International Society of the Learning Sciences

Downshifting: the nuanced antecedents of seeking easier content

Kirk Vanacore, Amanda Barany, Shruti Mehta, Jaclyn Ocumpaugh, Ryan S. Baker
vanacor1@upenn.edu, abarany@upenn.edu, shrmeh@upenn.edu, ojaclyn@upenn.edu,
ryanshaunbaker@gmail.com
University of Pennsylvania

Nabil Al Nahin Ch, Vishal Kuvar, Caitlin Mills
nabilch@umn.edu, kuvar001@umn.edu, cmills@umn.edu
University of Minnesota

Owen Henkel
owen.henkel@education.ox.ac.uk
University of Oxford

Abstract: Learning involves challenge, and students respond to challenging content in different ways. This study investigates the antecedents and contexts of downshifting: seeking out easier questions while using a digital learning platform. We study this behavior using student data from Rori – a chatbot-based math intelligent tutoring system that operates over WhatsApp and is widely used in several African countries. Specifically, we use Sequential Pattern Mining to study the behaviors that immediately precede and follow downshifting and Bayesian Knowledge Tracing to evaluate students' mastery of the content prior to downshifting. Results show that the behavior is often preceded by sequences of incorrect answers, suggesting that some students may be adaptively downshifting to avoid failure. However, some students seem to downshift despite being on track to master the current content. We also find that students using Rori in schools as part of their regular curricula are likelier to downshift than those who access Rori independently. We position downshifting within a general taxonomy of behaviors related to persistence in ITSs and suggest further research into the causes of the behavior.

Introduction

Learning often requires struggle through difficulty and even failure (Bjork & Bjork, 2020; Kapur, 2012; VanLehn, 1988). However, failure can lead students to disengage in a variety of ways. For example, Baker et al. (2008) hypothesized that gaming the system – exploiting system features to progress without learning – might be used to avoid failure when problems feel too difficult. There is also evidence that students tend to game the system more while working on ambiguous and abstract content, both features likely associated with problem difficulty (Baker et al., 2009). Similarly, stopout – early quitting of a learning activity – tends to be preceded by failure (Botelho et al., 2019). These examples suggest that experiences of failure may cause students to stop meaningfully engaging with the material.

Failure may lead a student to seek easier learning activities as an alternative to disengagement and, possibly, as a form of self-regulation (Higgins & Spiegel, 2004). Some platforms allow students to switch to easier content during a learning activity – a behavior we call downshifting. Various motivations may explain downshifting. In some cases, students may avoid challenging tasks (Baker et al., 2010), whereas in others, they choose not to persist in the face of difficulty (cf. Botelho et al., 2019). However, in other cases, students may recognize that they are not ready for the current content and choose more appropriate content (Zimmerman & Moylan, 2009). As such, downshifting behavior is not necessarily maladaptive or clear evidence of disengagement, as it could be an appropriate response to a misalignment between student abilities and content difficulty.

This study explores the contexts in which students downshift to understand the circumstances in which they exhibit the behavior and identify when it is likely adaptive or maladaptive. We examine downshifting in a chatbot-based, math-focused intelligent tutoring system, Rori (described below), which allows students to switch the grade level of content at any time. Specifically, we explore how students use this agency to navigate problem difficulty and respond to patterns of success and failure. First, we use Sequential Pattern Mining (SPM) to explore the behaviors that precede and follow downshifting. Next, we use Bayesian Knowledge Tracing (BKT) estimates to explore the relationship between downshifting and learning trajectories, allowing us to distinguish between students on track to have mastered the material had they continued with the lesson and

those who were struggling to a greater degree. Finally, we examine the differences in the prevalence of downshifting between students using the Rori voluntarily (in independently chosen activities) and students using Rori as part of their regular school curriculum.

Background

Challenges and failure often benefit learning (Bjork & Bjork, 2020; Kapur, 2012; VanLehn, 1988). Therefore, persistence – the drive to continue a task until its completion in the face of challenge or failure – is generally considered the optimal response to failure (DiNapoli et al., 2023; Ryan, 1938; Peterson & Seligman, 2004). However, students react differently to difficulty, negative feedback, and failure, with some students taking steps to avoid further failure (avoidance motivation) and others strategizing for future success (approach motivation) (Elliot, 2013). Avoidance may manifest in gaming the system (Baker et al., 2009) or stopout behaviors (Botelho et al., 2019). From this perspective, not all persistence is productive (Krumm et al., 2015); sometimes, students persist even when they are not on track to succeed during the learning activity (e.g., wheel-spinning; Beck & Gong, 2013). In these cases, the learner is unprepared for the current activity and may need to learn more foundational content first (e.g., Falmagne & Doignon, 2012).

Thus, there may be times when it is ideal to switch to an easier or prerequisite activity (i.e., downshifting). In other words, downshifting may often be an appropriate self-regulatory learning strategy in which students evaluate their performance and respond with foresight about whether to persist (Zimmerman & Moylan, 2009). Research on problem difficulty in intelligent tutoring systems demonstrates that providing students with problems with optimal difficulty can increase engagement and improve learning outcomes (Papoušek et al., 2016; Zhang & Goh, 2021). It is, therefore, plausible that students downshift to find problems that align with their ability to optimize learning. Furthermore, downshifting may signal appropriate emotional self-regulation as psychological factors (i.e., frustration, anxiety, fatigue) can inhibit their readiness to learn from more complex problems, even if they have the cognitive capacity to succeed (Boekaerts, 2011).

Therefore, the decision of whether or not to downshift is complicated not only by the challenge of a student inferring their readiness and predicting whether persistence will be fruitful, but also by the psychological experiences associated with struggle. For example, Zhang et al. (2021) found that students who reported feelings of difficulty while using a computer-based learning platform were less likely to use strategic behaviors in the learning activity. Furthermore, some students may experience failure aversion, leading them to avoid tasks when they believe they may not succeed (Zhang et al., 2022). Other students seek to avoid effort (Skaalvik, 1997), which could also lead to downshifting when a student perceives that the current content is difficult. In order to fully understand downshifting, we must account for the possibility that students may choose this behavior for a variety of reasons, from maladaptive failure aversion or limited motivation to persist, to adaptive and positive self-regulation strategies.

Method

Study Context

Rori is a free intelligent tutoring system, administered through WhatsApp and accessible with a 2G internet connection, developed to provide students with one-on-one math instruction (Henkel et al., 2024). Unlike many intelligent tutors, Rori interacts with students using natural language processing methods, including large language models. These methods allow students to have an open-ended conversation about math concepts and specific math problems, which Rori presents to students during micro-lessons (short problem sets) on specific topics. These lessons are based on the Global Proficiency Framework, an internationally recognized set of math learning standards (Ovsyannikova, 2019), and content is available for grades 1-10. Each micro-lesson provides instruction, tailored scaffolding, and worked examples. Figure 1a presents an example of Rori's interactions during a lesson on statistics. When students start interacting with Rori, they are placed into difficulty levels based on a system-administered pretest assessment. However, they have the agency to choose which topics to study and to adjust their grade level anytime by asking Rori for the menu. Rori reminds them of this option each time they start a lesson (Fig. 1b). Within this study, we define downshifting as a student moving to a different knowledge component in a lower grade level of content (e.g., 6^{th} to 5^{th}).

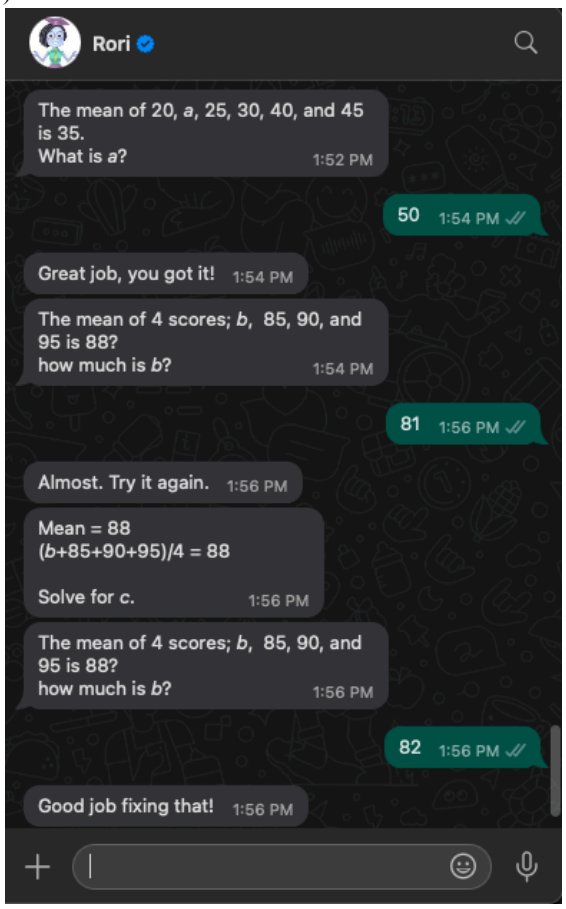
Rori is open and free to anyone using WhatsApp. However, it was specifically developed for use in Africa and is currently used by over 60 thousand users in Ghana, Kenya, Liberia, Nigeria, Rwanda, Sierra Leone, and Tanzania. Some students use Rori in conjunction with their curriculum, while others use it independently. A recent randomized control trial conducted in 11 *Rising Academy* schools in Ghana showed that using Rori during study halls significantly positively impacted math learning outcomes (Henkel et al., 2024).

Throughout this paper, we will refer to those in Ghana's *Rising Academies* as the *Rori In-School* sample and all others as the *Rori Open* sample.

Data

The current study uses problem-level log-file data, which include the lessons and topics students worked on, the grade level of the content, students' responses to Rori math questions, and whether those responses were correct or incorrect. The sample comprises 8,125 students, 5.86% *Rori In-School* and 94.14% *Rori Open.* These data were collected between May 2023 and October 2023. Within that time, students submitted 506,858 attempts to 30,5198 questions across 40,149 microlessons.

(a)



(b)

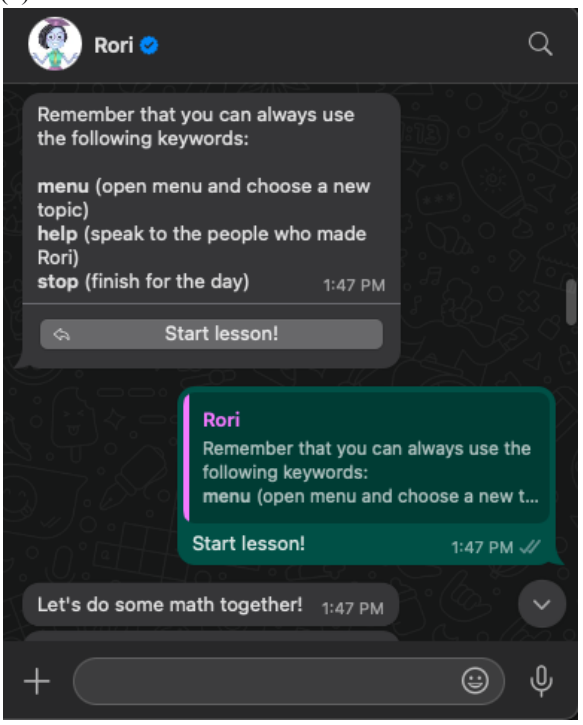


Figure 1

Example chats with Rori. Fig. 1a (left) shows how Rori presents problems and responds to correct and incorrect answers. Fig. 1b (right) shows a message the student receives at the beginning of each lesson, reminding them that they can use the keyword "menu" to choose other topics, including those with different levels of difficulty.

Analyses

Our analyses are primarily exploratory and descriptive as we seek to understand the patterns and frequencies of behaviors surrounding downshifting. We evaluate this behavior using Sequential Pattern Mining (SPM) and Bayesian Knowledge Tracing (BKT).

SPM is a data mining technique that identifies patterns or trends that occur in a specific order over time. It aims to identify more frequent sequences or subsequences of events while maintaining their order of occurrence. SPM is particularly helpful for understanding temporal data, such as a series of actions. In SPM, each data instance consists of an ordered list of items or events. The goal is to discover recurring subsequences that meet certain prevalence (referred to as "support" in SPM) thresholds, indicating that these patterns occur frequently enough across multiple sequences in the dataset. Our analysis used the SPADE algorithm, implemented through the *arulesSequences* package (Buchta, 2023) in *R* (Ihaka & Gentleman, 1996). The

SPADE algorithm was chosen for its efficiency and ease of implementation for real-world applications (Zaki, 2001). When implementing SPM, we restricted our data in three ways. First, we only looked at actions of correct or incorrect attempts on a problem. Multiple attempts on the same problem were included in the analysis. Second, we only examined sequences of actions immediately before and after students switched content – three attempts before the content change and two after. Thus, all the patterns were mutually exclusive. Third, we only examined sequences where the maximum time between problem attempts was less than 30 minutes to avoid considering sequences that span multiple learning sessions.

BKT is a modeling method that tracks a learner's mastery of specific skills or knowledge – known as knowledge components – over time (Corbett & Anderson, 1995). It applies a Hidden Markov Model to estimate the probability that students know a particular concept based on their performance on related tasks. The model uses four key parameters: the probability of learning on the first attempt, the probability of acquisition, the probability of making correct guesses, and the probability of making mistakes despite already having learned the knowledge components (i.e., a slip). BKT is widely used in ITSs to predict future performance (Abdelrahman et al., 2023). The BKT model used in this analysis performed satisfactorily with an AUC of 0.74 when predicting whether students get the next problem correct in a sequence of problems on the same topic, in line with or moderately higher than BKT's typical performance (Mao et al., 2018; Lee et al., 2023). Fifteen (8.33%) micro-lessons had insufficient data for training BKT models (Slater & Baker, 2018) and were excluded from BKT-related analyses.

In our application, we used BKT parameters to predict the likelihood that each student knows the knowledge component during each problem before and after downshifting. We then convert this likelihood to a binary label ("mastery") using a liberal threshold of 0.80. Although the original use of Bayesian Knowledge Tracing used a threshold of 0.95 to indicate mastery (Corbett & Anderson, 1994), subsequent work in some platforms has chosen a threshold of 0.80 (e.g., Wang et al., 2022). This liberal threshold can be seen as distinguishing between whether students are close to mastering the knowledge component, as opposed to more stringent thresholds used to determine with more certainty that the student has mastered the knowledge component. BKT estimates at or above this threshold are designated as instances in which students either mastered the knowledge component or are likely to master it within the next few problems (likely mastered). When estimates are below this threshold, students have not reached mastery (not mastered) and may need multiple subsequent problem attempts or alternative instruction to reach mastery.

Results

Frequency and context of downshifting

On average, students downshifted in 15.06% of the micro-lessons. A substantial minority (18.65%) of students downshifted at least once while using the platform. Nearly half (48.05%) of these students downshift more than once. Those who downshifted at least once averaged 3.99 switches. However, this distribution was positively skewed (SD = 5.57; skew = 2.82), indicating that some students were downshifting substantially more frequently. Downshifts per micro-lesson was positively correlated with the number of micro-lessons students completed (Spearman's rho = 0.42, $p < 0.001$).

Table 1 presents the ten most common sequences identified by the SPM. As stated above, we restricted the sequences to five actions: three problem attempts before and two after each downshift. These ten sequences represent 74% of the patterns surrounding the behavior of downshifting. All but one of the sequences included at least one incorrect response before the downshift. In eight of the sequences, the downshift is immediately preceded by an incorrect response. After downshifting, students are more likely to answer more questions correctly. However, some patterns suggest continuing struggles. Even though students are successful directly after they begin easier content in seven of the top ten sequences, students still gave an incorrect answer within the next two problems in half of these sequences. These patterns reveal that, in most cases, students struggle to produce correct answers before downshifting, and they generally, though not always, perform better after this switch.

Table 1
SPM Results of actions before and after switching to an easier level

	Action Relative to Easier Level Change					
Prevalence	Action -3	Action -2	Action -1	Action 0	Action +1	Action +2
0.19	correct	incorrect	incorrect	downshift	correct	correct
0.13	incorrect	incorrect	incorrect	downshift	correct	correct
0.12	correct	correct	correct	downshift	correct	correct
0.05	correct	incorrect	incorrect	downshift	correct	incorrect
0.05	incorrect	incorrect	incorrect	downshift	correct	incorrect
0.04	correct	correct	incorrect	downshift	correct	correct
0.04	incorrect	incorrect	incorrect	downshift	incorrect	incorrect
0.04	correct	incorrect	incorrect	downshift	incorrect	incorrect
0.04	incorrect	correct	correct	downshift	correct	correct
0.04	incorrect	incorrect	incorrect	downshift	incorrect	correct

These trends can also be seen in the overall average correct/incorrect rates for each action relative to the level change, as presented in Table 2. During the three actions preceding the incorrect response, students' accuracy decreases from 0.60 (Action -3) to 0.37 (Action -2) to 0.32 (Action -1). After switching to easier content, accuracy increases to 0.76 (Action +1) and 0.71 (Action +2).

Table 2
Performance on actions before and after switching to an easier level

Action Relative to Downshift	Correct	Incorrect
Action -3	0.60	0.40
Action -2	0.34	0.66
Action -1	0.32	0.68
Action +1	0.76	0.24
Action +2	0.71	0.29

Overall, these results show that although most students never downshift, the behavior is hardly rare; nearly one-fifth of learners exhibit the behavior. The SPM and descriptive statistics of correct and incorrect responses demonstrate that switching to lower levels is typically a response to failure and often multiple subsequent failures. Although these students are not quitting the system entirely (unlike in the stopout behavior reported in Botelho et al., 2019), they are quitting the current problem set and moving to easier content. Interestingly, students who downshift tend to complete more micro-lessons overall, perhaps indicating that downshifting is not a reflection of overall disengagement with the content.

Relation between downshifting & knowledge component mastery

While patterns of downshifting are a part of some users' experiences, it is unclear whether the behavior is harmful. Downshifting could represent a maladaptive avoidance of difficult problems from which the student could benefit. Alternatively, it may be an adaptive response to a lesson that is too difficult for the student. To differentiate between these two possibilities, we evaluate the behavior based on the probability that the student has mastered the lesson's knowledge component before downshifting, using the BKT estimates from the activity prior to downshifting.

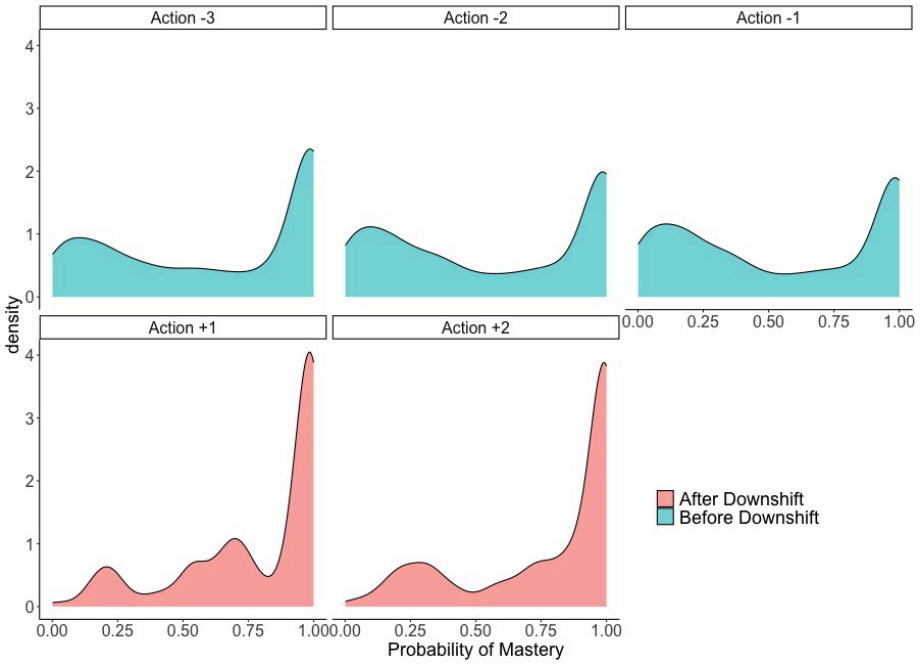


Figure 2
The density plot of the student's probability of mastery during actions preceding and after switching to easier content.

Figure 2 presents the probability distributions that students had mastered the knowledge component on each action preceding and following their downshifting. A visible bimodal distribution is present for the actions preceding the shift, showing that some students downshift when they are unlikely to have mastered a knowledge component, but others do so even when mastery is likely. If we threshold the BKT data by 0.80, as explained in the Method section, almost a third (32%) of these students were likely to master the lesson's knowledge commons on subsequent attempts.

Table 3 shows the prevalence of incorrect responses by the binary mastery estimates after the last action before downshifting (Action -1). The overwhelming majority (96%) of the students with lower mastery levels had at least one incorrect response within the previous three actions. However, even for students with higher mastery levels (over 0.80), downshifting is often preceded by errors; a majority of these students (63%) submitted at least one incorrect response. Thus, some students downshift after failure, even though they appear to be likely to know the skill and are likely to get the next problem correct.

Table 3
Prevalence of the number of failures by mastery estimates

Mastery Action -1	**n**	**Number of Incorrect Responses in the Last Three Actions**			
		0	**1**	**2**	**3**
Likely Mastered	967	0.37	0.24	0.33	0.06
Not Mastered	1474	0.04	0.14	0.45	0.37

Differences in downshifting between school implementation and open access

So far, we have focused on factors within the Rori platform associated with downshifting behavior, such as failure and probability of mastery; however, external contextual factors are also likely to play a role in this behavior. Although we do not have extensive information about each student's learning environment, we can tell whether students are in the *Rori In-School* or *Rori Open* populations. This is an important distinction, because most *Rori Open* students are studying independently and likely of their own volition, whereas *Rori In-School* students were required to use Rori as part of their normal school day. These contextual differences may result in student samples with different levels of motivation, which may impact when and how downshifting occurs. For

example, students who opt-in to using *Rori* may have more intrinsic motivation to persist compared to those with required usage.

Table 4 presents the average percentages of micro-lessons in which students downshifted across the two populations. Students in the *Rori In-School* sample downshifted more frequently than those in the *Rori Open* sample ($p < 0.001$). As Table 4 shows, this result was consistent regardless of whether students were likely or unlikely to master the knowledge component before downshifting.

Table 4

Means (standard deviations) and Mann-Whitney test of differences in the percent of micro-lessons in which students downshift behaviors between Rori populations

	Rori In-School	Rori Open	p-value
Downshifting	29.38% (19.95%)	4.00% (11.47%)	< 0.001
Downshifting - likely mastered	8.00% (8.51%)	1.37 (6.37%)	< 0.001
Downshifting - unlikely mastered	16.47% (14.88%)	2.19% (8.40%)	< 0.001

Discussion

This paper examines downshift – moving to easier content during a learning activity – which has been underexplored in learning sciences. Almost a fifth of our sample downshifted at least once, often after strings of incorrect responses. According to knowledge tracing estimates, most of these students had less than an 80% likelihood of mastering the knowledge component on which they were working. However, in almost one-third of these instances, knowledge tracing estimates indicated that students had mastered or were close to mastering the lesson's skill. These findings, considered together, suggest that although downshifting behaviors are likely a low-persistence response to failure, the behavior may be adaptive for some and maladaptive for others. Furthermore, downshifting prevalence also varied based on whether students accessed Rori as part of their regular curriculum, suggesting that external contextual factors influenced this behavior.

We propose that downshifting can be integrated into a taxonomy of persistence behaviors in mastery learning activities, which involves two dimensions: the level of persistence and the probability of mastering the knowledge component when persisting (see Table 5). Krumm et al. (2016) define productive persistence as a tendency to continue with a difficult task as well as a "good use of strategies." For strategies to be good, their application must lead to positive results (e.g., mastery). For example, persistence is likely to be productive if a student is highly likely to master the knowledge component. Alternatively, high persistence with a low probability of mastery leads to wheel spinning. Downshifting is positioned as a lower persistence behavior – between persistence and disengagement -- which can also be divided based on students' likelihood of mastery. Downshifting with a high probability of mastery is maladaptive, as the student would benefit from persistence. By contrast, when students have a lower probability of mastery, downshifting may prevent wheel spinning and improve learning. Notably, downshifting is not the lowest level of persistence, as students are likely still engaged in the content; students who downshifted more frequently tended to complete more micro-lessons.

Table 5

Taxonomy of persistence behaviors in mastery learning activities

	Probability of Mastery of Knowledge Components	
	Higher	**Lower**
Higher Persistence	productive persistence	wheel spinning
Lower Persistence	maladaptive downshifting	adaptive downshifting
Disengagement	gaming the system, stopout, etc.	

There are many plausible explanations for why students downshift even when they have are close to or have mastered the content. Our results show that even for these students, downshifting tends to be preceded by multiple incorrect responses. This aligns closely with the "soft underbelly" conceptualization of downshifting, where students simply avoid challenging content (Baker et al., 2010). Failure aversion provides a potential explanation for this behavior (Zhang, 2022); students might switch content to avoid the negative feelings associated with getting problems wrong. Fear of failure during learning activities is negatively associated with

metacognitive self-regulation (Bartles & Jackson, 2009), suggesting that failure aversion may disrupt students' decision-making processes when evaluating whether or not to downshift. Downshifting is also plausibly related to perfectionism, which is associated with anxiety and maladaptive behaviors such as cheating and procrastination, particularly when perfectionism is socially prescribed (Bong et al., 2014). This theory may explain the significant differences in behavior between the mandatory school implementations, where teachers were monitoring student progress and performance, and the opt-in use of Rori, where students were likely using the platform independently. At school, students may have responded to pressure to get problems correct by avoiding micro-lessons where they had previously failed.

Self-regulated learning also provides a framework for understanding downshifting as a learning strategy. In cases where downshifting is an adaptive learning strategy, students likely reflect on their performance with appropriate self-evaluations and have the foresight to downshift, allowing them to avoid unproductive persistence (Zimmerman & Moylan, 2009). However, this process is likely complex, as self-regulated learning requires students to manage their emotional states and protect their sense of self (Panadero, 2017). Thus, when appraising a learning experience, the student must apply coping strategies to handle any negative emotions from failure experiences (Boekaerts, 2011). In the case of downshifting, students must balance affective and cognitive appraisals to determine whether to proceed with the activity. Although downshifting may often be a product of failure aversion, in some cases, it can be a strategy for mitigating the emotional consequences of failure that disrupt the learning process, as well as the prospect of unproductive persistence.

Future Directions

The prevalence and context of downshifting suggest that it is a nuanced learning strategy with adaptive and maladaptive applications. Our analyses highlight that future research is warranted to document and understand the interplay of students' emotional regulation, self-efficacy, confusion, frustration, and cognitive abilities that may cause downshifting. Future research should consider variations in when students exhibit this behavior (e.g., Is it directly after placement into the content by the initial test or does it occur later, perhaps after wheel spinning?). More work can be done to differentiate between unproductive failure aversion and more constructive efforts to avoid wheel spinning. A deeper understanding of why the context of use predicts downshifting—possibly through methods like immediate post-downshifting interviews (Baker et al., in press)—could also help to reveal motivational factors driving this behavior. Finally, it is crucial to understand how downshifting is associated with distal outcomes (e.g., Do students eventually successfully return to the initial lesson? How do they perform on external learning measures?). With a better understanding of the factors linked to downshifting and its effects on learners, we can guide them in recognizing when downshifting is beneficial and when it is counterproductive.

Acknowledgments

We would like to thank the Rori and Rising Academies staff for their contributions. This work was supported by the Learning Engineering Virtual Institute (LEVI) program.

References

Abdelrahman, G., Wang, Q., & Nunes, B. (2023). Knowledge tracing: A survey. *ACM Computing Surveys*, 55(11), 1-37.

Roll, I., Baker, R. S., Aleven, V., & Koedinger, K. R. (2011). Modeling students' metacognitive errors in two intelligent tutoring systems. *User modeling and user-adapted interaction*, 21(4), 383-403.

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why Students Engage in "Gaming the System" Behavior in Interactive Learning Environments. *Journal of Interactive Learning Research*, 19(2), 185–224.

Baker, R., Carvalho, A., Raspat, J., Aleven, V., & Koedinger, K. R. (2009). Educational Software Features that Encourage and Discourage "Gaming the System." *Proceedings of the 14th International Conference on Artificial Intelligence in Education*, 14, 475–482.

Baker, R.S., Hutt, S., Bosch, N., Ocumpaugh, J., Biswas, G., Paquette, L., Andres, J.M.A., Nasiar, N., Munshi, A. (in press) Detector-Driven Classroom Interviewing: Focusing Qualitative Researcher Time by Selecting Cases in Situ. To appear in *Educational Technology Research & Development*.

Baker, R., Mitrović, A., & Mathews, M. (2010). Detecting Gaming the System in Constraint-Based Tutors. In P. De Bra, A. Kobsa, & D. Chin (Eds.), *User Modeling, Adaptation, and Personalization* (pp. 267–278).

Bartels, J. M., & Magun-Jackson, S. (2009). Approach–avoidance motivation and metacognitive self-regulation: The role of need for achievement and fear of failure. Learning & Individual Differences, 19(4), 459-463.

Beck, J. E., & Gong, Y. (2013). Wheel-Spinning: Students Who Fail to Master a Skill. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Artificial Intelligence in Education* (pp. 431–440). Springer.

Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, 9(4), 475–479. https://doi.org/10.1016/j.jarmac.2020.09.003

Bong, M., Hwang, A., Noh, A., & Kim, S.-i. (2014). Perfectionism and motivation of adolescents in academic contexts. *Journal of Educational Psychology,* 106(3), 711–729. https://doi.org/10.1037/a0035836

Botelho, A. F., Van Inwegen, E. G., Varatharaj, A., & Heffernan, N. T. (2019). Refusing to try: Characterizing early stopout on student assignments. *LAK19: Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 391–400. https://doi.org/10.1145/3303772.3303806

Buchta, C., Hahsler, M., Diaz, D., Buchta, M. C., & Zaki, M. J. (2023). Package 'arulesSequences'. https://bioconductor.statistik.tu-dortmund.de/cran/web/packages/arulesSequences/arulesSequences.pdf

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 4(4), 253–278.

DiNapoli, J. (2023). Distinguishing between Grit, Persistence, and Perseverance for Learning Mathematics with Understanding. *Education Sciences*, 13(4), Article 4. https://doi.org/10.3390/educsci13040402

Elliot, A. J. (2013). Handbook of approach and avoidance motivation. Psychology Press.

Henkel, O., Horne-Robinson, H., Kozhakhmetova, N., & Lee, A. (2024). Effective and Scalable Math Support: Experimental Evidence on the Impact of an AI-Math Tutor in Ghana. *In International Conference on Artificial Intelligence in Education* (pp. 373-381). Cham: Springer Nature Switzerland.

Higgins, E. T., & Spiegel, S. (2004). Promotion and prevention strategies for self-regulation. *Handbook of self-regulation: Research, theory, and applications*, 171-187. The Guilford Press.

Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. Journal of computational and graphical statistics, 5(3), 299-314.

Kapur, M., & Bielaczyc, K. (2012). Designing for Productive Failure. *Journal of the Learning Sciences*, 21(1), 45–83.

Krumm, A., Coulson, A., & Neisler, J. (2022, March). Defining productive struggle in st math: Implications for developing indicators of learning behaviors and strategies in digital learning environments. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (pp. 514-520).

Lee, M., Croteau, E., Gurung, A., Botelho, A. F., & Heffernan, N. T. (2023). Knowledge Tracing Over Time: A Longitudinal Analysis. Unpublished manuscript. https://doi.org/10.5281/ZENODO.8115788

Liu, C., Liu, J., Cole, M., Belkin, N. J., & Zhang, X. (2012). Task difficulty and domain knowledge effects on information search behaviors. *Proceedings of the American Society for Information Science and Technology, 49*(1), 1–10.

Mao, Y., Lin, C., & Chi, M. (2018). Deep Learning vs. Bayesian Knowledge Tracing: Student Models for Interventions. *Journal of Educational Data Mining*, 10(2), 28-54.

Ovsyannikova, O., 2019, *Global proficiency framework*, UNESCO Institute for Statistics, Montreal.

Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, 8. https://doi.org/10.3389/fpsyg.2017.00422

Papoušek, J., Stanislav, V., & Pelánek, R. (2016). Impact of question difficulty on engagement and learning. In *Proc. Intelligent Tutoring Systems: 13th International Conference*. (pp. 267-272).

Peterson, C.; Seligman, M.E.P. *Character Strengths and Virtues: A Handbook and Classification*; Oxford University Press/American Psychological Association: New York, NY, USA, 2004.

Ryans, D. G. (1938). The Meaning of Persistence. *The Journal of General Psychology.* (pp. 79–96)

Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of educational psychology, 89*(1), 71.

Slater, S., & Baker, R. S. (2018). Degree of error in Bayesian knowledge tracing estimates from differences in sample sizes. *Behaviormetrika*, 45(2), 475-493.

Thiele, C., & Hirschfeld, G. (2021). cutpointr: Improved Estimation and Validation of Optimal Cutpoints in R. Journal of Statistical Software, 98, 1–27.

Vanacore, K., Sales, A., Liu, A., & Ottmar, E. (2023, July). Benefit of gamification for persistent learners: Propensity to replay problems moderates algebra-game effectiveness. In *Proceedings of the Tenth ACM Conference on Learning@ Scale* (pp. 164-173).

VanLehn, K. (1988). Toward a Theory of Impasse-Driven Learning. In H. Mandl & A. Lesgold (Eds.), *Learning Issues for Intelligent Tutoring Systems* (pp. 19–41). https://doi.org/10.1007/978-1-4684-6350-7_2

Ventura, M., Shute, V., & Zhao, W. (2013). The relationship between video game use and a performance-based measure of persistence. *Computers & Education*, 60(1), 52–58.

Vygotsky, L. S., & Cole, M. (1978). *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press.

Wang, K., Ma, Z., Baker, R. S., & Li, Y. (2022). Iterative Refinement of an AIS Rewards System. In *International Conference on Human-Computer Interaction* (pp. 113-125). Cham: Springer International Publishing.

Zaki, M. J. (2001). SPADE: An Efficient Algorithm for Mining Frequent Sequences. Machine Learning, 42(1), 31–60.

Zhang, L., Wang, B., Feng, X., Zhang, Y., & Wang, W. (2022). Exploring the influence of failure aversion on learning from project failure: a Sensemaking perspective. *Frontiers in Psychology*, 13, 794390.

Zhang, Y., & Goh, W. B. (2021). Personalized task difficulty adaptation based on reinforcement learning. *User Modeling and User-Adapted Interaction*, 31(4), 753-784.

Zhang, Y., Paquette, L., Baker, R. S., Bosch, N., Ocumpaugh, J., & Biswas, G. (2023). How are feelings of difficulty and familiarity linked to learning behaviors and gains in a complex science learning task?. *European Journal of Psychology of Education*, *38*(2), 777-800.

Zimmerman, B. J., and Moylan, A. R. (2009). "Self-regulation: where metacognition and motivation intersect," in H*andbook of Metacognition in Education*, eds D. J. Hacker, J. Dunlosky, and A. C. Graesser (New York, NY: Routledge), 299–315.