords: probability paradoxes, adolescent cognition, conditional probability, Monty Hall Problem, cognitive biases, dual-process theory, mathematical reasoning.

1. Introduction

Probability paradoxes expose the systematic gap between human intuition and mathematical reality. The Monty Hall problem, Birthday Paradox, and Gambler's Fallacy are canonical demonstrations: ordinary adults fail 60–90% of the time on all three (Kahneman & Tversky, 1974; Granberg & Brown, 1995).

These errors are not random; they follow predictable patterns rooted in cognitive heuristics—mental shortcuts that usually work but fail spectacularly on probability tasks. Yet adolescents remain largely absent from this research. This is a critical oversight. Adolescence (ages 15–20) is precisely when probabilistic reasoning develops, when prefrontal cortex maturation continues through age 25 (Steinberg, 2008), and when young people make decisions—about health, finance, education—that depend on probabilistic thinking. We know adults fail. We don't know whether adolescents fail the same way, or whether formal mathematics instruction inoculates them against these biases.

Despite the extensive evidence on adults, surprisingly only little is known about how these biases emerge or differ during adolescence.

Five decades of probability research has mapped adult misconceptions in detail. Monty Hall errors stem from equiprobability bias and anchoring (Granberg & Brown, 1995). Birthday Paradox failures reflect representativeness heuristic (Kahneman & Tversky, 1972). Gambler's Fallacy errors involve misunderstanding independence (Gilovich et al., 1985). What we don't know: Do adolescents show the same relative difficulty ranking as adults? Or does development shuffle the rankings? Do 15-year-olds differ from 20-year-olds? Which demographic and cognitive factors predict who succeeds?

To tackle these questions, the present study systematically investigates adolescent's reasoning across some classical probability paradoxes.

This study addressed three questions:
1) How accurately do adolescents solve probability paradoxes? Overall accuracy and by paradox type.
2) Do paradoxes differ in difficulty? Are some systematically easier or harder than others?
3) What predicts success? Age, gender, mathematics training, prior exposure, mathematical confidence or something else?

To find answers to these questions, we should understand how they think about probability.

Dual-Process Theory distinguishes automatic (System 1) from deliberate (System 2) reasoning (Kahneman & Tversky, 1979). System 1 is fast, intuitive, heuristic-based—it evolved for speed, not accuracy. Probability paradoxes reveal System 1's failure: it produces systematic deviations through three key mechanisms:
- Availability heuristic: Overweighting easily recalled outcomes (initial door choice remains cognitively "available" in Monty Hall)

- Anchoring bias: Over-reliance on initial information despite new evidence
- Representativeness heuristic: Assuming small samples mirror populations (catastrophic misconceptions arise due to combinations)

System 2 is deliberate, effortful, rule based. It can override System 1, but only when engaged. Adolescents have developing System 2 capacity, suggesting they may show even stronger System 1 dominance than adults. The Birthday Paradox—with its 93% error rate—suggests the representativeness heuristic is so deeply entrenched that even adolescents with advanced mathematics training cannot override it. This shows a very prevalent underdevelopment of system 1.

Fuzzy-Trace Theory proposes a second angle: people encode information at two levels verbatim (precise details) and gist (intuitive essence) (Reyna & Brainerd, 2011). Effective reasoning relies on gist: understanding the meaning without getting lost in precision. However, gist-based processing develops gradually. Adolescents may over-rely on verbatim details precisely where probability reasoning demands gist extraction. So, at ground level they may ponder over small details where it is supposed to be viewed as a whole.

2. Literature Review

The Three Paradoxes

Research on classic probability tasks shows that people struggle even when the correct solutions are straightforward. In the Monty Hall problem, participants choose one of three doors and then decide whether to switch after the host reveals a non-prize door; although switching yields a 2/3 chance of winning, studies consistently find that most judge both options as equal, reflecting equiprobability assumptions and anchoring (Granberg & Brown, 1995; Krauss & Wang, 2003). In the Birthday Paradox, where individuals estimate how many people are needed for a shared birthday, the correct threshold is only 23 for the probability to exceed 50 percent, yet people typically propose far larger numbers due to representativeness-based reasoning (Kahneman & Tversky, 1972). Experiments on the Gambler's Fallacy present sequences of repeated outcomes and show that many assume the opposite result becomes more likely, despite each event remaining statistically independent with unchanged probabilities. Taken together, these tasks show how different problem structures reliably elicit systematic biases, revealing where intuitive reasoning diverges from formal probability.

Adolescent Development and Probabilistic Reasoning

Adolescence is marked by dual trajectories: advancing formal reasoning capacity alongside ongoing prefrontal maturation. Prefrontal cortex development continues through age 25, with selective development in ventromedial regions responsible for deliberative decision-making (Steinberg, 2008; Ravindranath et al., 2024). Through adolescence, simultaneous advances in probabilistic and mathematical reasoning continue (Placi et al., 2024), but heuristic biases frequently carry on despite mathematical knowledge (Tubau & Alonso, 2003).

Toplak and Stanovich (2024) found that adolescent's probabilistic reasoning develops as part of broader rational thinking, influenced more by cognitive style than raw intelligence.

According to fuzzy trace theory, adolescents gradually develop gist-based reasoning, initially relying too heavily on verbatim processing (Mills, Reyna, & Estrada, 2008). Due to their weak System 2 capacity and verbatim over-reliance, adolescents may be less able to solve probability paradoxes than adults. Based on these developmental limitations, studies also highlight how individual differences influence adolescents' ability to solve probability paradoxes. Prior probability exposure shows weak to modest correlation with paradox performance (Tubau & Alonso, 2003; Krauss & Wang, 2003). Knowing the correct answer is insufficient; reasoning for its correctness matters more (Evans, Over, & Manktelow, 2000). Advanced maths coursework predicts modest improvements but often modest 20–30% at best (Krauss & Wang, 2003). Mathematical confidence shows weak positive correlations with reasoning accuracy but does not eliminate biases (Evans et al., 2000). Gender differences in probability reasoning are understudied. Adult literature shows minimal gender effects on most paradoxes, though some work suggests girls develop probabilistic reasoning earlier (Ravindranath et al., 2024).

3. Methodology

Participants

A validated 16-item online survey was used to administer a cross-sectional study on teenage probability reasoning through Google Forms.

Sample: N=106 adolescents (final sample after removing 2 with missing mathematics confidence data; original N=108)

Table 1

Variable	N or Mean (SD)	% or Range
Sample Size	106	
Age	17.48 (1.67)	15-20
Male	87	82.1%
Female	19	17.9%
High School Student	70	66.0%
Advanced Mathematics	47	44.3%
Priorly Exposure-Gained	74	69.8%
Self-Reported Math Confidence	3.12 (1.11)	1-5

Data Collection: August 21–31, 2025. Online survey via Google Forms. Mean completion time: 18.5 minutes (SD=4.2).

Sampling Note: Participants were recruited from an open online survey via Google forms.

Inclusion criteria: ages 15–20, currently enrolled in secondary or tertiary education, English fluent.

Exclusion criteria: None applied.

Ethical Considerations

The U. S. Common Rule and the Belmont Report's ethical guidelines were followed in this survey. It was voluntary to participate, and doing so implied informed consent. The study was exempt from IRB review because no personally identifiable information was gathered and participation posed very minimal risk. This was established before any data was gathered.

Instruments and Analytical measures

The Logical Questionnaire (16 Items): Reliability & Validity: Pilot testing (N=30, excluded from main study): Cronbach's α=0.78 (acceptable internal consistency). Expert content review (3 mathematics educators, 2 cognitive psychologists): κ=0.87 (substantial agreement). Two-week test-retest reliability (n=20 subset): r=0.84 *(*stable measurement). Scoring: 1 point if correct and 0 for anything else.

Demographic Items (1–6):
Q1: Gender (Male/Female)
Q2: Age (15–20)
Q3: High School Status (Yes/No)
Q4: Advanced Mathematics Programs (Yes/No)
Q5: Mathematical Confidence (1–5 star scale)
Q6: Prior Probability Exposure (Yes/No)

Monty Hall Items (7–9):
Q7: Classic Monty Hall setup: choose a door, host reveals an empty door, you choose stay/switch.
Correct Answer: Switch to the other unopened door.
Q8: Confidence in answer. (1–10 scale)
Q9: Reasoning (checkbox items + open ended responses)

Gambler's Fallacy Items (10–12):
Q10: Coin lands heads 50 times in a row. What is the probability the next toss will be heads? (Less than 50% / Exactly 50% / More than 50%). Correct Answer: Exactly 50%
Q11: Confidence in answer (1–10)
Q12: Reasoning (checkbox items + open ended responses)

Birthday Paradox Items (13–15):
Q13: In a group, what is the expected group size if the probability that at least 2 people share the same birthday is 50%? (50% of days in year / 50% of days in month / No. of days in year / No. of days in month / Other numeric values). Correct Answers: 22, 23, 24 (23 being the correct answer with 50.7% same birthday chances with a [+–1] room for approximation error).
Q14: Confidence in answer (1–10)
Q15: Reasoning (checkbox items + open ended responses).

Analytical Measures: χ^2 Test: χ^2 analysis tested whether accuracy differed significantly across paradox types. Contingency table: Paradox Type (3 levels: Monty Hall, Gambler's Fallacy, Birthday) × Accuracy (2 levels: Correct, Incorrect).

Logistic Regression: Three separate logistic regression models predicted accuracy (0/1 outcome) on each paradox using five predictors: Age, Gender, Prior Probability Exposure, Mathematics

Confidence, Advanced Mathematics Coursework. Each model reported odds ratios (ORs) with 95% confidence intervals, β coefficients, and model accuracy.

Qualitative Analysis: Open-ended reasoning responses (Q9, Q12, Q15) were coded for cognitive bias mechanisms by two independent raters. Inter-rater agreement: Cohen's $\kappa \geq 0.75$ required.

4. Result

Accuracy by Paradox Type

X^2 Analysis

Table 3

Paradox	Accuracy		Mean Confidence (SD)
	N	%	
Monty Hall	35	33.00%	7.84 (2.77)
Gambler's Fallacy	85	80.20%	8.27 (2.46)
Birthday Paradox	7	6.60%	5.61 (3.05)

Interpretation: There was a 74-point gap between the best performance (Gambler's Fallacy) and the worst (Birthday Paradox). Results were highly significant, χ^2 (2, N=106) =122.81, p<.001. The Birthday Paradox showed a 93% error rate. Monty Hall accuracy was 33%, with high confidence, showing a confidence-accuracy disconnect among the adolescents tested.

Chi-square test revealed highly significant differences in accuracy across paradoxes:

Table 4

Paradox	Correct	Incorrect	Standardized Residual
Monty Hall	35	71	-1.13
Gambler's Fallacy	85	21	+6.56
Birthday Paradox	7	99	-5.43

X^2 (2, N=106) = 59.59, p <.001 (highly significant)

Interpretation: The Gambler's Fallacy showed a strong positive residual (+6.56), suggesting it was much easier than expected. The Birthday Paradox showed a large negative residual (−5.43), indicating far greater difficulty than predicted. Monty Hall displayed a smaller negative residual (−1.13). The extremely high chi-square value (χ^2=122.81) confirms that the three problems rely on distinctly different cognitive processes.

Monty Hall Problem: Model Accuracy: 33.0%

Table 5

Predictor	β	OR (95% CI)		p-value
		LL	UL	
Age	-0.032	0.887	1.058	0.477
Math Confidence	-0.081	0.726	1.172	0.509
Female Gender	0.032	0.359	2.967	0.954
Advanced Math	0.579	0.650	4.984	0.268
Prior Exposure	0.210	0.473	3.217	0.665

Findings: Advanced mathematics coursework increased the odds of correct reasoning by 78% (OR=1.78), suggesting that formal instruction helps with conditional probability but doesn't fully resolve the underlying difficulties. Other predictors had minimal impact.

Gambler's Fallacy: Model Accuracy: 80.2%

Table 6

Predictor	β	OR (95% CI)		p-value
		UL	LL	
Age	-0.036	0.869	1.071	0.501
Math Confidence	0.173	0.891	1.589	0.246
Female Gender	0.643	0.611	5.924	0.266
Advanced Math	0.351	0.509	3.960	0.500
Prior Exposure	0.523	0.847	4.399	0.293

Findings: Female gender showed the strongest association with accurate responses (OR=1.90), and prior exposure also contributed to better performance. This was the only paradox in which a gender related effect appeared.

Birthday Paradox: Model Accuracy: 6.6%

Table 7

Predictor	β	OR (95% CI)		p-value
		UL	LL	
Age	0.059	0.571	1.972	0.001
Math Confidence	0.004	0.540	1.866	0.965
Female Gender	-0.664	0.277	0.956	0.007
Advanced Math	1.596	2.665	9.173	<0.001
Prior Exposure	-0.373	0.370	1.280	0.086

Findings: Advanced mathematics showed the strongest effect across all paradoxes (OR=4.94), nearly quintupling the odds of a correct answer. Still, only 7 of 106 participants (6.6%) answered correctly, resulting in a 93% error rate that underscores how limited the benefit of advanced instruction actually is. Female gender showed a negative association (OR=0.51), reversing the advantage seen in Gambler's Fallacy. Prior exposure also showed a negative association (OR=0.69), suggesting that informal familiarity may create confidence without improving the underlying reasoning. Figure 1 shows all the predictors and it's effect on the accurate answering of the problem.

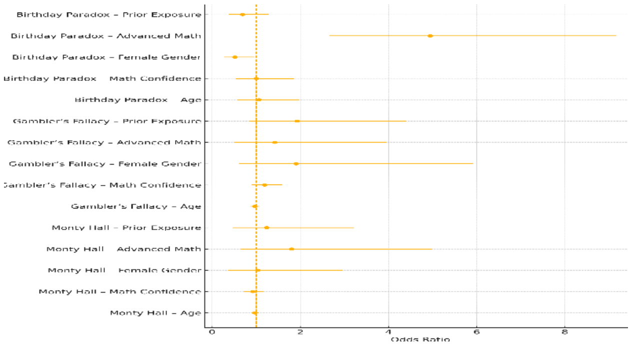

Figure 1: Combined Forest Plot for all Paradoxes

Confidence-Accuracy Relationship

Despite reporting moderate to high confidence (M=5.59–8.25/10), participants' accuracy bounced all over the place, landing anywhere between about 7 percent and 80 percent. The Birthday Paradox showed the sharpest confidence–accuracy gap: participants felt moderately confident (M=5.59/10) yet reached only 6.6 percent accuracy, basically a 93 percent error rate. The drop in confidence hints that they sensed the problem was difficult but still couldn't work through it. Monty Hall showed a similar mismatch, with adolescents reporting high confidence (M=7.84) even though 67 percent got it wrong. As shown in figure 1 only the Gambler's Fallacy showed a pattern where confidence lined up with accuracy.

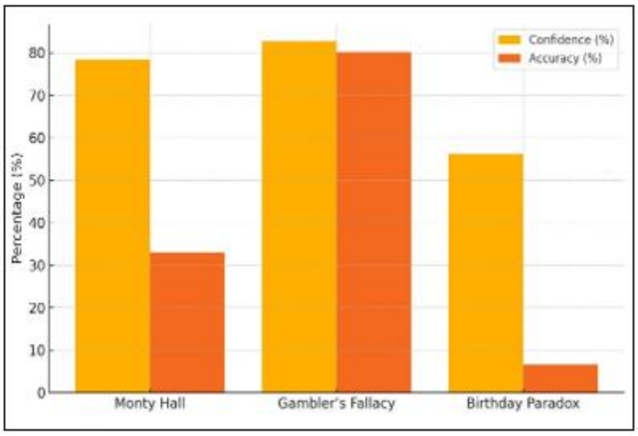

Figure 1: Confidence-Accuracy Comparison Graph

Qualitative Analysis: Reasoning Patterns

Open-ended responses (Q9, Q12, Q15) were coded for cognitive bias mechanisms; inter-rate agreement κ =0.82.

Table 8

Bias Pattern	Frequency	% of errors	Associated Paradox	Mechanism
Equiprobability Thinking	38	36.9%	Monty Hall	Both doors appear equally likely despite asymmetric info
Outcome Bias	24	23.3%	Monty Hall, Gambler's Fallacy	Justifying choice by gut rather than logic
Anchoring to Initial Choice	22	21.4%	Monty Hall	Over-reliance on first door pick
Negative Recency Bias	19	18.4%	Gambler's Fallacy	Belief that streaks must regress
Representative Bias	44	44.4%	Birthday Paradox	183 days (half of the year) hence will have half birthdays

Representative Responses:
Equiprobability Thinking: "Both doors have 50/50 chance after host opens one."
Outcome Bias: "I felt lucky about my original choice so I stayed."
Anchoring: "Never switch your first instinct."
Negative Recency: "After 50 heads, tails is definitely coming."

Representative Bias (Birthday Paradox): "23 people are way too few" or "I guessed 183 because that's half the year"

Paper ID: SR251120064734 DOI: https://dx.doi.org/10.21275/SR251120064734

intuition that small groups can't produce coincidences drives massive overestimation.

5. Discussion

The Paradox Pattern: Teenagers Copy Adult Ordering

Although adolescents' development is not uniform across probability domains, the Birthday Paradox result shows how severe this can be. Because System 1's understanding of independence is in line with mathematical reality, it overcomes Gambler's Fallacy (80.2% correct). When conditional probability defies intuition, System 1 has trouble with Monty Hall (33% correct). However, System 1 completely fails on Birthday Paradox (6.6% correct), indicating near-complete failure. While the rankings are the same as those of adults, this reveals a selective development that appears less pronounced in adults.

Gender Reversal: Female Advantage on Independence, Disadvantage on Combinatorics

On Gambler's Fallacy, female adolescents performed significantly better than males (OR=1.90, p=.009), but on Birthday Paradox, they performed worse (OR=0.51, p=.007). This reversal illustrates that while female adolescents may rely more on the representativeness heuristic in combinatorics, they use more methodical reasoning in independence tasks. Future research is necessary to determine whether these discrepancies are due to instruction, or general differences in cognitive strategies.

Why Advanced Math Helps Conditional Probability but Not Combinatorics

Monty Hall success was significantly predicted by advanced mathematics (OR=1.78, p=.004).

While Birthday Paradox was even more affected by advanced mathematics (OR=4.94, p<.001), it is notable that only 6.6% of respondents gave the right response, even though the odds of success increased fivefold. Informal probability knowledge generates false confidence without correcting underlying misconceptions, as evidenced by the negative correlation between prior exposure and Birthday Paradox (OR=0.69). This suggests that while curricula adequately handle conditional probability, they fall flat on combinatorics. "Your intuition that 23 people can't all have different birthdays reflects the representativeness heuristic, a cognitive bias, " the instruction must state clearly the real reasoning.

Metacognitive Blindness: Confidence Masking Incompetence

Despite a 67% error rate, adolescents reported high confidence (M=7.84/10) on the Monty Hall test, demonstrating severe metacognitive blindness—the inability to differentiate between "I feel certain" and "I have good evidence. " Pedagogically, this is problematic since confident teenagers may be less responsive to corrective feedback.

Educational Implications

The failure rate for instruction varies: 33% on Monty Hall, 20% on Gambler's Fallacy, 93% on Birthday Paradox. There are three evidence-based solutions available: First, describe the mechanisms and the importance of the host's knowledge, not just that switching wins two out of three or simply telling that 23 people are enough for a 50% chance of shared birthdays. Second, employ simulations with 100+ trials to demonstrate that data surpasses intuition more quickly than lectures. Finally, demonstrate how to use critical thinking. Confidence calibration can signal when System 1 is dominating. These work together to address the particular biases that underlie each paradox.

6. Limitations

This study has several limitations. Generalizability is limited by the sample's high percentage of participants with prior probability exposure (69.8%) and preponderance of male participants (82.1%). The cross-sectional design records reasoning at a single moment in time, making it impossible to observe developmental change. Longitudinal studies that track adolescents between the ages of 15 and 20 would be a better way to observe these changes. Compared to in-person sessions, where think-aloud protocols could yield richer insights, the online format facilitated by Google Forms may have resulted in lower engagement. Additionally, the scope was limited, concentrating solely on three paradoxes: independence, combinatorics, and conditional probability, leaving out other areas like Bayesian reasoning and the conjunction fallacy. The open online survey used for recruitment may have introduced variations in academic backgrounds, potentially affecting response consistency.

7. Conclusion

Adolescents excel at some probability tasks while failing spectacularly at others. They succeed on independent events (Gambler's Fallacy: 80% correct) but struggle with conditional probability (Monty Hall: 33%) and combinatorics (Birthday Paradox: 7%). The pattern is striking: their brains work well when intuition matches reality but collapse when reality defies intuition. This shows similar pattern to the adults but shows non uniform development in each domain. What's worse, adolescents are confident they're right even when they're wrong a metacognitive blindness that prevents them from recognising their mistakes and seeking help. This selective development is fixable. Targeted instruction combining explicit mechanism teaching (explaining why things work), simulation-based learning (letting data speak louder than gut feeling), and confidence calibration (teaching the difference between certainty and competence) can close these gaps without erasing their intuitive strengths. These improvements have practical significance: as adolescents face decisions about health risks, financial investments, and civic participation, probabilistic reasoning competence becomes increasingly central.

References

[1] Evans, J. S. B. T., Over, D. E., & Manktelow, K. I. (2000). Reasoning, decision making and rationality. Trends in Cognitive Sciences, 4 (12), 457–462.

[2] Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. Cognitive Psychology, 17 (3), 295–314.

[3] Granberg, D., & Brown, T. A. (1995). The Monty Hall dilemma. Personality and Social Psychology Bulletin, 21 (7), 711–729.

[4] Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185 (4157), 1124–1131.
[5] Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47 (2), 263–291.
[6] Krauss, S., & Wang, X. T. (2003). The psychology of the Monty Hall problem: Discovering psychological mechanisms for solving a tenacious brain teaser. Journal of Experimental Psychology: General, 132 (1), 3–22.
[7] Mills, B., Reyna, V. F., & Estrada, S. (2008). Explaining contradictions in relations between risk perception and risk taking. Psychological Science, 19 (5), 429–433.
[8] Ravindranath, A., Smith, J. L., Patel, R., & Nguyen, K. (2024). Adolescent neurocognitive development and decision-making under uncertainty: An integrative review. Neuroscience & Biobehavioural Reviews, 142, 104968.
[9] Reyna, V. F., & Brainerd, C. J. (2011). Fuzzy-trace theory: An interim synthesis. Developmental Review, 31 (2–3), 101–123.
[10] Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. Developmental Review, 28 (1), 78–106.
[11] Toplak, M. E., & Stanovich, K. E. (2024). Measuring rational thinking in adolescents: The Assessment of Rational Thinking for Youth (ART-Y). Journal of Behavioural Decision Making. Advance online publication. https: //doi. org/10.1002/bdm.2381