

International Journal of Learning, Teaching and Educational Research
 Vol. 25, No. 8, pp. 661-677, August 2026
<https://doi.org/10.26803/ijlter.25.8.26>
 Received Feb 5, 2026; Revised Jun 27, 2026; Accepted Aug 3, 2026

Systematic Prompt Fading to Foster Self-Questioning and Reading Comprehension in At-Risk Elementary Students: A Multiple-Baseline Single-Case Study

Matthias Grünke^{*} , Anne Barwasser  and Susanne Hoff 
 University of Cologne
 Cologne, Germany

Abstract. This study examined the effects of a structured reading comprehension intervention based on systematic prompt fading on three third-grade students from a socioeconomically disadvantaged elementary school in Germany. Self-questioning was incorporated as one instructional component of the intervention. A single-case AB design with multiple baselines across participants was employed to evaluate functional relations between the intervention and outcomes. Each student engaged in 15 individual sessions, during which comprehension was assessed through curriculum-based probes consisting of 250-word texts and ten open-ended questions. Data were analyzed using descriptive statistics, nonparametric effect size indices (PET, Tau-U), and piecewise regression models at both the individual and group level. Results demonstrated clear improvements for all participants, with median comprehension scores rising from 1 correct answer at baseline to 5 during the intervention. Nonparametric indices indicated strong or large effects, while regression analyses confirmed significant slopes and level changes in Phase B. Social validity outcomes were uniformly positive. Findings provide convergent evidence that systematic prompt fading can substantially enhance reading comprehension in ecologically valid school settings.

Keywords: reading comprehension; self-questioning strategies; prompt fading; at-risk learners; single-case research

Citation:

Grunke, M., Barwasser, A., & Hoff, S. (2026). Systematic Prompt Fading to Foster Self-Questioning and Reading Comprehension in At-Risk Elementary Students: A Multiple-Baseline Single-Case Study. *International Journal of Learning, Teaching and Educational Research*, 25(8), 661–677. <https://doi.org/10.26803/ijlter.25.8.26>

*Corresponding author: Matthias Grünke; matthias.gruenke@uni-koeln.de

1. Introduction

1.1 Reading Comprehension as a Central Educational Outcome

Reading is a foundational skill in contemporary societies. In educational contexts, it is not limited to language arts instruction but functions as a prerequisite for learning across virtually all subject areas, including mathematics, science, and social studies. Beyond formal schooling, reading competence is essential for professional qualification, occupational advancement, and meaningful participation in civic and digital life. Everyday activities likewise presuppose adequate literacy, such as understanding written instructions or navigating online communication (e.g., emails, online forms, and text messages). Leisure pursuits also rely on literacy, including following blogs, participating in online forums, or engaging with social media content. This competence strongly influences educational attainment, career prospects, and social participation, as individuals with lower literacy skills may struggle to access essential information and opportunities (Cremin et al., 2023; Watkins, 2018).

From an educational-developmental perspective, reading proficiency typically unfolds along a developmental trajectory that begins with phonological awareness, progresses through grapheme-phoneme correspondences and decoding, and advances toward fluent word recognition and ultimately comprehension of connected text (Jackson & Coltheart, 2001; Stuart & Stainthorp, 2015). Widely accepted frameworks, such as the Simple View of Reading, conceptualize comprehension as the product of decoding and language understanding (Gough & Tunmer, 1986). Thus, constructing meaning from written language represents the culmination of reading development and a key competence expected by the end of elementary school (Cartwright, 2023; Connor, 2016).

For most individuals, reading comprehension develops successfully during the first years of schooling. By the end of elementary education, the majority of students are able to decode texts fluently and construct meaning (Kilpatrick et al., 2019). However, a substantial proportion of children fail to attain adequate levels of reading comprehension. Syntheses of large-scale international assessments, particularly PIRLS 2021, show that in several major education systems, more than one quarter of fourth-grade students perform below minimum proficiency levels.

This pattern extends beyond primary education: results from the Programme for International Student Assessment (PISA) likewise indicate that a considerable proportion of 15-year-olds across OECD countries do not reach baseline proficiency in reading, pointing to persistent and systemic challenges in literacy development (OECD, 2023). Even when decoding skills are sufficient, these students often experience pronounced comprehension difficulties, with negative implications for academic achievement and participation. Multilingual learners and students from socioeconomically disadvantaged backgrounds face especially pronounced risks (Delaney et al., 2023).

1.2 Evidence-Based Instruction and the Role of Scaffolded Strategy Use

Encouragingly, research over the past several decades has identified a range of effective instructional approaches to support students who struggle with reading comprehension. Multiple meta-analyses conducted over the past decades consistently indicate that explicit, strategy-oriented instruction is particularly effective in improving reading comprehension outcomes for students with difficulties in this area (Anderson et al., 2023; Budianto et al., 2022; Hudler et al., 2025; Joseph et al., 2016; Maimunsyah et al., 2025). Across studies, instructional methods that emphasize modeling, guided practice, scaffolding, and a gradual progression from simpler to more complex tasks lead to substantial gains in students' understanding of texts.

Interventions are especially effective when students actively engage with reading materials, for example, by monitoring their understanding, identifying key information, generating questions, and summarizing, and when responsibility is systematically transferred from teacher to learner. Research syntheses further show that modeled strategy use, corrective feedback, and systematic prompt fading improve outcomes even for learners with substantial challenges. Overall, the evidence converges on a core principle: strategic behaviors should be explicitly modeled, scaffolded, and gradually withdrawn (Vaughn et al., 2024).

1.3 Prompt Fading, Research Gaps, and the Purpose of the Present Study

One instructional technique that instantiates these principles with particular conceptual precision, yet remains comparatively underexamined in reading research, is prompt fading. Originating from behavior-analytic theories of stimulus control, this approach is designed to promote responding to the natural discriminative stimulus rather than to supplemental prompts through the systematic reduction of prompts (Cengher et al., 2018; Deitz & Malone, 1985). In doing so, it operationalizes the gradual release of responsibility not merely as a pedagogical metaphor, but as a programmed mechanism grounded in well-established learning principles such as blocking and overshadowing. When used to teach reading comprehension, prompt fading is a theoretically sound way to teach students how to ask themselves questions. They start with text-based questions, then move to self-generated ones, and finally to independent comprehension monitoring in a step-by-step, verifiable way.

The limited empirical literature aligns closely with this logic. Two single-case experimental studies demonstrated that systematic prompt fading can effectively teach fourth- and fifth-grade students with learning difficulties to generate their comprehension questions, resulting in clear improvements in reading comprehension and evidence of maintenance and partial generalization to novel texts (Rouse-Billman & Alber-Morgan, 2019; Rouse et al., 2014). In both studies, students first engaged with texts containing embedded questions; these were then replaced by prompts to self-generate questions, which were subsequently faded, precisely mirroring the modeled strategy use, scaffolded practice, and stepwise release highlighted as effective in broader intervention research.

A systematic search across major scholarly databases (ERIC, PsycINFO, Scopus, Web of Science) identified no additional empirical studies investigating prompt fading in the context of reading comprehension, underscoring the remarkable scarcity of research on an instructional technique that so closely matches the core features of the most effective interventions (search date: December 10, 2025). Although theoretically well-grounded and conceptually aligned with established models of gradual release of responsibility, empirical research on prompt fading in reading comprehension, particularly in non-English-speaking contexts, remains limited.

The present study explicitly focuses on teaching practices, in addition to contributing to the empirical literature on reading comprehension interventions. By specifying concrete instructional procedures, decision rules for performance-contingent prompt fading, and criteria for reducing support, the study is intended to inform educational practice in school-based settings. The focus is on the intervention's effectiveness and its feasibility for implementation in typical classroom and intervention contexts. Accordingly, the present study seeks to address this gap by examining the effects of a structured prompt-fading-based self-questioning approach on the reading comprehension performance of third-grade students at risk for learning difficulties in a German elementary school context.

1.4 Research Questions

Against this background, we aimed to examine the effects of a structured self-questioning intervention with systematic prompt fading on the reading comprehension performance of elementary school students at risk for learning difficulties. Specifically, the following research questions guided the investigation:

1. To what extent does a prompt-fading intervention that incorporates self-questioning improve reading comprehension performance in third-grade students with adequate decoding skills but pronounced comprehension difficulties?
2. Does the introduction of prompt fading lead to clear and replicable changes in level, trend, and variability of reading comprehension performance across participants in a multiple-baseline design?
3. How do participating students perceive the acceptability, feasibility, and usefulness of the prompt-fading-based self-questioning intervention (social validity)?

These questions aim to determine whether prompt fading is an effective and socially valid mechanism for fostering independent reading comprehension strategies.

2. Method

2.1 Participants and Setting

We conducted this study in a public elementary school located in a metropolitan region of western Germany. At the time of data collection, it served a culturally and linguistically diverse student population. It was characterized by elevated

socioeconomic risk indicators, as documented by the regional school monitoring system. To identify suitable candidates for participation, curriculum-based measurement data from the entire third-grade cohort at the school, routinely collected as part of school-wide assessment procedures, were used. These data had been obtained approximately two weeks before the selection process. The measures constituted an internal, non-published screening procedure based on school-specific norms. They were used to identify students whose reading comprehension performance fell within the lowest 10% of their local peer group.

Subsequently, these third graders completed the one-minute fluency subtests of the Salzburg Reading and Spelling Battery-2 (SLRT-2; Moll & Landerl, 2014), which assesses accuracy and speed in reading real words and pseudowords. Eligibility required a minimum percentile rank of 30 on the SLRT-2, thereby ensuring sufficient reading fluency to meet the cognitive demands of the instructional procedures. This multistep process led to a final sample of three students, all of whom had serious problems with reading comprehension but were able to decode words well enough. Participation was voluntary. Written parental consent and child assent were obtained before study onset, and ethical approval was granted by the Learning Disabilities Worldwide Ethics Committee (approved June 19, 2025).

To ensure full compliance with ethical standards and protect participant confidentiality, all identifying details were removed from the dataset. The names used throughout the manuscript are researcher-assigned pseudonyms that bear no relation to the participants' actual identities and were selected solely for illustrative purposes. Furthermore, potentially identifying characteristics were either omitted or modified to reduce the likelihood of participant recognition. Key information about the students is summarized in Table 1.

Table 1: Participant Characteristics

Participant	Age	Gender	Language Background	Reading Fluency
Ayhan	8	Male	Turkish German	32
Bhavya	9	Female	Hindi-German	38
Cyra	9	Female	Arabic-German	40

Note. Reading fluency values are reported as percentile ranks (PR)

The following descriptions provide additional qualitative details to complement the participant characteristics presented in Table 1: Ayhan was described by his teacher as introverted and soft-spoken, with notable weaknesses in German expressive language, including a limited vocabulary and frequent grammatical errors. Despite these challenges, he demonstrated strong mathematical competence and was able to complete academic tasks when appropriately scaffolded. His error rates in reading were sufficiently low to permit engagement with connected text.

Bhavya showed academic performance ranging from average to slightly below average. Her teacher consistently reported literacy difficulties, but she excelled in artistic and physical activities. Although her reading performance was below age expectations and associated with elevated error rates, she exhibited high

motivation and performed comparably to the other participants during initial probes.

Cyra was described as easily distracted and somewhat insecure in academic contexts, often requiring repetition, structured scaffolding, and positive reinforcement. In addition to literacy challenges, she required support in mathematics and general classroom engagement. Her decoding profile was characterized by relatively faster reading combined with elevated error rates, indicating a speed-accuracy imbalance commonly associated with reading comprehension difficulties.

All sessions were conducted individually in quiet school rooms typically used for small-group instruction or tutoring. Although room assignment varied slightly depending on school availability, the instructional environment remained consistently calm and well-controlled. Each participant sat side-by-side with one or two researchers at a table, facilitating modeling and access to instructional materials while maintaining a standardized physical arrangement. Except for an occasional co-occupant from another support program, sessions were conducted without others present, minimizing external distractions. Session scheduling, seating arrangement, and materials placement were standardized across participants to minimize contextual variability. Instructional materials, daily routines, and researcher prompts were kept uniform across sessions to ensure procedural fidelity.

2.2 Design

We implemented a concurrent multiple-baseline single-case experimental design across participants (Ledford & Gast, 2024). Baseline lengths were staggered to allow three time-lagged demonstrations of effect (Ayhan = 3 sessions, Bhavya = 4 sessions, Cyra = 5 sessions). All participants began baseline concurrently and transitioned individually to intervention once a stable performance pattern had been established. Whenever possible, we conducted one session per school day, resulting in an approximate 24-hour interval between measurement occasions.

Outcomes were measured daily during both phases to capture within- and between-phase variability. Internal validity was supported by phase separation, systematic staggering, consistent instrumentation (randomized texts and questions), and control of treatment diffusion. Standardized instructions, consistent scoring rubrics, and cross-checking by two researchers enhanced objectivity and reliability.

2.3 Dependent Variable

The dependent variable was students' reading comprehension performance, operationalized as the number of correctly answered open-ended comprehension questions per session. Reading comprehension was assessed using written responses to standardized comprehension probes administered at the end of each session. We scored responses dichotomously (0 = incorrect, 1 = correct), yielding a maximum score of 10 per session. A response was scored as correct if it accurately reflected the information stated in the text and addressed the question without contradiction or omission of essential elements. We disregarded minor

spelling or grammatical errors as long as the intended meaning was unambiguous.

To ensure scoring objectivity and reliability, a second independent rater scored at least 25% of all comprehension probes across participants and phases using the same scoring criteria. Point-by-point interrater agreement exceeded 95% across all double-scored sessions.

2.4 Materials

Baseline sessions utilized non-literacy activities (e.g., Sudokus, visual search tasks, and arithmetic problems without text) delivered on tablets to control for session duration and cognitive engagement while minimizing influences on literacy-related outcomes. The standardized comprehension probes used to assess the dependent variable consisted of randomly selected expository texts of approximately 250 words with a Flesch Reading Ease score of 80, verified using the readability statistics function in Microsoft Word. Each text was accompanied by 10 open-ended comprehension questions; however, responses were restricted to a single word per question. Texts and comprehension questions were not repeated across sessions.

Instructional texts used during the intervention phase were initially shorter and less complex (approximately 130 words; Flesch Reading Ease $\approx$ 90) and increased gradually in length and syntactic complexity to approximate the difficulty of the comprehension probes. In early sessions, comprehension questions were provided explicitly; in later sessions, these were systematically replaced with blank spaces and stop-sign symbols, prompting students to generate their own questions.

To enhance motivation, students monitored their daily performance using individual line graphs (self-graphing) and earned stickers for stable or improved performance. Every fifth sticker (indicated by a crown symbol) could be exchanged for small prizes (e.g., stationery, sweets). Reinforcement was explicitly tied to demonstrated comprehension-related behaviors.

2.5 Procedures

2.5.1 Baseline Phase

Each baseline session lasted 20 minutes. To control for individual attention, session duration, and task engagement while deliberately excluding reading-related instruction, sessions began with 15 minutes of non-literacy activities (e.g., puzzles, visual search tasks, arithmetic problems without text), selected based on students' interests. These matched the time-on-task structure of the intervention sessions but did not involve any reading comprehension practice or feedback.

During the final five minutes of each session, students independently read a 250-word expository text and answered 10 open-ended comprehension questions in written form. No instructional support or feedback was provided. Responses were scored dichotomously (0 = incorrect, 1 = correct), yielding a maximum possible score of 10 per session. This procedure established a stable baseline of

reading comprehension performance under conditions of controlled attention and equivalent session length.

2.5.2 Intervention Phase

Each intervention session also lasted 20 minutes and implemented a structured self-questioning approach with systematic, performance-contingent prompt fading. Sessions began with a brief review of the student's performance chart displaying comprehension probe scores from previous sessions. The researcher provided standardized motivational feedback, explicitly attributing changes in performance to effort and strategy use.

The instructional phase lasted approximately 15 minutes and used expository texts that increased gradually in length and syntactic complexity to approximate the difficulty of the assessment probes. Instruction began with explicit modeling, during which the researcher read the text aloud and verbalized comprehension processes, including identifying key information, formulating relevant questions, and locating answers directly in the text. The researcher generated sample questions during this phase, allowing the student to observe without the need to respond.

Guided practice followed, with the researcher and student reading the text together. Stop-sign symbols embedded in the text cued the student to pause, reflect, and attempt to formulate and answer content-related questions. Student responses were supported through a standardized prompting hierarchy consisting of (a) a general prompt to ask a question, (b) a specific prompt directing attention to a sentence or idea unit, and (c) a model prompt. The researcher modeled an appropriate question, explained how it was derived from the text, and asked the student to rephrase it when errors or omissions occurred. Instructional interactions remained highly structured and non-dialogic throughout, with explicit guidance provided at each step and the student rehearsing the strategy under systematic support.

Transition to independent self-questioning occurred when the student generated at least two consecutive, text-relevant questions without assistance. In this phase, the child was expected to generate and record their comprehension questions at each stop-sign marker. The researcher intervened only when hesitation exceeded five seconds or when questions were off-topic or incomplete, applying the least intrusive prompt necessary. As sessions progressed, we systematically removed prewritten questions, requiring increasing reliance on self-generated questioning. Prompt fading was performance-contingent rather than time-based; we reduced support only when students demonstrated consistent and accurate strategy use and temporarily reinstated higher levels of prompting if performance declined.

Across all sessions, students were reminded to pause after each sentence or paragraph, summarize aloud, and continue reading only after verifying comprehension. Observed difficulties, such as limited retention of key details or overly rapid reading, were addressed through brief, targeted strategy coaching embedded within instruction.

Each session concluded with the standardized comprehension probe administered under identical conditions as in baseline. Students independently read a 250-word expository text and answered 10 open-ended comprehension questions without feedback. Performance was recorded immediately on the student's self-graphing chart. To reinforce sustained engagement, students earned stickers for stable or improved performance, with every fifth sticker exchanged for a small prize. Reinforcement was explicitly linked to effort and strategy application rather than outcomes alone. No adverse events occurred during the study. By the end of the intervention phase, students applied self-questioning strategies with minimal external support, indicating increased autonomy in monitoring their reading comprehension.

2.6 Treatment Fidelity

We assessed procedural fidelity to ensure that the independent variable was implemented as intended and consistently across phases. Fidelity data were collected during three independently observed sessions per participant (one during baseline, two during intervention), following recommendations for intermittent but systematic fidelity checks in single-case research.

A structured fidelity checklist was used, covering (a) session duration and sequencing, (b) modeling procedures, (c) adherence to the prompting hierarchy, (d) application of the error-correction protocol, (e) delivery of reinforcement, and (f) standardized administration of outcome probes. Each component was scored dichotomously (implemented correctly/not implemented), and fidelity was calculated as the percentage of correctly implemented steps per session.

Across all observed sessions and components, treatment fidelity exceeded 90%, indicating high procedural integrity. No systematic deviations were observed, supporting the internal validity of the functional relationship between the intervention and observed changes in reading comprehension.

2.7 Social Validity

Consistent with the original conceptualization of social validity as a three-component construct encompassing goals, procedures, and outcomes (Wolf, 1978), the present study assessed the acceptability, feasibility, and perceived effectiveness of the intervention from the perspective of the direct participants. Social validity was assessed at the conclusion of the study using an eight-item child-friendly questionnaire developed for this study. The questionnaire employed three-point response options ("yes," "a little," "no") and was designed to provide a simple and age-appropriate measure of children's perceptions of the intervention.

Items addressed (a) perceived usefulness and learning gains (outcome social validity), (b) enjoyment and acceptability of the instructional procedures (procedural social validity), and (c) willingness to continue or recommend the intervention (goal relevance and overall acceptability). Item scores were summed to yield a total score ranging from 8 to 24, with higher scores indicating greater perceived social validity. To ensure comprehension and reduce measurement error, all items were read aloud when necessary, and children were encouraged

to ask clarifying questions before responding. Administration procedures were standardized across participants.

3. Results

Descriptive statistics, effect size calculations, regression analyses, and graphical displays were computed using the SCAN package for R (Wilbert, 2026). Figure 1 presents the session-by-session reading comprehension performance of all three participants across baseline (A) and instructional (B) phases. Individual line graphs display the number of correctly answered comprehension questions per session, allowing inspection of the level, trend, variability, and immediacy of change associated with the phase transitions.

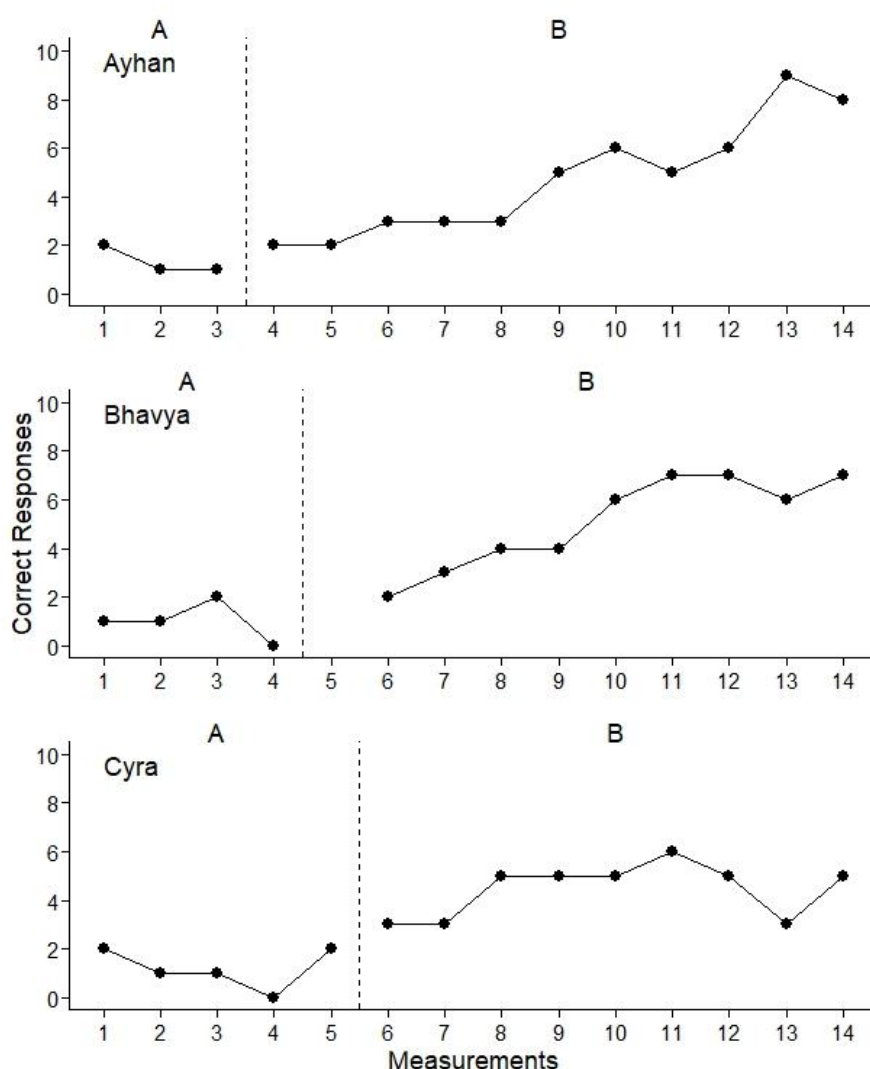


Figure 1: Number of correctly answered comprehension questions

Table 2 provides complementary descriptive statistics for each participant and phase, including sample size (N), measures of central tendency (mean M, median Md), variability (interquartile range IQR), and observed score ranges (minimum and maximum). Together, the graphical displays and descriptive statistics offer

an overview of performance patterns prior to and following the introduction of the self-questioning procedures with prompt fading.

Table 2: Descriptive statistics for correctly answered comprehension questions

Student	N		M		Md		IQR		min		max	
	A	B	A	B	A	B	A	B	A	B	A	B
Ayhan	3	11	1.33	4.73	1	5	1	3	1	2	2	9
Bhavya	4	9	1.00	5.11	1	6	1	2	0	2	2	7
Cyra	5	9	1.20	4.44	1	5	1	2	0	3	2	6

Note. IQR = interquartile range

During baseline (A), all three participants demonstrated consistently low and stable reading comprehension performance, with scores confined to the lower end of the scale (0–2 correct responses emphasized). Central tendency was uniformly low across participants (Ayhan: $M = 1.33$, $Md = 1$; Bhavya: $M = 1.00$, $Md = 1$; Cyra: $M = 1.20$, $Md = 1$), and variability was minimal, as indicated by identical and narrow interquartile ranges ($IQR = 1$ for all). Baseline ranges were restricted (Ayhan: 1–2; Bhavya: 0–2; Cyra: 0–2), and no positive trends were observed, indicating stable but persistently weak performance before instruction.

Ayhan's baseline responding was stable and low ($M = 1.33$, $Md = 1$, $IQR = 1$, range = 1–2). With the introduction of instruction (B), an immediate and sustained increase was observed. Mean performance increased to $M = 4.73$ ($Md = 5$), corresponding to an average performance gain of approximately 256% relative to baseline. This improvement was accompanied by a broader response distribution ($IQR = 3$, range = 2–9). Scores consistently exceeded baseline levels, with later sessions reaching high levels of accuracy, indicating both immediacy of effect and continued improvement over time.

Bhavya showed a similarly constrained baseline ($M = 1.00$, $Md = 1$, $IQR = 1$, range = 0–2). After the phase change, her performance first matched the baseline and then steadily improved. During Phase B, mean performance rose to $M = 5.11$ ($Md = 6$), representing an average increase of approximately 411% relative to baseline. Variability was moderate ($IQR = 2$, range = 2–7), reflecting a gradual but robust acquisition of the targeted strategies, with performance trajectories clearly trending upward and stabilizing at higher levels toward the end of the phase.

Cyra's baseline performance was likewise low and stable ($M = 1.20$, $Md = 1$, $IQR = 1$, range = 0–2). At the onset of Phase B, an immediate level increase occurred, followed by consistently higher scores throughout the instructional phase. Mean performance improved to $M = 4.44$ ($Md = 5$), corresponding to an average gain of approximately 270% relative to baseline. Variability remained relatively limited ($IQR = 2$, range = 2–6), indicating substantial gains accompanied by stable responding over time.

Across participants, all performance trajectories showed clear upward shifts following the phase change, with data paths characterized by positive trends and convergence toward relatively high end-of-phase performance levels. Rather than isolated score increases, the results reflect sustained growth over sessions, with

each student demonstrating improved and increasingly reliable responses at the conclusion of Phase B. Taken together, these patterns are consistent with a strong phase-related enhancement in reading comprehension performance.

To complement visual inspection and descriptive analyses, two nonparametric overlap-based indices, Percentage Exceeding the Trend (PET) and Tau-U with baseline trend adjustment (Tau-U_BA), were calculated. Consistent with WWC (Maggin et al., 2021) and SCRIBE (Tate et al., 2016) recommendations, these indices were used to quantify instructional effects while acknowledging that causal inferences in single-case research are primarily established through visual analysis. PET and Tau-U were selected because they are robust to small sample sizes, do not rely on distributional assumptions, and are well-suited to time-series data with limited observations.

Importantly, both indices are sensitive to gradual, trend-based changes in performance and are therefore appropriate in contexts in which effects are expected to emerge progressively rather than as abrupt level shifts. PET indexes the extent to which performance during the instructional phase exceeds values predicted from baseline trend, whereas Tau-U_BA estimates phase nonoverlap while controlling for baseline trend.

PET values were 100% for Ayhan, 90% for Bhavya, and 100% for Cyra, indicating that most intervention data points exceeded baseline trend projections. Corresponding Tau-U_BA values were 0.90 for Ayhan, 0.79 for Bhavya, and 0.67 for Cyra, suggesting that between approximately two-thirds and nine-tenths of all baseline-intervention comparisons favored intervention performance after baseline-trend adjustment.

These indices provide converging quantitative evidence in support of the patterns observed in visual analysis. Ayhan showed the most consistent improvement, Bhavya exhibited strong but slightly attenuated effects consistent with a delayed response, and Cyra demonstrated reliable gains with greater variability. In line with WWC guidance, effect size estimates are interpreted as complementary evidence. They should be considered in conjunction with visual, descriptive, and design features when evaluating the strength of the functional relation.

As a supplementary model-based analysis, piecewise regression models were estimated at both the case and group levels to further examine changes in level and trend across phases. At the case level, results were consistent with visual inspection, indicating either immediate level changes (Cyra) or positive within-intervention growth (Ayhan and Bhavya), while baseline trends were uniformly non-significant. Model fit was strong across participants (adjusted $R^2 = .68-.90$), suggesting that the segmented models captured a substantial proportion of variance in performance over time.

At the group level, a hierarchical piecewise model indicated a significant positive slope during Phase B ($B = 0.574$, $p = .003$), reflecting reliable improvement over the course of intervention across participants, whereas the average level change

at phase transition was not significant ($B = 0.897$, $p = .169$). Given the small number of cases, this regression results are interpreted cautiously and serve to corroborate, rather than supplant, the conclusions drawn from visual analysis and overlap-based indices.

All three participants reported high levels of social validity. Total scores were uniformly elevated (Ayhan = 24; Bhavya = 22; Cyra = 23), indicating strong endorsement of the intervention across dimensions. Across items, students reported that (a) the intervention helped them understand texts better, (b) they felt more confident in their reading abilities, and (c) performance feedback supported their learning. Ayhan, Bhavya, and Cyra all indicated that they would recommend the program to peers with similar difficulties. Two children reported positive engagement during the instructional phase, while the third described increased interest relative to baseline and moderate motivation to continue using the approach. Overall, these results suggest that the prompt-fading-based self-questioning approach was not only effective but also perceived as feasible and beneficial from the students' perspective.

4. Discussion

4.1 Main Findings

The intervention yielded consistent improvements in reading comprehension across all three participants, thereby directly addressing the first research question. Students who entered the study with adequate decoding skills but pronounced difficulties in understanding text showed marked gains from baseline to intervention. Median comprehension scores increased sharply, maximum scores rose considerably, and session-to-session performance shifted from uniformly low and stable responding during baseline to consistently higher levels during intervention, indicating improvements that were unlikely to be attributable solely to incidental fluctuation.

Clear and replicable changes in level and trend reading comprehension performance were observed across participants, directly addressing the second research question. A visual inspection showed that two participants' levels went up right away and that the third participant's level went up slowly but steadily. All of the intervention data series were higher than the baseline levels. Variability increased modestly during this phase, reflecting an expansion of performance ranges rather than instability. Nonparametric overlap-based indices (PET and Tau-U with baseline trend adjustment) confirmed these patterns by showing strong effects across cases. Piecewise regression analyses also showed positive growth within the intervention at both the individual and group levels. Taken together, these findings support a functional relationship between the introduction of a prompt-fading-based self-questioning approach and improvements in reading comprehension performance.

With respect to the third research question, the findings indicate high levels of social validity. All participants evaluated the instructional approach as advantageous and acceptable, reported increased confidence in their reading abilities, and indicated that it supported their understanding of texts. Two

participants reported enjoyment during the intervention, while the third described increased enjoyment relative to baseline and moderate interest in continued use. Overall, these results suggest that the prompt-fading-based self-questioning approach was associated with positive student perceptions and may represent a feasible and beneficial instructional option for learners with similar characteristics.

4.2 Limitations

When interpreting the findings of this study, one should consider several limitations. First, the investigation employed a single-case experimental design with a small number of participants, which necessarily limits the generalizability of the results. Although the multiple-baseline design provides strong internal validity through replicated phase changes across individuals, conclusions are restricted to students with profiles similar to those included in this study, namely third graders with adequate decoding skills but pronounced difficulties in reading comprehension. To determine the broader applicability of the findings, researchers should replicate the study with larger samples and across different educational contexts.

Second, the study focused exclusively on short-term effects. No follow-up or maintenance data were collected to examine whether gains in reading comprehension persisted once the instructional supports were withdrawn. Consequently, it remains unclear to what extent students maintained independent self-questioning behaviors over time or generalized these strategies to classroom reading tasks and curricular materials.

Third, instruction was delivered individually in a controlled setting by trained researchers. While this method ensured high treatment fidelity and minimized contextual variability, it does not fully reflect typical classroom conditions, where teachers must manage multiple students simultaneously and instructional time is constrained. Therefore, the question of how effectively the prompt-fading approach can be implemented in small-group or whole-class formats remains open. In addition, students received small rewards for stable or improved performance. Consequently, the effects of reinforcement cannot be fully separated from those of the intervention itself. Future research should examine the extent to which outcomes can be replicated under different reinforcement conditions.

Finally, although multiple complementary analytic approaches were employed, including visual analysis, overlap indices, and regression-based models, causal inferences in single-case research ultimately rest on design logic and replication rather than statistical significance. All quantitative analyses should thus be interpreted as supportive rather than definitive evidence.

4.3 Practical Implications

Despite these limitations, the findings have several practical implications for instruction and intervention. The results suggest that systematic prompt fading represents a feasible and instructionally efficient mechanism for supporting students with reading comprehension difficulties. By gradually reducing external supports, the approach aligns closely with established principles of explicit

instruction and the gradual release of responsibility, while simultaneously fostering independent strategic use.

From a practical standpoint, the approach requires relatively few materials and can be integrated into existing reading instruction with modest preparation. The structured progression from fully supported questioning to increasingly autonomous self-questioning offers a clear framework for scaffolding strategy use while systematically reducing the need for continuous instructional support. Although the present study was conducted under highly controlled, individualized conditions, the consistent effects observed suggest that prompt fading may represent a particularly robust instructional component. This raises the possibility that the core procedures could be adapted to more scalable formats, such as peer tutoring or video-based modeling, in which prompts and their gradual withdrawal can be implemented with limited ongoing teacher involvement.

Importantly, the positive social validity ratings indicate that students perceived the intervention as advantageous and acceptable. This is a critical consideration for classroom adoption, as instructional approaches that are rejected by learners are unlikely to be sustained, regardless of their measured effectiveness.

4.4 Future Research

Future research should extend the present findings in several directions. First, we need replication studies with larger samples and more diverse learner populations to establish the robustness of prompt fading as a reading comprehension intervention. In particular, examining its effectiveness with multilingual learners, students with co-occurring language impairments, and students at different grade levels would help clarify boundary conditions.

Second, future studies should incorporate maintenance and generalization measures to determine whether students continue to apply self-questioning strategies independently and whether gains transfer to authentic classroom reading tasks. Investigating implementation in small-group or whole-class formats would further enhance ecological validity and inform instructional decision-making within multitiered systems of support.

Third, comparative studies examining prompt fading relative to other forms of scaffolded comprehension instruction could help isolate its unique contribution and determine whether it offers added value beyond traditional guided practice.

4.5 Conclusion

The present study provides preliminary but compelling evidence that a structured self-questioning intervention with systematic prompt fading can improve reading comprehension performance among elementary students at risk for learning difficulties. By combining strong theoretical grounding with high social acceptability, prompt fading emerges as a promising instructional mechanism worthy of further investigation and cautious application in educational practice.

5. Declaration on the Use of Artificial Intelligence

ChatGPT (OpenAI) was used solely to improve the linguistic quality and readability of the manuscript. The authors developed and verified all ideas, interpretations, analyses, and conclusions, taking full responsibility for the content of the paper.

6. References

- Anderson, R. C., Chaparro, E. A., Smolkowski, K., & Cameron, R. (2023). Visual thinking and argumentative writing: A social-cognitive pairing for student writing development. *Assessing Writing*, 55, Article 100694. <https://doi.org/10.1016/j.asw.2023.100694>
- Budianto, A., Setyosari, P., Kuswandi, D., & Ulfa, S. (2022). Summaries writing to enhance reading comprehension: Systematic literature review from 2014 to 2021. *Eurasian Journal of Applied Linguistics*, 8(1), 149–161. <https://doi.org/10.32601/ejal.911526>
- Cartwright, K. B. (2023). *Executive skills and reading comprehension: A guide for educators* (2nd ed.). Guilford Press.
- Cengher, M., Budd, A., Farrell, N., & Fienup, D. M. (2018). A review of prompt-fading procedures: Implications for effective and efficient skill acquisition. *Journal of Developmental and Physical Disabilities*, 30(2), 155–173. <https://doi.org/10.1007/s10882-017-9575-8>
- Connor, C. M. (Ed.). (2016). *The cognitive development of reading and reading comprehension*. Routledge.
- Cremin, T., Hendry, H., Rodriguez Leon, L., & Kucirkova, N. (Eds.). (2023). *Reading teachers: Nurturing reading for pleasure*. Routledge.
- Deitz, S. M., & Malone, L. W. (1985). On terms: Stimulus control terminology. *The Behavior Analyst*, 8(2), 259–264. <https://doi.org/10.1007/BF03392826>
- Delaney, E., McAteer, S., Delaney, M., McHugh, G., & O'Neill, B. (2023). *PIRLS 2021: Reading results for Ireland*. Educational Research Centre.
- Gough, P. B., & Tunmer, W. E. (1986). *Decoding, reading, and reading disability*. *Remedial and Special Education*, 7(1), 6–10. <https://doi.org/10.1177/074193258600700104>
- Hudler, S. E., Hurlburt, G., & Brock, M. E. (2025). A systematic literature review of reading instruction for students with complex communication needs. *Research and Practice for Persons with Severe Disabilities*, 50(4), 305–321. <https://doi.org/10.1177/15407969251333963>
- Jackson, N. E., & Coltheart, M. (2001). *Routes to reading success and failure: Toward an integrated cognitive psychology of atypical reading*. Psychology Press.
- Joseph, L. M., Alber-Morgan, S., Cullen, J., & Rouse, C. (2016). The effects of self-questioning on reading comprehension: A literature review. *Reading & Writing Quarterly*, 32(2), 152–173. <https://doi.org/10.1080/10573569.2014.891449>
- Kilpatrick, D. A., Joshi, R. M., & Wagner, R. K. (Eds.). (2019). *Reading development and difficulties: Bridging the gap between research and practice*. Springer.
- Ledford, J. R., & Gast, D. L. (2024). *Single case research methodology: Applications in special education and behavioral sciences* (4th ed.). Routledge.
- Maggin, D. M., Barton, E., Reichow, B., Lane, K. L. & Shogren, K. A. (2021). Commentary on the What Works Clearinghouse Standards and Procedures Handbook for the review of single-case research. *Remedial and Special Education*, 43(6), 421–433. <https://doi.org/10.1177/07419325211051317>
- Maimunsyah, Rahmat, A., & Nuruddin. (2025). CLIL in reading comprehension learning: A systematic literature review. *TPM: Testing, Psychometrics, Methodology in Applied Psychology*, 32(S6), 1382–1388. <https://tpmap.org/submission/index.php/tpm/article/view/2047>
- Moll, K., & Landerl, K. (2014). *SLRT-II: Lese- und Rechtschreibtest [SLRT-II: Salzburg Reading and Spelling Battery-2]*. Hogrefe.

- OECD (2023). Education at a glance 2023: OECD indicators. OECD Publishing.
- Rouse, C. A., Alber-Morgan, S. R., Cullen, J. M., & Sawyer, M. (2014). Using prompt fading to teach self-questioning to fifth graders with LD: Effects on reading comprehension. *Learning Disabilities Research & Practice, 29*(3), 117–125. <https://doi.org/10.1111/ldrp.12036>
- Rouse-Billman, C., & Alber-Morgan, S. (2019). Teaching self-questioning using systematic prompt fading: Effects on fourth graders' reading comprehension. Preventing School Failure. *Alternative Education for Children and Youth, 63*(4), 352–358. <https://doi.org/10.1080/1045988X.2019.1619508>
- Stuart, M., & Stainthorp, R. (2015). *Reading development and teaching*. SAGE Publications.
- Tate, R. L., Perdices, M., Rosenkoetter, U., McDonald, S., Togher, L., Shadish, W., Horner, R., Kratochwill, T., Barlow, D. H., Kazdin, A., Sampson, M., Shamseer, L., & Vohra, S. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016: Explanation and elaboration. *Archives of Scientific Psychology, 4*(1), 10–31. <https://doi.org/10.1037/arc0000027>
- Vaughn, S., Boardman, A., & Klingner, J. K. (2024). *Teaching reading comprehension to students with learning difficulties* (3rd ed.). Guilford Press.
- Watkins, P. (2018). *Teaching and developing reading skills*. Cambridge University Press.
- Wilbert, J. (2026). *Analyzing single-case data with R and scan*. <https://doi.org/10.5281/zenodo.5713559>
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*(2), 203–214. <https://doi.org/10.1901/jaba.1978.11-203>