Mixing Backward- with Forward-Chaining for Metacognitive Skill Acquisition and Transfer

Mark Abdelshiheed, John Wesley Hostetter, Xi Yang,
Tiffany Barnes, and Min Chi
{mnabdels, jwhostet, yxi2, tmbarnes, mchi}@ncsu.edu

North Carolina State University, Raleigh, NC 27695, USA

Abstract. Metacognitive skills have been commonly associated with preparation for future learning in deductive domains. Many researchers have regarded *strategy-* and *time-awareness* as two metacognitive skills that address *how* and *when* to use a problem-solving strategy, respectively. It was shown that students who are both strategy- and time-aware (*StrTime*) outperformed their *nonStrTime* peers across deductive domains. In this work, students were trained on a logic tutor that supports a default forward-chaining (FC) and a backward-chaining (BC) strategy. We investigated the impact of mixing BC with FC on teaching strategy- and time-awareness for *nonStrTime* students. During the logic instruction, the experimental students (*Exp*) were provided with two BC worked examples and some problems in BC to practice *how* and *when* to use BC. Meanwhile, their control (*Ctrl*) and *StrTime* peers received no such intervention. Six weeks later, all students went through a probability tutor that only supports BC to evaluate whether the acquired metacognitive skills are transferred from logic. Our results show that on both tutors, *Exp* outperformed *Ctrl* and caught up with *StrTime*.

Keywords: Strategy Awareness · Time Awareness · Metacognitive Skill Instruction · Preparation for Future Learning · Backward Chaining.

1 Introduction

One fundamental goal of education is being prepared for future learning [6] by transferring acquired skills and problem-solving strategies across different domains. Despite the difficulty of achieving such transfer [6], prior research has shown it can be facilitated by obtaining metacognitive skills [1–3, 8]. It has been believed that metacognitive skills are essential for academic achievements [5], and teaching such skills impacts learning outcomes [8] and strategy use [13]. Much prior research has categorized knowing *how* and *when* to use a problem-solving strategy as two metacognitive skills [15], referred to as strategy- and time-awareness, respectively. Our prior work found that students who were both strategy- and time-aware —referred to as *StrTime*— outperformed their *nonStrTime* peers across deductive domains [1, 2]. In the current work, we provide interventions for the latter students to catch up with their *StrTime* peers.

Deductive domains such as logic, physics and probability usually require multiple problem-solving strategies. Two common strategies in these domains are forward-chaining (FC) and backward-chaining (BC). Early studies showed that experts often use a mixture of FC and BC to execute their strategies [12]. This work investigates the impact of mixing FC and BC on teaching strategy- and time-awareness for $nonStrTime$ students.

Our study involved two intelligent tutoring systems (ITSs): logic and probability. Students were first assigned to a logic tutor that supports FC and BC, with FC being the default, then to a probability tutor six weeks later that only supports BC. During the logic instruction, $nonStrTime$ students were split into experimental (Exp) and control ($Ctrl$) conditions. For Exp, the tutor provided two worked examples solved in BC and presented some problems in BC to practice *how* and *when* to use BC. $Ctrl$ received no such intervention as each problem was presented in FC by default with the ability to switch to BC. Our goal is to inspect whether our intervention would make Exp catch up with the golden standard —$StrTime$ students— who already have the two metacognitive skills and thus need no intervention. All students went through the same probability tutor to evaluate whether the acquired metacognitive skills are transferred from logic. Our results show that Exp outperformed $Ctrl$ and caught up with $StrTime$ on both tutors.

1.1   Metacognitive Skill Instruction

Metacognitive skills regulate one's awareness and control of their cognition [7]. Many studies have demonstrated the significance of metacognitive skills instruction on academic performance [5], learning outcomes [2, 3, 8] and regulating strategy use [13]. Schraw and Gutierrez [13] argued that metacognitive skill instruction involves feeling what is known and not known about a task. They stated that such instruction should further compare strategies according to their feasibility and familiarity from the learner's perspective. Chi and VanLehn [8] found that teaching students principle-emphasis skills closed the gap between high and low learners, not only in the domain where they were taught (probability) but also in a second domain where they were not taught (physics).

Strategy- and time-awareness have been considered metacognitive skills as they respectively address *how* and *when* to use a problem-solving strategy [5, 15]. Researchers have emphasized the role of strategy awareness in preparation for future learning [2, 4] and the impact of time awareness on planning skills and academic performance [5, 9]. Belenky and Nokes [4] showed that students who had a higher aim to master presented materials and strategies outperformed their peers on a transfer task. Fazio et al. [9] revealed that students who knew when to use each strategy to pick the largest fraction magnitude had higher mathematical proficiency than their peers. de Boer et al. [5] showed that students who knew when and why to use a given strategy exhibit long-term metacognitive knowledge that improves their academic performance. de Boer et al. emphasized that knowing *when* and *why* has the same importance as knowing *how* when it comes to strategy choice in multi-strategy domains.

1.2   Forward- and Backward-Chaining

FC and BC are two standard problem-solving strategies in deductive domains. In FC, the reasoning proceeds from the given propositions toward the target goal, whereas BC is goal-driven in that it works backward from a goal state to a given state. Substantial work has investigated the impact of FC and BC strategies in two research categories: empirical studies and post-hoc observations.

Prior empirical studies have shown the significance of FC over BC in learning physics [10] and weightlifting movements [11]. Moore and Quintero [11] compared FC and BC in teaching the clean and snatch movements to novice weight lifters. The participants showed mastery performance with the FC training but showed substantially fewer improvements in performance accuracy via the BC training. All participants mastered the movements when some BC lifts were changed to FC. Conversely, some studies reported no significant difference between the two strategies [14]. Slocum and Tiger [14] assessed the children's FC and BC strategy preferences on various learning tasks. They found that children were equally efficient on both strategies and had similar mixed strategy preferences.

Early research has observed the impact of mixing FC and BC strategies [12]. Priest and Lindsay [12] compared how experts and novices solve physics problems. Although both groups used a mixture of FC and BC, *only* the experts knew how and when to use each strategy and significantly produced more complete plans and stages than their novice peers. In brief, while no consensus has been reached on whether FC or BC is most effective in problem-solving, prior work has observed that the mixture of FC and BC yields the highest performance accuracy as learners know how and when to use each strategy.

2   Methods

Participants They are Computer Science undergraduates at North Carolina State University. Students were assigned each tutor as a class assignment and told that completion is required for full credit. Similar to our prior work, we utilize the random forest classifier (RFC) that, based on pre-test performance, predicts the metacognitive label ($StrTime$ or otherwise) before training on logic and was previously shown to be 96% accurate [2]. Specifically, $StrTime$ students frequently follow the desired behavior of switching *early* (within the first 30 actions) to BC, while their peers either frequently switch late (after the first 30 actions) or stick to the default FC [1–3]. A total of 121 students finished both tutors and were classified by the RFC into 26 $StrTime$ and 95 otherwise. The latter students were randomly assigned to $Experimental$ ($Exp : N = 49$) and $Control$ ($Ctrl : N = 46$) conditions. The RFC was 97% accurate in classifying students who received no intervention —$Ctrl$ and $StrTime$.

Logic Tutor and Our Intervention The logic tutor teaches propositional logic proofs by applying inference rules such as Modus Ponens. A student can solve any problem by either a **FC** or **BC** strategy. Students derive a conclusion at the bottom from givens at the top in FC (Fig. 1a), while they derive a contradiction

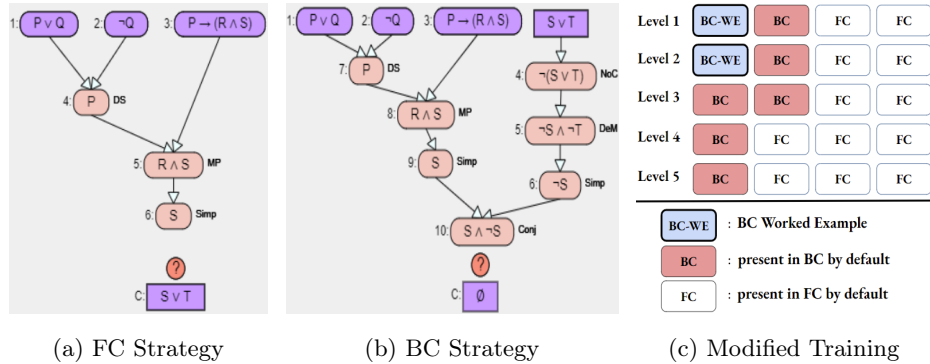(a) FC Strategy   (b) BC Strategy   (c) Modified Training

Fig. 1. Logic Tutor

from givens and the *negation* of the conclusion in BC (Fig. 1b). A problem is presented by *default* in FC with the ability to switch to BC by clicking a button. The tutor consists of two pre-test, 20 training and six post-test problems. The post-test is *much harder* than the pre-test, and the first two post-test problems are isomorphic to the two pre-test problems. The *pre-* and *post-test* scores are calculated by averaging the pre- and post-test problem scores, where a problem score is a function of time, accuracy, and solution length. The training consists of five ordered levels in an *incremental degree of difficulty*, and each level consists of four problems. We modified the training section to mix BC with FC (Fig. 1c). Specifically, two worked examples (WE) on BC were implemented, where the tutor provided a step-by-step solution, and six problems were presented in BC by default. The two WEs and the six problems are expected to teach students *how* and *when* to use BC. Note that the colored problems in Figure 1c were selected based on the historical strategy switches in our data [1].

Probability Tutor It teaches how to solve probability problems using ten principles, such as the Complement Theorem. The tutor consists of a textbook, pretest, training, and post-test. The textbook introduces the domain principles, while training consists of 12 problems, each of which can *only* be solved by *BC* as it requires deriving an answer by *writing and solving equations* until the target is ultimately reduced to the givens. In pre- and post-test, students solve 14 and 20 open-ended problems graded by experienced graders in a double-blind manner using a partial-credit rubric. The *pre-* and *post-test* scores are the average grades in their respective sections, where grades are based *only* on accuracy. Like the logic tutor, the post-test is much harder than the pre-test, and each pre-test problem has a corresponding isomorphic post-test problem.

Procedure Students were assigned to the logic tutor and went through the pre-test, training and post-test. Before training on logic, the RFC predicted the metacognitive label for each student, as described in the Participants section. During training, *Exp* received the modified tutor shown in Figure 1c, while *Ctrl*

and $StrTime$ received the original tutor, where all problems are presented in FC by default. Six weeks later, students were trained on the probability tutor.

3   Results

Table 1. Comparing Groups across Tutors

	$Experimental\,(Exp)$ $(N = 49)$	$Control\,(Ctrl)$ $(N = 46)$	$StrTime$ $(N = 26)$
	Logic Tutor		
Pre	61.7 (18)	58.7 (20)	62.1 (20)
$Iso\text{-}Post$	81 (11)	70.4 (14)	81.3 (10)
$Iso\text{-}NLG$	0.27 (.12)	0.09 (.31)	0.29 (.16)
$Post$	77.4 (11)	66.7 (14)	79 (9)
NLG	0.24 (.15)	0.06 (.37)	0.25 (.18)
	Probability Tutor		
Pre	74.8 (14)	74.2 (16)	75.8 (15)
$Iso\text{-}Post$	90.4 (10)	65.3 (16)	90.6 (8)
$Iso\text{-}NLG$	0.29 (.19)	-0.02 (.27)	0.26 (.17)
$Post$	89.5 (15)	62.5 (18)	88.8 (7)
NLG	0.26 (.21)	-0.08 (.3)	0.24 (.15)

Table 1 compares the groups' performance across the two tutors showing the mean and standard deviation of pre- and post-test scores, isomorphic scores, and the learning outcome in terms of the normalized learning gain (NLG) defined as ($NLG = \frac{Post-Pre}{\sqrt{100-Pre}}$), where 100 is the maximum test score. We refer to pre-test, post-test and NLG scores as Pre, $Post$ and NLG, respectively. On both tutors, a one-way ANOVA found no significant difference on Pre between the groups.

To measure the improvement on isomorphic problems, repeated measures ANOVA tests were conducted using $\{Pre, Iso\text{-}Post\}$ as factor. Results showed that Exp and $StrTime$ learned significantly with $p < 0.0001$ on both tutors, while $Ctrl$ did not perform significantly higher on $Iso\text{-}Post$ than Pre on both tutors. These findings verify the RFC's accuracy, as $StrTime$ learned significantly on both tutors, while $Ctrl$ did not, despite both receiving no intervention.

A comprehensive comparison between the three groups was essential to evaluate our intervention. On the logic tutor, A one-way ANCOVA using Pre as covariate and group as factor found a significant effect on $Post$: $F(2, 117) = 14.5$, $p < .0001$, $\eta^2 = .18$. Subsequent post-hoc analyses with Bonferroni correction ($\alpha = .05/3$) revealed that Exp and $StrTime$ significantly outperformed $Ctrl$: $t(93) = 3.8$, $p < .001$ and $t(70) = 3.9$, $p < .001$, respectively. Similar patterns were observed on NLG using ANOVA and the post-hoc comparisons.

On the probability tutor, a one-way ANCOVA using Pre as covariate and group as factor showed a significant effect on $Post$: $F(2, 117) = 48.1$, $p < .0001$, $\eta^2 = .35$. Follow-up pairwise comparisons with Bonferroni adjustment showed that Exp and $StrTime$ significantly surpassed $Ctrl$: $t(93) = 6.1$, $p < .0001$ and $t(70) = 5.9$, $p < .0001$, respectively. Similar results were found on NLG using ANOVA and the post-hoc comparisons.

4   Conclusion

We showed that mixing BC with FC on the logic tutor improved the experimental students' learning outcomes, as *Exp* significantly outperformed *Ctrl* on logic and on a probability tutor that only supports BC. Additionally, *Exp* caught up with *StrTime* on both tutors suggesting that *Exp* students are prepared for future learning [6] as they acquired BC mastery skills on logic and transferred them to probability, where they received no intervention. There is at least one caveat in our study. The probability tutor supported only one strategy. A more convincing testbed would be having the tutors support both strategies. The future work involves implementing FC on the probability tutor.

References

1. Abdelshiheed, M., et al.: Metacognition and motivation: The role of time-awareness in preparation for future learning. In: Cogsci. vol. 42 (2020)
2. Abdelshiheed, M., et al.: Preparing unprepared students for future learning. In: Cogsci. vol. 43 (2021)
3. Abdelshiheed, M., et al.: The power of nudging: Exploring three interventions for metacognitive skills instruction across intelligent tutoring systems. In: Cogsci. vol. 44 (2022)
4. Belenky, D.M., Nokes, T.J.: Motivation and transfer: The role of mastery-approach goals in preparation for future learning. J. Learn. Sci. **21**(3), 399–432 (2012)
5. de Boer, H., et al.: Long-term effects of metacognitive strategy instruction on student academic performance: A meta-analysis. Educ. Res. Rev. **24**, 98–115 (2018)
6. Bransford, J.D., Schwartz, D.L.: Rethinking transfer: A simple proposal with multiple implications. Review of research in education **24**(1), 61–100 (1999)
7. Chambres, P., et al.: Metacognition: Process, function, and use. Kluwer Academic Publishers (2002)
8. Chi, M., VanLehn, K.: Meta-cognitive strategy instruction in intelligent tutoring systems: How, when, and why. J Educ Techno Soc **13**(1), 25–39 (2010)
9. Fazio, L.K., et al.: Strategy use and strategy choice in fraction magnitude comparison. J. Exp. Psychol.: Learn. Mem. Cogn. **42**(1), 1 (2016)
10. Larkin, J., et al.: Expert and novice performance in solving physics problems. Science **208**, 1335–1342 (1980)
11. Moore, J.W., Quintero, L.M.: Comparing forward and backward chaining in teaching olympic weightlifting. Journal of applied behavior analysis **52**(1), 50–59 (2019)
12. Priest, A., Lindsay, R.: New light on novice—expert differences in physics problem solving. British journal of Psychology **83**(3), 389–405 (1992)
13. Schraw, G., Gutierrez, A.P.: Metacognitive strategy instruction that highlights the role of monitoring and control processes. In: Metacognition: Fundaments, applications, and trends, pp. 3–16. Springer (2015)
14. Slocum, S.K., Tiger, J.H.: An assessment of the efficiency of and child preference for forward and backward chaining. JABA **44**(4), 793–805 (2011)
15. Winne, P.H., Azevedo, R.: Metacognition. In: Sawyer, R.K. (ed.) The Cambridge Handbook of the Learning Sciences. Cambridge Handbooks in Psychology, Cambridge University Press, 2 edn. (2014)