Title

Effects of encouraging vs. reflective feedback on children's explore-exploit decisions

Permalink

https://escholarship.org/uc/item/3xc356bq

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Fang, Yijin

Chang, Ariel

Christie, Stella

Publication Date

2026

Effects of encouraging vs. reflective feedback on children's explore-exploit decisions

Yijin Fang[1] **(fangyj22@mails.tsinghua.edu.cn)**
Ariel Chang[1] **(ariel.chang8@gmail.com)**
Stella Christie[1,2] **(christie@mail.tsinghua.edu.cn)**
[1]Department of Psychological and Cognitive Sciences, Tsinghua University, Beijing, China
[2]Ministry of Higher Education, Science, and Techonology, Republic of Indonesia

Abstract

Explore-exploit decisions are often made without accurate awareness of one's own behavioral tendencies, limiting effective metacognitive regulation. Feedback has the potential to improve metacognitive calibration, but its effectiveness depends on its informative value. This study examined how encouraging and reflective feedback–differing in informative value–influence children's explore-exploit decisions, while also considering the role of feedback frequency. Results showed that high-frequency reflective feedback reduced children's exploration but led to higher reward performance, suggesting the development of more adaptive search strategies. In contrast, children exhibited a more exploratory search pattern under low-frequent reflective feedback. Although encouraging feedback provided relatively low informative value, both low- and high-frequency delivery sustained high exploration, potentially by fostering emotional safety. Overall, these findings offer practical implications for education, highlighting the importance of strategically balancing affective reassurance and metacognitive control to support children's effective learning and exploration.

Keywords: explore-exploit decisions; feedback; encouragement, reflection, metacognitive regulation

Introduction

Decisions between exploration and exploitation emerge early and apply widely across diverse domains of daily life (Kim and Carlson, 2024; Harms et al., 2024; Hills et al., 2015), requiring individuals to choose between exploring unknown options versus exploiting familiar ones (Cohen et al., 2007; Wilson et al., 2021; Mehlhorn et al., 2015). For instance, young children not only know to directly take their favorite candies, but also are eager to pick out a new flavor that they've never tried before. On the one hand, more exploitative options tend to lead to a steady reward level but may result in the agent being stuck in a "good enough" state due to the more constrained nature of such options. On the other hand, a broader and more exploratory search strategy may increase the potential for achieving a better outcome but can also create higher uncertainty at the same time (Stojić et al., 2020; Wu et al., 2018; Wilson et al., 2014).

In explore–exploit decision-making contexts aimed at reward maximization, individuals typically begin with exploration and gradually shift toward exploitation as knowledge about the context accumulates (Wilson et al., 2014; Blanco and Sloutsky, 2021). This initial exploratory phase functions as a learning period, during which individuals form expectations about potential outcomes (Wu et al., 2018; Schulz et al., 2019). The reduced uncertainty–whether through truly accurate or biased beliefs–ultimately enables individuals to drive the explore-to-exploit shift (Blanco and Sloutsky, 2021; Liquin and Gopnik, 2022). This general tendency has been observed in both adults and children, although the explore-to-exploit shift tends to be more pronounced and systematic in adults (Kim and Carlson, 2024; Schulz et al., 2019; Meder et al., 2021).

Yet, in most cases, individuals usually stay with an "unskilled and unaware" state, where the metacognitive regulation deficits or biases, classically captured by the Dunning-Kruger effect (Kruger and Dunning, 1999). This posits that less competent individuals substantially overestimate their abilities or become confidence, whereas more competent individuals tend to moderate- or even under-estimate their own competence (Dunning, 2011). A substantial body of evidence indicates that metacognitive inaccuracy—particularly overconfidence—often undermines learning motivation and impairs achievement across both adult and child populations (Dunlosky and Rawson, 2012; Metcalfe et al., 2010; Finn and Metcalfe, 2014).

A promising approach to improving individual's self-evaluation accuracy is to provide feedback within rich social contexts (Hattie and Timperley, 2007; Miller and Geraci, 2011). Effective feedback–such as performance feedback–clarifies the criteria for a fully correct response and explicitly supports calibration and self-regulation of their actual performance (Raaijmakers et al., 2017; Van Loon and Roebers, 2017; Grinschgl et al., 2021; Venables and Fairclough, 2009). However, not all forms of explicit feedback are equally beneficial for learning or exploration behavior (Mandouit and Hattie, 2023). A comprehensive meta-analysis indicated that feedback effectiveness increases substantially with informational richness: low-information feedback–such as simple praises–yields minimal or even negative effects for learning motivation; whereas high-information feedback that explains errors, reveals underlying cognitive processes, and offers strategies for improvement produces robust benefits (Wisniewski et al., 2020).

Particularly, in the feedback model proposed by Hattie and Timperley (2007), four levels of feedback are distinguished. Within this framework, "self-level" feedback–such as general praise–is typically the least effective, as it provides lit-

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

tle actionable information about task performance, despite offering emotional support. In contrast, feedback operating at the “self-regulation-level” is considerably more powerful. Rather than supplying correct answers or direct corrective information, this form of feedback encourages learners to engage in self-monitoring, adaptive evaluation, and error detection. In the present study, we focus on comparing the effects of self-level feedback (i.e., encouraging feedback) and self-regulation-level feedback (i.e., reflective feedback) on children’s explore–exploit decisions.

Beyond feedback type, the frequency of feedback delivery may also play a crucial role in shaping children’s decisions and behaviors (Petranek et al., 2019; Chhokar and Wallin, 1984). Evidence showed that learning is not enhanced by increasingly frequent feedback; rather, a moderate level of feedback frequency yields the greater benefits (Lowe and Buchwald, 2017). Consistent with self-determination theory, overly frequent feedback may undermine intrinsic motivation by shifting the perceived locus of control outward, thereby discouraging active learning and exploration (Deci and Ryan, 1985). Moreover, the effects of feedback frequency are conditional on task characteristics, such that when task difficulty is low, less frequent feedback may be sufficient, whereas more challenging tasks may benefit from more frequent feedback (Sidaway et al., 2012). Together, these findings indicate that feedback frequency, alongside feedback type, jointly contributes to the regulation of children’s explore–exploit decisions.

Overall, the present study aims to investigate how feedback type (encouraging vs. reflective) and feedback frequency (low vs. high) jointly influence children’s explore-exploit decisions using a 2×2 design. Grounded in a metacognitive calibration framework, we particularly conceptualize reflective feedback as a self-regulatory mechanism that elicits metacognitive regulation on one’s knowledge state. By contrasting encouraging feedback with reflective feedback across different frequencies, it can help us understand when and how feedback influences children decision-making in explore-exploit trade-off contexts. We used a modified version of a spatially correlated multi-armed bandit task (Wu et al., 2017), in which children searched for and collected rewards over multiple rounds. In each round, children faced a limited number of opportunities and had to choose between exploration, by selecting new tiles with unknown rewards, and exploitation, by selecting previously chosen tiles with known rewards. At the same time, in each round, children received feedback from the experimenter, which was either “Great job!” (i.e., encouraging feedback) or “Are you sure?” (i.e. reflective feedback). Half of the children received feedback after every choice (i.e. high frequency), while the other half received feedback only after some of their choices (i.e. low frequency). We hypothesize that, compared to the encouraging feedback, the reflective feedback would promote a more adaptive and effective strategies in explore-exploit decisions, due to the metacognitive calibration. Although encouragement tend to be regarded as a low-information feedback, it can provide affective support and a sense of safety. As a result, even if encouraging feedback does not directly enhance strategy efficiency and task performance, it may foster more intensive exploration preferences under the uncertainty. By examining these questions, this research advances our understanding of how the social-contextual factor–explicit feedback–regulate children’s preferences in explore-exploit decisions.

Methods

Participants

71 6- to 8-year-old children (*M* = 6.94 years, *SD* = 0.63; 37 girls) were recruited. We obtained informed consent from the children’s parents, and all children provided verbal assent before starting the experiment. Children were randomly assigned to one of totally four experimental conditions. Detailed children’s distribution in four conditions was showed in Table 1.

Table 1: Children’s demographic information in 4 conditions

Feedback(type×freq)	Boys(Girls)	Mean age(SD)
Encouraging×High	9(9)	6.86(0.56)
Encouraging×Low	7(10)	7.09(0.58)
Reflective×High	10(6)	6.96(0.73)
Reflective×Low	9(11)	6.87(0.68)

Design

The experiment employed a 2×2 between-subjects design, manipulating feedback type (encouraging vs. reflective) and feedback frequency (high vs. low). Feedback type determined the content of the experimenter’s verbal responses to children’s choices, whereas feedback frequency determined how often feedback was delivered during the task.

feedback instruction

Feedback was delivered verbally by the experimenter and followed standardized scripts to ensure consistency across participants. In the encouraging feedback condition, children received affective praise aimed at providing emotional support, such as “Great job!”, “Well done!”, and “Keep going!”. In contrast, in the reflective feedback condition, children received prompts designed to encourage self-monitoring and evaluation of their decisions, including “Are you sure?”, “Is this the one?”, and “Why did you make this choice?”. Feedback was administered in accordance with the assigned feedback frequency condition (high vs. low).

Task

The explore-exploit task contains a clicking game for children to search and collect rewards, which was revised from a spatially correlated multi-armed bandit paradigm originally presented by Wu et al. (2017). In our experiment, we present children with a total of 64 tiles in a 8×8 grid world (see Figure 1). Each tile contains a certain number of rewards, with the reward ranging from 1 to 32. The reward of a tile is assigned following a "smooth" rule, which states that nearby tiles have similar rewards. The tiles have different colors depending on the reward they contain, with tiles containing greater rewards having darker colors. The rewards of tiles were hidden from children initially, with all the tiles displaying a gray color. Children can interact with this grid world by clicking on one tile at a time, revealing and collecting the reward of that tile. When a tile is clicked on for the first time, its color changes from gray to its assigned color, which signifies its reward. The tiles that have already been clicked on will remain in their assigned color and can be clicked on again. After every click, children receive information about the amount of reward they received in this click, the total amount of reward they have received, and the number of clicking opportunities left for this round. Children were told that they needed to engage in 3 rounds of the game, aiming to maximize their total rewards within 25 clicks per round. The distribution of rewards is different in these 3 rounds. In this task, children can click on a new tile (exploration) or re-click an already-revealed one (exploitation), representing two distinct options in the explore-exploit decision process. Children's behavioral patterns, as well as the total rewards they earned, are assessed.

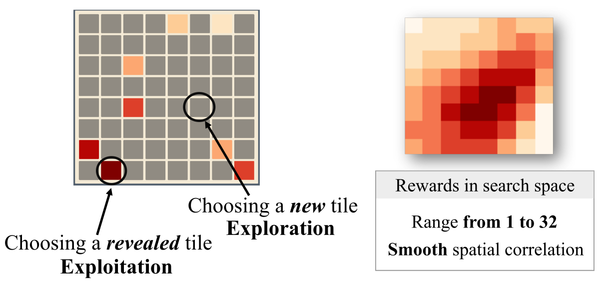

Figure 1: Explore-exploit task design. A 8×8 grid world where children can search and collect rewards. Darker-colored tiles have greater rewards. Children could choose a new tile (exploration choice) or choose a revealed one (exploitation choice). Choosing a new tile changes the tile to its assigned color. Possible reward of a tile ranges from 1-32, and the grid world has a smooth spatial correlation.

Procedure

Instructional phase. The experimenter explained the goal and rules of the task to the children. The experimenter made two emphases: (1) darker-colored tiles contain higher rewards, and (2) any tile could be re-clicked multiple times. Then, to help children better understand the rules, the experimenter gave simple demonstrations correspondingly: (1) The experimenter showed a picture that four revealed tiles with different colors in the whole grid world and told children that these were several example tiles with their respective colors and rewards. (2) The experimenter demonstrated the behavior of repeatedly clicking on a tile and told children that this behavior was allowed but not mandatory. Finally, children were told that their goal was to accumulate as many total rewards as possible in a total of 4 rounds, with each round offering 25 clicking opportunities.

At the end of the instructional phase, children were asked to repeat the goal of the task and the rules they were told, with a particular focus on the two key rules. If the child was unable to correctly repeat the rules, the experimenter would explain to the child one more time. After that, if the child still could not repeat the rules, the experimenter would take note of that. All children were allowed to continue to the next phase.

Practice phase. After the instructional phase, children play one round of the game as practice. After the practice, children complete a comprehension test to confirm that they have truly understood the rules. This test includes 2 questions assessing children's understanding of the relationship between the shades of the colors and the number of rewards, as well as 2 questions evaluating their understanding of the rule regarding whether re-clicking one revealed tile is allowed. Only children who answer all the questions correctly are valid and included in the final analysis.

Play phase with feedback. After the practice phase and the comprehension test, children enter the formal play phase, containing a total of 4 rounds. Children assigned to the encouraging feedback condition received affective, non-informative feedback (e.g., "Great job!"), whereas children in the reflective feedback condition received feedback designed to prompt self-monitoring and reflection (e.g., "Are you sure?"). Feedback content was consistent across rounds and independent of children's true choices. Additionally, in the high-frequency feedback condition, children received feedback after every choice. In the low-frequency feedback condition, children received feedback after a predetermined subset of choices, specifically once every five choices, for a total of five feedback times per round.

Post-task phase. After the children finished all 4 rounds of the game, we included measures to approximate children's explore–exploit preferences in a new, non-performance-oriented context. Children were asked (3) whether they would prefer to play the same, a similar, or a different game (i.e., their willingness to play), and to make a forced choice among these three game types. Unlike our explore-exploit task, which required performance-oriented decision-making, this measure assessed children's preferences for familiarity versus novelty in the absence of performance goals. All these questions allowed us to have a better understanding of children's motivation and preference for explore-exploit decisions.

Results

Children explore least when they received reflective feedback with a high frequency

We first examined children's exploration ratio in each round. The exploration ratio is defined as the percentage of chosen tiles that are new tiles (number of exploration choices / 25). We fitted a linear mixed-effects model predicting children's exploration ratio from type, frequency, and round, including subject as a random intercept. The descriptive statistics showed that children who received high-frequent reflective feedback explored least (M = 77.88%, SD = 0.26) than another three conditions (high-frequent encouraging: M = 83.72%, SD = 0.23; low-frequent encouraging: M = 86.00%, SD = 16.63; low-frequent reflective: M = 89.30%, SD = 12.73). Particularly, in the model, there was a significant three-way interaction between type, frequency, and round: the tendency of lowest exploration in high-frequent reflective feedback group was obvious in later rounds (round 3: β = -0.26, 95% CI [-0.40, -0.13], $t(266) = -3.83$, $p < 0.001$; round 4: $\beta = -0.14$, 95% CI [-0.28, -0.01], $t(266) = -2.08$, $p = 0.039$; see Figure 2A and B). These results suggested that the combined effect of type and frequency on children's exploration ratio changed across rounds, emerging primarily in the later stages of the task. That is, within rounds, children gradually reduced their exploration when they received high-frequently reflective feedback.

Children earned more rewards in the condition of the high-frequent reflective feedback

Two-factor ANOVA analysis showed the main effect of feedback frequency on reward performance, $F(1,280) = 5.02$, $p = 0.026$, $\eta^2 = 0.03$, and an interactive effect between feedback type and frequency, ($F(1,280) = 3.87$, $p = 0.050$, $\eta^2 = 0.01$. Children who received a high-frequent reflective feedback earned the most reward in a round (M = 520.33, SD = 118.92) than another 3 conditions (high-frequent encouraging: M = 502.76, SD = 91.38; low-frequent encouraging: M = 500.00, SD = 90.87; low-frequent reflective: M = 475.30, SD = 55.37, see Figure 2C). No differences in children's reward performance by rounds were observed from the generalized linear regression model with random subject effects across 4 conditions within 4 rounds (see Figure 2D).

Reflective feedback with low-frequency delivery promoted distant search pattern

To understand children's exploration patterns in the four conditions in more detail, we calculated the distance between each pair of consecutive choices and classified each choice into one of three categories: repeat, near, and far. Repeat choices: the child clicked the exact same tile as in the previous click (distance = 0). Near choices: the child clicked a tile in the immediately adjacent neighborhood of the previous click (distance = 1 and $\sqrt{2}$). Far choices: the child clicked a more distant tile (distance $> \sqrt{2}$). For repeat choices, the pattern of repeat choice selection mirrored those observed for exploration choices, as the ratio of repeat choices was operated as the inverse of exploration (i.e., 1-exploration ratio).

We then fitted a multinomial logistic regression model predicting choice type (repeat, near, far), with repeat choices serving as the reference category. Predictors included feedback type (encouraging vs. reflective), feedback frequency (low vs. high), their interaction, and round as a covariate. Encouraging feedback delivered at low frequency served as the baseline condition. The results showed that children were

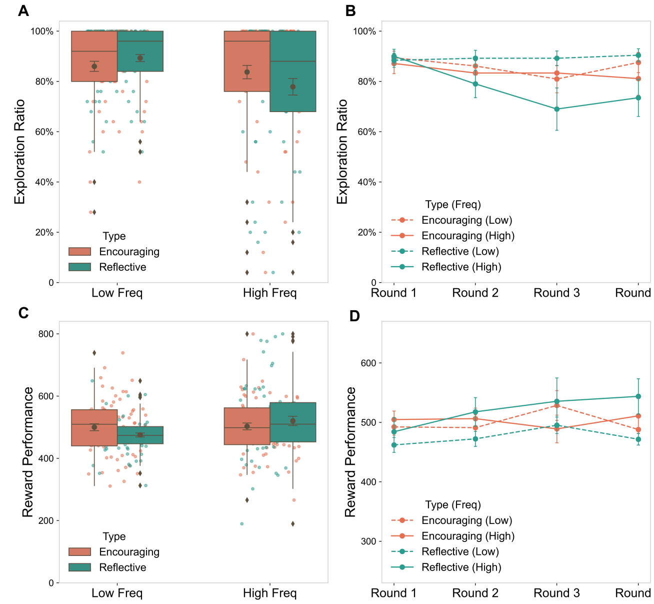

Figure 2: (A) The average exploration ratio across 4 feedback conditions: children explore least when they received reflective feedback with a high frequency. (B) The exploration ratio across 4 rounds in 4 conditions. (C) The average reward performance across 4 feedback conditions: children earned no significantly different rewards across 4 conditions. (D) The reward performance across 4 rounds in 4 conditions.

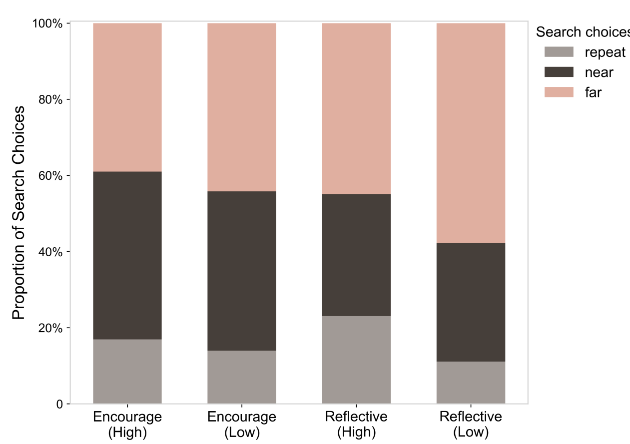

Figure 3: Proportions of three search choice categories (repeat, near, far) across the four feedback conditions. Children were more likely to select near and far choices in encouraging feedback. In contrast, the effects of reflective feedback on children's search choices more depended on feedback frequency–particularly reflective feedback delivered at low frequency was associated with an increased likelihood of far choice selection.

more likely to select near choices than repeat choices under the encouraging feedback (OR = 3.36, 95% CI [2.69, 4.20], $p < 0.001$). For far choices, children who received encouraging feedback also showed a strong tendency toward far choices relative to repeat choices (OR = 5.69, 95% CI [4.58, 7.07], $p < 0.001$). Compared with encouraging feedback, reflective feedback significantly increased the odds of selecting far choices (OR = 1.66, 95% CI [1.35, 2.04], $p < 0.001$). In contrast, the facilitative effect of reflective feedback on far exploration was significantly decreased when feedback was delivered frequently (OR = 0.51, 95% CI [0.39, 0.68], $p < 0.001$; see Figure 3). Finally, far choices declined across rounds (OR = 0.79, 95% CI [0.74, 0.84], $p < 0.001$), indicating a shift from exploration to exploitation over time.

Taken together, these findings demonstrated that the impact of feedback type, especially reflective feedback, on children's search choices is contingent on feedback frequency. Specifically, reflective feedback delivered at low frequency exhibited a more distant search pattern, as evidenced by a marked increase in far choices.

Children continued to preferentially explore non-repeated games

To examine how feedback type and feedback frequency influenced children's willingness to engage with different game options, we fitted a linear model predicting willingness ratings from feedback type, feedback frequency, and game category (same, similar, different), including all interactions. Across feedback conditions, children reported significantly lower willingness to play a different game, β = -0.67, 95% CI [-1.31, -0.03], p = 0.041 (see Figure 4A). These results indicate that children consistently showed reduced willingness to engage in the novel game options relative to familiar ones. Neither feedback type nor feedback frequency had significant main effects on willingness, and none of the interactions reached significance, suggesting that the observed preference pattern was stable across feedback conditions.

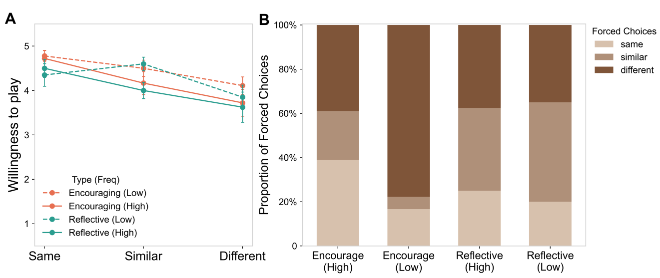

Figure 4: (A) Children's willingness to play games from 3 categories (same, similar, different) across 4 feedback conditions. Across conditions, children consistently showed the lowest willingness to play a novel (different) game. (B) Proportions of children's forced-choice selections among the 3 game categories across the 4 feedback conditions.

However, children's forced-choice selections among the three game categories differing in novelty revealed an interestingly distinctive pattern across the four feedback conditions. Firstly, children in all conditions showed a very low percentage (only 25.35% on average) of choosing the same game to play again. Moreover, children who received high-frequency encouraging feedback were most likely to choose the same game again, whereas those receiving low-frequency feedback were most likely to select a novel game (see Figure 4B). However, a Pearson chi-square test indicated only a marginal association between feedback condition and forced-choice selections, $\chi^2(6) = 12.01, p = 0.062, V = 0.21$. Given the absence of a strong statistically significance, these patterns should be interpreted cautiously. Overall, the results suggest that feedback type and frequency might potentially influence children's exploration preference among things differing in novelty.

General Discussion

How do feedback type (encouraging vs. reflective) and feedback frequency (low vs. high) jointly influence children's explore-exploit decisions? Our results suggested that, children who received reflective feedback with a high-frequency delivery decreased their exploration preference, whereas they earned relatively greater rewards. Under the strongly extensive reflection delivered by the high-frequent reflective feedback, children developed a more adaptive and efficient exploration strategies, achieving a better performance. In line with the "self-regulation level" feedback in Hattie and Timperley (2007)'s model–which emphasizes reflection and promotes metacognitive regulation, reflective feedback helps learners evaluate their choices more deliberately and attend more closely to the consequences of their actions (Nederhand et al., 2020; Hacker et al., 2008).

Notably, when receiving a low-frequent reflective feedback, children tended to develop a most broad search pattern that chose high percentage of far options rather than repeat or near options, which refers to a more extensive exploration tendency. This results suggested that occasional reflective feedback may not strongly heighten individuals' attention to performance and action efficacy, nor does it necessarily induce self-doubt or deeper reflection on the function of one's own decisions. This interpretation, however, remains speculative and warrants further empirical validation.

Under encouraging feedback, either with a low- or high-frequency delivery, children generally kept a high-exploratory pattern, with making less exploitative choices and more near and far choices. In Hattie and Timperley (2007)'s model, encouraging feedback was more likely to be defined as the "self-level" feedback. In this view, although praise tends to be a low-information form of feedback and does not necessarily enhance learning motivation and effectiveness, it serves as a source of affective and encouraging support. Drawing from attachment theory, the secure base hypothesis proposes that when individuals feel supported and safe, they become more willing to exploration into uncertainty (Bowlby, 2012).

By buffering against the fear of negative outcomes, encouraging feedback may tilt children toward exploration, enabling them to search for uncertain options without immediate rewards.

Interestingly, our results showed that children preferred to continue exploring novel games despite reporting a decreased willingness to play them. Specifically, In the forced-choice test, children who received highly frequent encouraging feedback more preferred to replay the same game they had just played, compared to those in another conditions. One possible explanation is that abundant encouragement fostered strong feelings of confidence and self-efficacy. This form of emotional support may have temporarily led children to develop an affective reliance on, and preference for, familiar activities (Henderlong and Lepper, 2002; Deci et al., 1999). In contrast, children who received a lower level of encouraging feedback showed the highest proportion of choices for completely novel games. It is possible that only a moderate level of encouragement elicits positive emotional experiences that effectively support exploratory motivation, as such encouragement may be perceived as more sincere and credible (Henderlong and Lepper, 2002; Mueller and Dweck, 1998).

Taken together, these findings demonstrate that children's explore–exploit decisions are jointly shaped by both the type and frequency of feedback, revealing the role of affective support and metacognitive regulation on one's exploration preference and performance. High-frequency reflective feedback shifted children toward less exploratory but more efficient strategies that yielded higher rewards, consistent with self-regulation–level feedback that promotes deliberate evaluation and adaptive control. In contrast, low-frequency reflective feedback preserved broad exploratory search, suggesting that occasional reflection may support exploration without constraining behavior through strongly performance monitoring. Although containing low-information, encouraging feedback with different frequencies generally sustained high exploratory tendencies, likely by providing emotional safety. However, excessive encouragement appeared to foster a short-term affective reliance on familiar activities, whereas moderate encouragement more strongly supported continued exploration of novel options. Overall, this work contributes to cognitive science by demonstrating how external feedback structures shape children's adaptive search strategies under uncertainty. It also offers practical implications for educational practice, suggesting that educators should strategically balance affective reassurance and metacognitive control, and calibrate the frequency of feedback, to support both effective learning and exploratory motivation.

Acknowledgments

We would like to thank the Tsinghua-Feihe Joint Research Project on "Brain Development and Cognition" for funding this research. We also thank all the children and parents who participated in and supported our research.

References

Blanco, N. J. and Sloutsky, V. M. (2021). Systematic exploration and uncertainty dominate young children's choices. *Developmental Science*, 24(2):e13026.

Bowlby, J. (2012). *A Secure Base*. Routledge, 0 edition.

Chhokar, J. S. and Wallin, J. A. (1984). A field study of the effect of feedback frequency on performance. *Journal of Applied Psychology*, 69(3):524–530.

Cohen, J. D., McClure, S. M., and Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481):933–942.

Deci, E. L., Koestner, R., and Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6):627–668.

Deci, E. L. and Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2):109–134.

Dunlosky, J. and Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4):271–280.

Dunning, D. (2011). The Dunning–Kruger Effect. In *Advances in Experimental Social Psychology*, volume 44, pages 247–296. Elsevier.

Finn, B. and Metcalfe, J. (2014). Overconfidence in children's multi-trial judgments of learning. *Learning and Instruction*, 32:1–9.

Grinschgl, S., Meyerhoff, H. S., Schwan, S., and Papenmeier, F. (2021). From metacognitive beliefs to strategy selection: does fake performance feedback influence cognitive offloading? *Psychological Research*, 85(7):2654–2666.

Hacker, D. J., Bol, L., and Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: the effects of incentives, reflection, and explanatory style. *Metacognition and Learning*, 3(2):101–121.

Harms, M. B., Xu, Y., Green, C. S., Woodard, K., Wilson, R., and Pollak, S. D. (2024). The structure and development of explore-exploit decision making. *Cognitive Psychology*, 150:101650.

Hattie, J. and Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1):81–112.

Henderlong, J. and Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, 128(5):774–795.

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., and Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, 19(1):46–54.

Kim, S. and Carlson, S. M. (2024). Understanding explore-exploit dynamics in child development: current insights and future directions. *Frontiers in Developmental Psychology*, 2:1467880.

Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6):1121–1134.

Liquin, E. G. and Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, 218:104940.

Lowe, M. S. and Buchwald, A. (2017). The Impact of Feedback Frequency on Performance in a Novel Speech Motor Learning Task. *Journal of Speech, Language, and Hearing Research*, 60(6S):1712–1725.

Mandouit, L. and Hattie, J. (2023). Revisiting "The Power of Feedback" from the perspective of the learner. *Learning and Instruction*, 84:101718.

Meder, B., Wu, C. M., Schulz, E., and Ruggeri, A. (2021). Development of directed and random exploration in children. *Developmental Science*, 24(4):e13095.

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., Hausmann, D., Fiedler, K., and Gonzalez, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, 2(3):191–215.

Metcalfe, J., Eich, T. S., and Castel, A. D. (2010). Metacognition of agency across the lifespan. *Cognition*, 116(2):267–282.

Miller, T. M. and Geraci, L. (2011). Training metacognition in the classroom: the influence of incentives and feedback on exam predictions. *Metacognition and Learning*, 6(3):303–314.

Mueller, C. M. and Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75(1):33–52.

Nederhand, M. L., Tabbers, H. K., Jongerling, J., and Rikers, R. M. J. P. (2020). Reflection on exam grades to improve calibration of secondary school students: a longitudinal study. *Metacognition and Learning*, 15(3):291–317.

Petranek, L. J., Bolter, N. D., and Bell, K. (2019). Attentional Focus and Feedback Frequency Among First Graders in Physical Education. *Journal of Teaching in Physical Education*, 38(3):199–206.

Raaijmakers, S. F., Baars, M., Schaap, L., Paas, F., and Van Gog, T. (2017). Effects of performance feedback valence on perceptions of invested mental effort. *Learning and Instruction*, 51:36–46.

Schulz, E., Wu, C. M., Ruggeri, A., and Meder, B. (2019). Searching for Rewards Like a Child Means Less Generalization and More Directed Exploration. *Psychological Science*, 30(11):1561–1572.

Sidaway, B., Bates, J., Occhiogrosso, B., Schlagenhaufer, J., and Wilkes, D. (2012). Interaction of Feedback Frequency and Task Difficulty in Children's Motor Skill Learning. *Physical Therapy*, 92(7):948–957.

Stojić, H., Schulz, E., P. Analytis, P., and Speekenbrink, M. (2020). It's new, but is it good? How generalization and uncertainty guide the exploration of novel options. *Journal of Experimental Psychology: General*, 149(10):1878–1907.

Van Loon, M. H. and Roebers, C. M. (2017). Effects of Feedback on Self-Evaluations and Self-Regulation in Elementary School. *Applied Cognitive Psychology*, 31(5):508–519.

Venables, L. and Fairclough, S. H. (2009). The influence of performance feedback on goal-setting and mental effort regulation. *Motivation and Emotion*, 33(1):63–74.

Wilson, R. C., Bonawitz, E., Costa, V. D., and Ebitz, R. B. (2021). Balancing exploration and exploitation with information and randomization. *Current Opinion in Behavioral Sciences*, 38:49–56.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., and Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, 143(6):2074–2081.

Wisniewski, B., Zierer, K., and Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10:3087.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., and Meder, B. (2017). Mapping the unknown: The spatially correlated multi-armed bandit.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., and Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, 2(12):915–924.