

Rethinking Self-Regulated Learning in Online Education: Typologies, Affective Dynamics, and Design Implications from Learner Feedback

Henry Khiat

Abstract: As self-directed online learning becomes more prevalent, understanding how learners self-regulate in these environments is critical for ensuring engagement, persistence, and success. This study examines how self-regulated learning (SRL) behaviours are expressed in authentic learner feedback and explores their associations with emotional tone and satisfaction. A total of 826 anonymised comments were analysed from participants in asynchronous professional trade finance courses. Drawing on established SRL models, the study combined thematic analysis with natural language processing (TF-IDF vectorisation, KMeans clustering), sentiment analysis (TextBlob), and statistical correlations to identify behavioural patterns. Results showed that SRL behaviours varied widely and often clustered into compound strategies. Help-Seeking and Motivational Beliefs were positively correlated with satisfaction, while Planning, Time Management, and Environmental Structuring were linked to frustration, particularly where platform design failed to support learners' regulatory intentions. Sentiment analysis confirmed that emotional tone aligned with regulatory challenges or supports. Six learner typologies and five cluster-derived themes were identified, reflecting distinct SRL profiles and systemic pain points such as unclear structure, feedback gaps, and usability issues. These findings affirm that SRL is not merely an internal trait but a contextually shaped, co-regulated process. Implications include the need for emotionally intelligent, scaffolded platforms tailored to diverse learner profiles, and policy shifts toward feedback-rich, adaptive learning environments. While the study is situated in a specific professional learning context, its integrated methods offer a scalable model for analysing authentic SRL behaviours in digital education. This approach demonstrates how learner feedback can serve as a scalable diagnostic lens for SRL-informed design.

Keywords: self-regulated learning, self-directed learning, online learning, distance education, learner typologies, sentiment analysis, natural language processing, machine learning, adaptive learning design, trade finance education

What is already known about this topic

- SRL is a key predictor of success in online learning, where autonomy is high and external guidance is limited.
- Many platforms lack scaffolds for SRL, such as support for planning, feedback, and motivation.

What this paper contributes

- Identifies six online learner archetypes with distinct SRL patterns and support needs.
- Combines coding, sentiment analysis, and clustering to analyse real learner feedback at scale.
- Shows that sentiment and satisfaction align with SRL, linking emotion and regulation.
- Provides evidence of compound SRL strategies, supporting its cyclical, integrated nature.

Implications for theory, practice, and/or policy

- SRL is co-regulated, shaped by both learner traits and system design features.
- Designers should build adaptive tools like feedback loops, pacing aids, and reflection prompts.
- Dual analysis offers a framework for SRL-informed, emotionally responsive course design.



Introduction

As online learning environments increasingly dominate educational landscapes, self-regulated learning (SRL) has become a critical determinant of learner persistence, performance, and engagement, particularly in asynchronous contexts requiring high autonomy (Pintrich, 2000; Zimmerman & Schunk, 2011; Broadbent & Poon, 2015; Dai, Luan, & Lin, 2023).

Despite its well-established theoretical significance, SRL is inconsistently integrated into online instructional design. Many platforms rely on passive content delivery, lacking embedded scaffolds for metacognitive awareness, time management, and adaptive help-seeking, thereby impeding effective self-regulation and diminishing engagement and satisfaction (Prasse, Sailer, & Fischer, 2024; Broadbent & Poon, 2015).

While traditional SRL research has relied heavily on self-reports or controlled settings—often overlooking authentic manifestations of regulatory processes—recent studies emphasize that SRL is not a fixed personal trait but a contextually contingent, dynamically co-regulated phenomenon (Hadwin, Järvelä, & Miller, 2018; Panadero, 2017; Roll & Winne, 2015; D’Mello, Dieterle, & Duckworth, 2020). Learners’ regulatory capacity depends not only on internal strategies but also on contextual supports—task design, feedback availability, and digital affordances—and on affective states such as motivation, frustration, and confusion that shape planning, monitoring, and reflection.

This interplay between cognition, emotion, and context highlights the need to reconceptualize SRL as a situated, co-regulated phenomenon rather than a solitary psychological process. To address these conceptual and methodological gaps, this study analyses a large corpus of open-ended learner feedback from asynchronous professional courses using a theoretically grounded SRL coding scheme, sentiment analysis, and unsupervised clustering (TF-IDF, KMeans) to reveal how regulatory behaviours and emotions interact in authentic contexts. This approach captures how learners describe their regulatory experiences, frustrations, and support needs in their own words, showing how strategies cluster, how emotional tone relates to satisfaction, and how design misalignments constrain regulatory success.

The next section reviews key developments in SRL theory and its implications for online learning design.

Literature

To contextualize the dual challenge of misaligned instructional design and limited methodological approaches identified in the introduction, this section synthesizes key developments in three interconnected domains: (1) theoretical models of self-regulated learning (SRL), (2) SRL within digital and online instructional design, and (3) emerging data-driven and computational approaches to studying SRL in authentic learning contexts.

Theoretical Foundations of Self-Regulated Learning

Self-regulated learning (SRL) is a cornerstone of educational psychology and a well-established predictor of academic success across contexts. Zimmerman’s (2000) cyclical model, comprising forethought, performance, and self-reflection, remains the most widely cited conceptualization, emphasizing the proactive and iterative nature of learner regulation. Complementing this, Winne and Hadwin’s (2008) cognitive and computational model highlights the dynamic interplay between monitoring, control, and metacognitive strategies as learners interact with instructional environments. Together, these models underscore SRL as both a process and a system, integrating cognitive, motivational, and behavioural dimensions.

Subsequent syntheses have deepened this perspective. Panadero's (2017) meta-review of six major SRL models identified common core elements—goal setting, strategy use, monitoring, and reflection—while underscoring the affective and motivational underpinnings of regulation. Efklides' (2011) MASRL framework further positioned emotion as central to metacognitive functioning, arguing that affective states such as frustration, curiosity, or satisfaction modulate the learner's ability to self-monitor and adapt. These theoretical advancements establish SRL as a multidimensional and dynamic system, where cognition, motivation, and emotion jointly shape learning outcomes.

SRL in Online and Digital Learning Contexts

As learning increasingly migrates to asynchronous and self-paced platforms, SRL has become a decisive factor in learner success. Meta-analytic and empirical evidence consistently supports this association: Broadbent and Poon (2015) found that time management, metacognitive regulation, and effort control are among the strongest predictors of online achievement; Yu (2023) and Dai, Luan, and Lin (2023) similarly demonstrated that SRL readiness and preparatory scaffolding substantially influence online performance and persistence.

However, translating SRL theory into practice within digital environments remains inconsistent. Many platforms still prioritize content delivery over metacognitive support, offering limited tools for planning, feedback, or reflection (Prasse, Sailer, & Fischer, 2024). Learners often face distractions and fragmented attention (Martin, Long, Haywood, & Xie, 2025; Wang, Salisbury-Glennon, Dai, Lee, & Dong, 2022), which disrupt their ability to sustain regulation. These findings highlight a design paradox: while online learning demands high autonomy, most systems under-scaffold the very skills required for such autonomy.

Adult and professional learners face particular challenges. Khat (2017) argues that self-direction functions as a developmental precursor to SRL, requiring explicit guidance in goal setting, reflection, and feedback integration. Without these supports, autonomy becomes a burden rather than a facilitator of learning. This perspective aligns with Hadwin, Järvelä, and Miller's (2018) notion of socially shared and co-regulated learning, which extends SRL beyond the individual to collaborative and professional learning contexts. Together, these perspectives underscore that SRL is not merely an internal competence but a relational and context-dependent process, mediated by both design and social interaction.

Data-Driven and Computational Approaches to SRL Research

The methodological evolution of SRL research reflects an increasing emphasis on authentic, data-rich representations of regulation. Traditional self-report surveys and controlled laboratory tasks, while informative, often fail to capture the spontaneous, situated manifestations of SRL in real learning environments (Panadero, 2017). In response, scholars have called for ecologically valid methodologies capable of capturing how regulation unfolds dynamically and contextually (Greene & Azevedo, 2010; Järvelä et al., 2019).

Recent advancements in learning analytics and artificial intelligence have opened new pathways. Steinert et al. (2023) demonstrated that large language models can function as real-time formative feedback agents, supporting learners' metacognitive reflection and motivation. Similarly, Chatti, Yücepur, Muslim, Guesmi, and Joarder (2023) emphasize the potential of personalized, analytics-driven feedback systems to scaffold autonomy and adapt to individual regulatory patterns. Roll and Winne (2015) further highlight how digital trace analytics can shift SRL research from measurement to active scaffolding, embedding metacognitive prompts directly within learning systems.

Emerging multimodal and computational approaches extend this trajectory. Lämsä et al. (2025) and Alghamdi, Bull, and Kharrufa (2025) advocate for the integration of text analytics, digital traces, and

behavioural data to model regulation at scale. Multimodal studies by Järvelä, Malmberg, Haataja, Sobocinski, and Kirschner (2019) show how eye-tracking, discourse, and physiological data reveal the complex, recursive nature of SRL processes. These developments underscore a paradigm shift, from static, introspective measures toward dynamic, machine-assisted analyses that capture regulation as it naturally occurs.

Positioning the Present Study

Building on this trajectory, the current study adopts a computational mixed-methods approach that combines SRL-grounded qualitative coding with natural language processing (NLP), sentiment analysis, and unsupervised clustering. By analyzing authentic learner feedback rather than experimental or self-reported data, this research contributes to the emerging field of ecologically valid SRL analytics. It advances existing scholarship in three ways:

1. Theoretical integration: framing SRL as a co-regulated, affective, and contextual phenomenon rather than a purely cognitive trait.
2. Methodological innovation: leveraging computational text analysis to identify compound SRL strategies and emotional dynamics in authentic learner discourse.
3. Design relevance: translating these analytical insights into practical frameworks for adaptive and emotionally responsive online learning environments.

Through this synthesis, the study bridges the theoretical richness of SRL research with the pragmatic demands of digital learning design, providing both a methodological model for large-scale SRL analysis and a conceptual foundation for developing more responsive, data-informed learning systems.

Theoretical Background

Grounded in these theoretical advancements and practical implementation gaps, this study draws on established models of self-regulated learning (SRL) to frame its investigation. Zimmerman's cyclical model (2000) outlines three interrelated phases of forethought, performance, and self-reflection, while Winne and Hadwin's computational model (2008) highlights the interactive roles of monitoring, control, and metacognitive strategies. Together, these frameworks illustrate how learners set goals, deploy strategies, monitor progress, and evaluate outcomes within dynamic learning environments. Panadero (2017) further synthesized six SRL models and identified converging elements across them, while Efklides' (2011) MASRL framework added the critical layer of motivational and affective processes, highlighting the role of emotions in sustaining regulation.

Building on these foundations, SRL is conceptualized in this study as a set of seven interlinked components: Forethought (Planning), Performance (Monitoring and Strategy Use), Self-Reflection (Evaluation), Motivational Beliefs, Help-Seeking and Resource Use, Time Management, and Environmental Structuring. These categories, drawn from the SRL literature, provide the theoretical basis for coding learner feedback and serve as the framework for analysing the authentic manifestations of SRL in online environments. This categorization also resonates with Hadwin, Järvelä, and Miller's (2018) notion of regulation as both an individual and socially shared phenomenon, acknowledging that even in self-directed environments, help-seeking and collaborative elements are integral to the regulatory cycle.

The central aim of this study is to investigate how self-regulated learning (SRL) behaviours are expressed, experienced, and evaluated by learners within self-directed online learning environments, and how these behavioural patterns relate to learners' emotional sentiment, satisfaction, and design needs. Recognizing that SRL is both a personal and contextual process, this study adopts a mixed-methods approach, integrating thematic coding, correlation analysis, sentiment detection, typology

development, and unsupervised clustering, to provide a multidimensional account of learner regulation and experience.

To address this aim, the study is guided by the following eight research questions (RQs), structured to move from identification to explanation to practical application:

RQ1: How do learners' SRL behaviours manifest in open-ended feedback?

This RQ explores the ways in which the seven core SRL components are represented in learners' own words.

RQ2: What interrelationships exist between different SRL components, and how do they reflect compound regulatory strategies?

This RQ investigates how different SRL behaviours tend to co-occur, forming integrated patterns of regulation.

RQ3: How do sentiments (positive, neutral, negative) align with expressed SRL behaviours?

This RQ explores how learners' emotional tone varies across different aspects of self-regulation, revealing areas of frustration or satisfaction.

RQ4: How are SRL behaviours and sentiments associated with learner satisfaction scores?

This RQ examines the relationship between what learners express (cognitively and emotionally) and how they rate their learning experience.

RQ5: What distinct learner typologies emerge from SRL-themed feedback, and how do these archetypes reflect differing self-regulation challenges and support needs in self-directed online learning environments?

This RQ identifies categories of learners based on their SRL orientations, highlighting the diversity of learner profiles and their support needs.

RQ6: To what extent do learner typologies correlate with specific self-regulated learning (SRL) components, and how do these associations validate the behavioural underpinnings of each archetype?

This RQ explores whether the identified learner types are meaningfully linked to specific SRL behaviours.

RQ7: What latent themes can be uncovered through broader patterns in learner feedback, and how do these themes reinforce or challenge the SRL components and learner typologies identified in previous analyses?

This RQ validates and complements prior interpretations by surfacing common learner concerns that span across individual differences.

RQ8: How can SRL-based learner typologies and feedback themes inform practical instructional design strategies in self-directed online learning environments?

This RQ translates the findings into actionable insights for course designers, with an emphasis on learner variability and support mechanisms.

Together, these research questions follow a progressive arc: beginning with learner-expressed SRL behaviours (RQ1–2), integrating emotional and experiential dimensions (RQ3–4), identifying regulatory archetypes (RQ5–6), validating themes through unsupervised learning (RQ7), and informing learner-responsive design strategies (RQ8). This structure supports both theoretical development and pedagogical relevance, advancing our understanding of SRL as it occurs in real-world online learning environments.

Methodology

This study adopts a mixed-methods design that integrates qualitative analysis, natural language processing (NLP), and quantitative correlation techniques to examine patterns of self-regulated learning (SRL) in learner feedback. The integration of human-coded and machine-driven approaches enabled both depth and breadth in capturing the cognitive, behavioural, and affective dimensions of SRL.

Research Design

The research followed a five-phase design. First, a pedagogically grounded SRL coding scheme was applied to classify learner comments into seven core components: Forethought (Planning), Performance (Monitoring and Strategy Use), Self-Reflection (Evaluation), Motivational Beliefs, Help-Seeking and Resource Use, Time Management, and Environmental Structuring. Second, heuristic keyword expansion and clustering techniques refined the categorization into more nuanced subcategories. Third, sentiment polarity and intensity were analysed to uncover emotional dimensions. Fourth, statistical correlations between SRL categories, sentiment, and satisfaction scores were computed. Finally, findings were synthesized into practical and theoretical implications for online learning design.

Data Collecting Tools

Data for this study were drawn from anonymized, open-ended survey responses collected through post-course evaluation forms administered by the case educational institution. These evaluation forms are a standard part of its quality assurance and learner engagement processes, distributed immediately after course completion to ensure that feedback reflects learners' most recent experiences.

The primary data consisted of qualitative comments in response to open-ended questions that invited learners to describe their experiences, challenges, and suggestions for improvement. This format allowed participants to articulate their perceptions in their own words, offering a rich and authentic account of their regulatory strategies, motivational beliefs, and emotional reactions. Unlike fixed-response items, open-text feedback provided rich insights into how learners navigated both the content and the digital learning environment, making it especially suitable for examining self-regulated learning (SRL) processes.

In addition to qualitative comments, the dataset included quantitative satisfaction ratings on a 10-point Likert-type scale, where 1 indicated "very dissatisfied" and 10 indicated "very satisfied." These ratings served two purposes: first, as a measure of overall learner satisfaction, and second, as a variable against which textual features (e.g., regulatory strategies, emotional tones) could be statistically correlated. By linking free-text comments to satisfaction scores, the study was able to explore not only what learners reported but also how these reports related to their perceived outcomes.

The combination of open-ended qualitative feedback and structured quantitative ratings created a mixed-methods dataset that captured both the depth of learner experiences and the breadth of their evaluative judgments. This dual structure provided a strong basis for applying computational text analysis techniques, as it allowed linguistic and affective markers in the comments to be directly connected to satisfaction outcomes, thereby enhancing the interpretive power of the analysis.

Sampling or Study Group

The dataset comprised 826 anonymized open-ended responses from learners who participated in asynchronous online training modules delivered by the case educational institution between 2022 and 2024. These courses are globally recognized professional development programs, designed to build applied knowledge in areas such as trade finance, compliance, risk management, and international business operations.

Participants represented a diverse professional audience, spanning both early-career practitioners seeking foundational knowledge and experienced senior professionals aiming to update or broaden their expertise. This diversity created meaningful variation in learning needs, regulatory strategies, and expectations of digital learning platforms. For example, younger professionals often approach training with a focus on skill acquisition and credentialing, while senior participants tend to emphasize applicability to complex, real-world scenarios.

The study group also reflected the international reach of the case educational institution drawing participants from multiple geographical regions across Asia, Europe, Africa, and the Americas. This global spread ensured that the dataset captured a wide spectrum of cultural and organizational perspectives on learning, self-regulation, and online engagement. Such diversity strengthens the generalizability of the findings, as regulatory behaviours and satisfaction drivers can differ significantly across professional and cultural contexts.

By focusing on learners enrolled in asynchronous, self-directed modules, the study targeted a population for whom self-regulated learning (SRL) is particularly critical. Unlike synchronous or instructor-led environments, these modules place greater responsibility on learners to set goals, monitor progress, manage time, and seek resources independently. The sample was therefore ideally suited to exploring how SRL processes manifest in authentic learner feedback and how they align with satisfaction and affective outcomes.

Data Analysis

To address the research objectives, the study adopted a mixed computational approach designed to uncover both the linguistic and emotional dimensions of learner feedback. Traditional self-report surveys often fail to capture the richness and spontaneity of learners' reflections; therefore, more nuanced analytic methods were employed to explore how self-regulated learning (SRL) manifests in authentic comments. By combining natural language processing (NLP) techniques with clustering algorithms and statistical tests, the analysis was able to move beyond surface-level impressions, systematically identifying salient terms, latent themes, emotional tones, and their relationships to satisfaction scores. This multi-layered approach ensured that the dataset was examined from complementary perspectives, allowing both cognitive and affective aspects of SRL to emerge from the learners' own words.

The workflow of analysis proceeded through four main stages. First, TF-IDF vectorization was applied to highlight the most salient lexical features in the learner feedback. Second, KMeans clustering grouped these features into latent themes, surfacing common patterns and learner archetypes. Third, sentiment analysis using the TextBlob library was conducted to classify the emotional tone of each comment. Finally, statistical analyses, including Pearson correlation and frequency distribution, were employed to link linguistic and emotional markers to learners' satisfaction scores. Together, these stages provided a coherent process: beginning with extracting meaningful words, organizing them into themes, identifying emotional tone, and finally validating these patterns against satisfaction outcomes.

Term Frequency–Inverse Document Frequency (TF-IDF) vectorization was first applied to the dataset. This technique assigns a weight to each word based on how often it appears in a particular comment compared to how frequently it occurs across all comments. Words that are common everywhere, such as “course” or “lesson,” are given less importance, while words that occur frequently in some comments but rarely across the entire dataset are highlighted as more meaningful. For example, in one learner's comment—“I think navigation can be improved. It is not very user friendly”—TF-IDF marked “navigation” and “user friendly” as distinctive terms that highlight structural concerns.

KMeans clustering was then used to group the comments into clusters that reflect underlying themes. KMeans works by placing “centers” in the data and then repeatedly assigning each comment to the

nearest center based on its word patterns. Over several iterations, the algorithm adjusts these centers until stable groups form. The result is that comments with similar wording or concerns end up in the same cluster. For instance, comments like “Sometimes it is difficult to get 100% per lesson” and “Sometimes progress is not captured and I need to repeat” were grouped into a progress tracking and assessment cluster, while comments such as “More interaction, it feels very expensive for what it is” and “A community could be created for engagements” aligned into a social support and interaction cluster.

To capture the emotional dimension of the feedback, sentiment analysis was performed using the TextBlob library. TextBlob evaluates each comment for the polarity of its language, scoring it along a scale from negative to positive. For example, “Loss of track record every time I disconnected from the system” was classified as negative, “The content continues to engage my interest with each lesson” was classified as positive, and “A community could be created for engagements” was classified as neutral, since it suggests an idea without expressing strong emotion. This process quantified the affective stance of learners and allowed us to connect their emotions to reported satisfaction levels.

Finally, statistical analyses were conducted using Python-based packages, primarily pandas, NumPy, SciPy, and matplotlib/seaborn. Two key techniques were employed: Pearson correlation and frequency distribution. The Pearson correlation was computed using the Pearson *r* function from the SciPy. Stats module to measure the strength and direction of the relationship between two variables. For instance, we tested whether positive sentiment comments such as “Thanks for this very interesting chapter” were associated with higher satisfaction ratings, and whether negative statements like “Most of the examples are outdated while relevance is critical in this field” corresponded with lower ratings. The frequency distribution analysis was implemented using pandas and collections. Counter (from Python’s standard library), with data visualized through matplotlib or seaborn. This approach counted how often particular words, phrases, or categories appeared in the dataset. For example, terms relating to time management (e.g., “Kindly increase the time frame in which to finish the course”) were flagged as recurrent, showing that this was a common source of concern among learners.

This integrated toolset made it possible to capture both the linguistic nuances of learner expression and the structural patterns embedded in the feedback, providing a comprehensive basis for interpretation. Importantly, these analyses were not conducted in isolation but were explicitly linked to the SRL framework, allowing learner comments to be examined in terms of how they reflected forethought, performance, reflection, motivation, help-seeking, time management, and environmental structuring. This ensured that the findings could be meaningfully interpreted within the broader theoretical landscape of self-regulated learning.

Validity and Reliability

To ensure the rigor and trustworthiness of the findings, multiple strategies were employed to establish construct validity, confirm model coherence, and strengthen reliability.

Construct validity was supported by grounding the analytical framework in established SRL theory. The coding schema was explicitly derived from Zimmerman’s (2000) cyclical model and Winne and Hadwin’s (2008) computational model, both of which provide widely recognized conceptual foundations for categorizing self-regulatory behaviours. By aligning the data analysis directly with these models, the study ensured that categories such as forethought, performance, reflection, and motivational beliefs were not arbitrarily imposed but rather theoretically justified.

Cluster validity was addressed through careful tuning and evaluation of the KMeans algorithm used for unsupervised clustering. Different numbers of clusters (*k*-values) were iteratively tested, and coherence was examined using silhouette analysis, a metric that measures how similar each comment is to its own cluster compared to other clusters. Higher silhouette scores indicated more distinct and internally

consistent groupings, confirming that the resulting clusters represented meaningful latent themes rather than random aggregations.

Reliability was strengthened through methodological triangulation across three complementary approaches: (1) human coding of comments into SRL subcomponents, (2) machine-based clustering of textual data, and (3) automated sentiment analysis of emotional tone. This triangulation reduced the risk that the findings were dependent on a single analytic lens. While TextBlob's lexicon-based sentiment classification presents known limitations in detecting complex emotional nuance, its outputs were interpreted conservatively and cross-validated against qualitative themes to ensure interpretive alignment. For instance, where human coding identified themes of time management, KMeans clustering independently grouped related feedback, and sentiment analysis further revealed the emotional valence attached to these concerns. The convergence of results across methods enhanced confidence in the robustness of the findings. In addition, formal inter-rater reliability was assessed on a stratified random subsample of 165 comments selected to ensure representation across all seven SRL components. A second coder independently recoded the comments using the operational coding definitions and was blinded to the original classifications. Because comments could reflect more than one SRL component, Cohen's kappa was calculated separately for each component using binary presence/absence coding. The coefficients were Environmental Structuring ($\kappa = .645$), Forethought (Planning) ($\kappa = .878$), Help-Seeking and Resource Use ($\kappa = .680$), Motivational Beliefs ($\kappa = .606$), Performance (Monitoring and Strategy Use) ($\kappa = .714$), Self-Reflection (Evaluation) ($\kappa = .553$), and Time Management ($\kappa = .948$). Observed agreement ranged from 84.8% to 98.2%. These results indicate moderate to very strong consistency in applying the coding framework, while also reflecting expected variation across conceptually overlapping components.

Finally, a manual review of the cluster outputs was conducted to ensure thematic accuracy and interpretive reliability. Exemplar comments within each cluster were inspected to verify that the assigned theme accurately reflected the underlying learner concerns. Where inconsistencies were observed, clusters were re-examined and labels refined to maintain alignment with both the data and the theoretical framework. This process provided an additional layer of quality control, ensuring that computational outputs were meaningfully interpreted within their context.

Together, these strategies established a rigorous and transparent analytic process, enhancing both the internal validity of the findings and their credibility for application in online learning design and SRL research.

Research Procedures

The research procedures followed a structured, multi-stage sequence designed to ensure systematic data collection, processing, and interpretation.

Data acquisition and preparation began with the export of anonymized learner feedback and satisfaction scores from the case institution's learning management system (LMS). Records were cleaned to remove duplicates, incomplete responses, and identifiable information to preserve confidentiality. The resulting dataset comprised 826 open-ended comments linked to corresponding 10-point satisfaction ratings, forming the foundation for both qualitative and quantitative analyses. All procedures complied with institutional data protection guidelines; formal ethics approval was not required due to the use of archival, anonymized data.

Qualitative coding was conducted using a schema derived from Zimmerman's (2000) cyclical model and Winne and Hadwin's (2008) computational model of self-regulated learning (SRL). Each comment was coded for the presence of one or more of seven SRL components: forethought (planning), performance/strategy use, self-reflection (evaluation), motivational beliefs, help-seeking and resource use, time management, and environmental structuring. To assess coding consistency, a stratified

random subsample of 165 comments, selected to ensure representation across all seven SRL components, was independently recoded by a second coder who was blinded to the original classifications. Because comments could receive more than one SRL code, inter-rater reliability was assessed separately for each component using binary presence/absence decisions and Cohen's kappa. Reliability coefficients and observed agreement are reported in the Validity and Reliability section. The reliability estimates were calculated from the independent classifications before comparison with the original coding. In the original coding process, ambiguous cases were reviewed iteratively and recoded where required to maintain conceptual fidelity to the SRL framework. This framework provided the theoretical basis for interpreting both human-coded and machine-generated outputs.

Computational text analysis was applied in three phases. First, Term Frequency–Inverse Document Frequency (TF-IDF) vectorization was used to extract salient lexical features from the learner feedback. Second, KMeans clustering grouped comments into latent thematic clusters, such as feedback access, time constraints, or interface usability. Silhouette scores were used to evaluate the internal coherence and distinctiveness of clusters. Third, sentiment analysis using the TextBlob library assigned polarity scores to each comment (positive, neutral, or negative). While TextBlob's lexicon-based approach may miss subtle affective nuances, outputs were conservatively interpreted and cross-checked against qualitative codes for alignment.

Integration and interpretation of these outputs followed a theory-driven approach. Themes identified via clustering and sentiment scores were analysed in relation to the seven SRL categories to explore how learners' expressed behaviours and emotions mapped onto known regulatory processes. This cross-comparison enabled triangulation across analytical methods.

Quantitative analysis was performed using Pearson correlation to examine relationships between SRL behaviours, sentiment polarity, and satisfaction ratings. Frequency distributions of SRL-coded behaviours and common themes were also analysed to identify dominant patterns and potential pain points in the learner experience.

Synthesis and typology construction comprised the final stage. By integrating SRL codes, sentiment profiles, thematic clusters, and satisfaction scores, a set of learner archetypes was inductively developed. These typologies captured distinct self-regulation profiles and associated emotional patterns, such as strategic seekers or frustrated navigators. Each typology was translated into actionable design recommendations for self-directed online learning environments, ensuring practical relevance alongside theoretical insight.

Limitations

While this study offers valuable insights into how self-regulated learning (SRL) is expressed and experienced in online professional education, several limitations—grouped by theme—should be acknowledged.

Data Characteristics and Scope

The study relied on post-course open-ended learner feedback, which captures retrospective reflections rather than real-time SRL processes. As such, spontaneous or unconscious regulatory behaviours such as in-the-moment strategy shifts or emotional fluctuations, may be underrepresented. In addition, the voluntary nature of survey responses introduces potential response bias, possibly overrepresenting highly satisfied or dissatisfied participants while underrepresenting those with neutral or disengaged views. Finally, sentiment analysis was conducted using TextBlob, a lexicon-based tool that may not fully capture nuanced emotional tones, such as mixed frustration or culturally specific expressions of dissatisfaction.

Methodological Considerations

Although the study employed a triangulated design, combining SRL coding, sentiment analysis, and unsupervised clustering, some degree of interpretive subjectivity was inevitable, particularly in labelling clusters and aligning them with SRL components. Formal inter-rater reliability for the SRL coding was generally satisfactory but varied across components, indicating that some conceptual boundaries remain more open to interpretation than others. Findings involving the lower-agreement components should therefore be interpreted with appropriate caution. While safeguards such as theoretical grounding and manual validation were applied, this interpretive layer may influence thematic boundaries. Moreover, the absence of multimodal data (e.g., clickstream logs, eye-tracking, or biometric signals) limits the ecological depth of SRL behaviour tracking. Future studies should consider real-time or process-tracing methods to complement text-based feedback.

Contextual Boundaries and Transferability

Participants in this study were predominantly professionals enrolled in trade finance-related online courses, which limits the generalisability of findings to other learner groups such as undergraduates, novices, or learners in non-financial fields. Regulatory strategies such as pragmatic goal-setting or time management under work pressures may be more salient in this demographic than in others. Additionally, the study did not incorporate data on co-regulated or socially shared learning processes, which are increasingly recognised as essential in digital and professional learning settings.

Findings and Discussions

This section presents the findings of the study in alignment with the eight research questions, moving from component-level manifestations of SRL behaviours (RQ1–RQ2), to their affective and evaluative associations (RQ3–RQ4), to learner-level variation through typologies (RQ5–RQ6), and finally to cross-cutting latent themes and design implications (RQ7–RQ8). Each set of results is interpreted in relation to established SRL theory and contextualised by the affordances and constraints of the learning environment. In doing so, the section not only reports observed patterns in the data but also situates these patterns within broader debates on how SRL is enacted, experienced, and supported in self-directed online learning contexts.

RQ1: How do learners' SRL behaviours manifest in open-ended feedback?

Analysis of 826 open-ended learner comments revealed that self-regulated learning (SRL) behaviours were prominently expressed across three core components: Help-Seeking and Resource Use (33.2%), Forethought (Planning) (21.1%), and Self-Reflection (Evaluation) (16.8%). These categories, derived from well-established SRL models (Zimmerman, 2000; Winne & Hadwin, 2008), collectively accounted for over 70% of the coded dataset, indicating that learner feedback frequently engages with strategic, preparatory, and reflective dimensions of regulation.

A deeper examination of subcategories within these components surfaced five dominant concerns, each representing a distinct behavioural expression of SRL:

Table 1. Top five learner concerns

Top Concern	SRL Component	Subcategory	Representative Quote
Materials and Resources	Help-Seeking and Resource Use	Materials/Resources	"Have unlimited access to all the study materials at all times."

Top Concern	SRL Component	Subcategory	Representative Quote
Question-Level Explanation Access	Forethought (Planning) / Help-Seeking	Answer Review / Explanation	"Need more explanation for wrong answers."
Post-Assessment Performance Feedback	Self-Reflection (Evaluation)	Overall Content / Assessment Evaluation	"Would be great to get feedback on quiz attempts to know mistakes."
Content Relevance and Clarity	Environmental Structuring / Forethought	Content Relevance and Clarity	"Some parts of the content were unclear or not well-explained."
Guidance Requests	Help-Seeking and Resource Use	Guidance Request	"Should have clearer navigation and more direct support."

These five concerns demonstrate how learners actively attempt to regulate their learning environments and processes, engaging in behaviours that align with key phases of the SRL cycle:

- **Help-Seeking and Resource Use:** Evident in high volumes of requests for downloadable materials, clearer instructions, and just-in-time support. This suggests a proactive attempt to compensate for perceived instructional gaps by mobilizing external resources, consistent with strategic regulation models (Winne & Hadwin, 2008).
- **Forethought (Planning):** Reflected in concerns about ambiguous case studies, insufficient goal clarity, and lack of explanatory guidance prior to assessments. These comments reveal how learners seek anticipatory support to align expectations with instructional intent—a central tenet of the SRL forethought phase (Zimmerman, 2000).
- **Self-Reflection (Evaluation):** Manifested through calls for meaningful post-assessment feedback, pointing to a desire to evaluate performance, identify mistakes, and inform future learning strategies. This behaviour aligns with metacognitive monitoring and reflective recalibration.

Critically, these expressions of SRL were not siloed. Many concerns straddled multiple components, highlighting the interdependence of SRL processes. For instance, "Need more explanation for wrong answers" was coded under both Forethought (indicating unmet anticipatory needs) and Help-Seeking (indicating reliance on feedback mechanisms). This blurring of boundaries reflects how learners dynamically adapt strategies in response to evolving task demands, a behaviour also highlighted in SRL process-tracing literature (Hadwin et al., 2018).

Moreover, the findings challenge traditional portrayals of SRL as a stable learner attribute or isolated strategy. Instead, they suggest that SRL is context-sensitive and environmentally modulated, with learners articulating regulatory needs in response to perceived breakdowns in clarity, pacing, scaffolding, or access.

From a design perspective, these insights point toward the need for SRL-sensitive affordances that go beyond static instructional content. Learners require embedded tools, such as interactive answer rationales, feedback dashboards, and navigational supports, that respond dynamically to their regulatory needs. The convergence of SRL categories in feedback also reinforces the call for integrated design features that can simultaneously support planning, help-seeking, and reflection rather than treating these as separate instructional targets.

RQ2: What interrelationships exist between different SRL components, and how do they reflect compound strategies?

To explore the interconnectedness of learners' self-regulated learning (SRL) behaviours, a comprehensive correlation analysis was conducted among seven SRL components derived from the coded dataset: Forethought (Planning), Performance (Monitoring and Strategy Use), Self-Reflection (Evaluation), Motivational Beliefs, Help-Seeking and Resource Use, Time Management, and Environmental Structuring. Pearson correlation coefficients were computed to assess co-occurrence patterns and detect statistically meaningful behavioural pairings. This approach aimed to identify compound regulatory strategies—bundled patterns of SRL behaviours that tend to co-manifest in learner feedback, thereby extending prior work on SRL cycles from isolated dimensions to dynamic, interactive configurations (Zimmerman, 2000; Winne & Hadwin, 2008).

The table shows Pearson correlation coefficients between the seven core SRL components:

Table 2. SRL Component Correlation Matrix

	Forethought	Performance	Self-Reflection	Motivational Beliefs	Help-Seeking	Time Management	Environmental Structuring
Forethought	1.00	0.21	0.25	0.18	0.30	0.47	0.12
Performance	0.21	1.00	0.31	0.19	0.22	0.26	0.10
Self-Reflection	0.25	0.31	1.00	0.24	0.28	0.20	0.15
Motivational Beliefs	0.18	0.19	0.24	1.00	0.36	0.14	0.11
Help-Seeking	0.30	0.22	0.28	0.36	1.00	0.33	0.09
Time Management	0.47	0.26	0.20	0.14	0.33	1.00	0.13
Environmental Structuring	0.12	0.10	0.15	0.11	0.09	0.13	1.00

The strongest positive correlation emerged between Forethought (Planning) and Time Management ($r = 0.47$). This robust association suggests that learners who actively engage in goal-setting, sequencing, and anticipatory thinking are also those who manage their time and pacing deliberately. Supporting qualitative evidence includes comments such as:

- *"The course flow needs to be more logical and segmented."*
- *"Kindly increase the time frame to complete the course."*

Such comments reveal how temporal and structural clarity are intertwined in learners' mental models of task preparation, underscoring the interdependence of cognitive and behavioural planning mechanisms. This finding aligns with Zimmerman's (2000) cyclical SRL model, which emphasizes the co-functionality of planning and time regulation during the preparatory phase of learning.

A second moderate correlation was observed between Help-Seeking and Resource Use and Motivational Beliefs ($r = 0.36$). This relationship highlights the affective drivers behind resource engagement: learners with higher levels of confidence, persistence, and emotional investment are more likely to request explanations, use supplementary materials, or seek clarification. Representative comments include:

- *"The explanation for wrong answers would help boost my confidence."*
- *"I really enjoyed the course but needed more clarification on some cases."*

These expressions resonate with findings by Cho and Heron (2015), who observed that motivation significantly predicts help-seeking behaviour in digital learning contexts. The co-activation of affective and behavioural regulation suggests that positive emotional states not only sustain engagement but also

catalyse strategic outreach for support, a dynamic often overlooked in SRL models that prioritize cognitive control.

A third meaningful correlation was found between Performance Monitoring and Self-Reflection (Evaluation) ($r = 0.31$), indicating that learners who track their progress during learning also tend to engage in post-task reflection and performance recalibration. Illustrative feedback includes:

- *“The module helped me evaluate how I apply concepts in real life.”*
- *“It’s helpful to track which parts I got wrong and why.”*

This correlation supports cyclical SRL models (Winne & Hadwin, 2008), where monitoring and reflection are not discrete events but reciprocally reinforcing processes. Learners iteratively evaluate their strategy effectiveness and adjust behaviours, indicating a deepening metacognitive maturity.

In contrast, Environmental Structuring showed weak or near-zero correlations with all other SRL components, suggesting a functional independence of this dimension. Rather than integrating into metacognitive cycles, environmental management appears as a logistical or infrastructural concern, particularly in online contexts. Learners voiced frustration in statements like:

- *“It was difficult to know if I had completed everything.”*
- *“The interface should clearly mark what’s done and what’s pending.”*

These remarks point to platform-level usability issues—distinct from higher-order regulation—and signal a need for instructional designers to distinguish interface clarity from pedagogical scaffolding.

To supplement the correlation analysis and further explore the multidimensional clustering of SRL behaviours, unsupervised machine learning techniques were applied to a subset of general or ambiguous learner comments. Term Frequency–Inverse Document Frequency (TF-IDF) vectorization followed by KMeans clustering revealed five latent behavioural themes, each linked to multiple SRL components:

Table 3. Five latent themes

Cluster Theme	Aligned SRL Components	Representative Comments
Content Clarity	Forethought, Self-Reflection	“Please improve the clarity of explanations in the answers.”
Technical Usability	Environmental Structuring	“Platform always shows incomplete despite completion.”
Resource Access	Help-Seeking and Resource Use	“Make more downloadable materials and references available.”
Structure/Flow	Forethought, Time Management	“The sequence of lessons feels disorganized and needs reordering.”
Engagement	Motivational Beliefs, Help-Seeking	“Felt very isolated during the course. Maybe add a discussion forum.”

These unsupervised clusters corroborated the correlation-based findings, providing convergent validity from a computational perspective. They show that SRL components do not exist in silos but converge around recurring instructional challenges, such as explanation clarity, structural navigation, and emotional engagement. Each theme blends cognitive, behavioural, and affective dimensions, illustrating that SRL in practice is not a linear checklist but an adaptive orchestration of strategies triggered by contextual affordances and constraints.

Taken together, both the statistical and machine learning analyses suggest that self-regulated learning in online contexts frequently manifests as compound strategies, with learners concurrently engaging in planning, help-seeking, motivation, and reflection when affordances permit. These behaviours appear mutually reinforcing, particularly where platform design supports real-time strategy execution (e.g., clear course flow, embedded feedback, accessible support channels).

Conversely, components such as Environmental Structuring tend to operate outside this integrative loop, likely because they reflect external friction (technical barriers, interface confusion) rather than internal regulatory dynamics. This distinction is critical for SRL-sensitive course design: while compound strategies can be supported through instructional scaffolds, infrastructural challenges require interface-level solutions.

In sum, this analysis affirms that SRL is not simply a set of discrete skills, but a dynamic system of co-activated strategies. To enable these, instructional designers must focus on integrative scaffolds—dashboards that unify feedback, tools that support reflection and planning, and motivation-sensitive interventions—that mirror the interdependent nature of learner regulation in digital environments.

RQ3: How do sentiments (positive, neutral, negative) align with expressed SRL behaviours?

To examine the relationship between emotional tone and self-regulated learning (SRL) processes, sentiment analysis was applied to all 826 learner comments using the TextBlob Python library. Each comment was assigned a polarity score and categorised into three sentiment classes: Positive, Neutral, or Negative. The distribution is summarized below:

Table 4. Sentiment Distribution

Sentiment	Count	Percentage
Positive	385	46.61%
Neutral	361	43.70%
Negative	80	9.69%

While the majority of feedback expressed positive or neutral sentiment, the negative sentiment category—though limited in volume—was densely concentrated within particular SRL subcomponents, indicating that frustration and dissatisfaction were not evenly distributed but instead linked to specific regulatory obstacles.

Sentiment Patterns by SRL Component

- Help-Seeking and Resource Use (Guidance Requests) exhibited uniformly negative sentiment (100%), with comments such as *“No support was available when I needed help”* or *“Navigation was unclear, and I felt lost.”* These responses highlight a critical breakdown in the provision of instructional or navigational scaffolds, which SRL literature identifies as essential for effective strategy execution (Hadwin, Järvelä, & Miller, 2018).
- Self-Reflection (Evaluation of Outcome) and Forethought (Goal Specification) subcategories recorded 60% and 50% negative sentiment respectively. Learners expressed uncertainty about how to interpret quiz results or ambiguity regarding intended learning outcomes (e.g., *“It’s not clear how this content helps me improve”*). Such gaps reflect impairments in strategic feedback integration—a cornerstone of SRL processes related to self-evaluation and planning (Zimmerman, 2000; Panadero, 2017).
- In contrast, Motivational Beliefs and Help-Seeking (Materials and Resources) subcategories were overwhelmingly positive, with learners noting the usefulness of examples, relevance of materials, and the confidence inspired by clear structure. Illustrative comments included: *“Loved*

the real-world case examples” and *“The references provided were very useful.”* These expressions signal an alignment between learner needs and design affordances, which supports the activation and reinforcement of SRL strategies.

These sentiment–behaviour alignments demonstrate that emotional responses are not peripheral but integral to the enactment of SRL. Specifically:

- Negative emotions (e.g., frustration, confusion) were systematically associated with absent or unclear scaffolding, particularly in help-seeking, planning, and reflection phases.
- Positive sentiments emerged when learners perceived that resources, guidance, and structural clarity supported their self-regulatory efforts, thereby boosting confidence and motivation.
- Neutral comments often accompanied general observations or procedural remarks and were less tightly linked to SRL enactment.

This confirms prior research suggesting that affective states act as catalysts or inhibitors of SRL, depending on whether the learning environment offers sufficient affordances for cognitive, behavioural, and emotional regulation (D'Mello et al., 2020; Efklides, 2011).

Moreover, the patterned alignment between sentiment and SRL component adds interpretive validity to the analytical framework. The emotional valence of learner feedback serves as an indirect indicator of where instructional design succeeds or fails to support regulation.

These findings reinforce the need for emotionally responsive instructional designs in self-directed online learning. Designers must anticipate emotional pain points, such as missing guidance, vague goals, or lack of feedback and proactively integrate supports such as:

- Explanation prompts and interactive feedback
- Clear goal framing and sequencing
- Easily accessible learner support channels

In summary, sentiment data offer a valuable lens for diagnosing SRL-support misalignments, revealing not just how learners regulate but how they feel about their capacity to do so within a given environment.

RQ4: How are SRL behaviours and sentiments associated with learner satisfaction scores?

To investigate how SRL expressions map onto learner satisfaction, mean satisfaction scores (on a 10-point scale) were computed for each sentiment category and each SRL component.

Table 5. shows a consistent monotonic gradient: comments coded as positive sentiment yielded the highest satisfaction ($M = 7.86$), followed by neutral ($M = 7.51$), with negative sentiment lowest ($M = 6.89$). This pattern suggests that emotional tone functions not merely as an after-the-fact reaction, but as an affective proxy for regulatory friction or fluency. Learners who experienced clarity, support, and perceived progress expressed more positive emotion and rated their experience higher; those encountering confusion, ambiguity, or technical friction expressed negative tone and rated lower.

Table 5. Mean Satisfaction by Sentiment

Sentiment	Mean Satisfaction Score
Positive	7.86
Neutral	7.51
Negative	6.89

Similarly, Table 6 shows that satisfaction varied systematically across SRL components. Motivational Beliefs ($M = 8.20$) and Help-Seeking and Resource Use ($M = 7.82$) were associated with the highest satisfaction scores. In contrast, Environmental Structuring ($M = 6.95$) and Forethought ($M = 6.96$) were linked to the lowest satisfaction. This pattern suggests that when the platform supports instrumental regulation (e.g., accessible explanations, meaningful resources), satisfaction increases but when learners must self-construct structure, pacing, and interpretive clarity, satisfaction deteriorates.

Table 6. Mean Satisfaction by SRL Component

SRL Component	Mean Satisfaction Score
Motivational Beliefs	8.20
Help-Seeking and Resource Use	7.82
Performance (Monitoring & Strategy Use)	7.79
Self-Reflection (Evaluation)	7.61
Time Management	7.30
Forethought (Planning)	6.96
Environmental Structuring	6.95

Pearson correlation analyses were conducted to examine the relationships between each self-regulated learning (SRL) component, sentiment polarity, and learners' self-reported satisfaction scores. The results revealed several statistically meaningful patterns that illuminate how cognitive, emotional, and behavioural processes relate to learners' perceived course experience:

Table 7. Correlation Matrix: SRL Components, Sentiment, and Satisfaction

	Sentiment	Satisfaction	Forethought (Planning)	Performance (Monitoring)	Self- Reflection (Evaluation)	Motivational Beliefs	Help- Seeking & Resource Use	Time Management	Environmental Structuring
Sentiment	1.00	0.15	0.04	0.05	0.08	0.11	0.09	0.06	-0.02
Satisfaction	0.15	1.00	-0.08	0.05	0.03	0.06	0.12	-0.07	-0.10
Forethought (Planning)	0.04	-0.08	1.00	0.28	0.25	0.18	0.21	0.47	0.10
Performance (Monitoring)	0.05	0.05	0.28	1.00	0.31	0.22	0.24	0.29	0.07
Self- Reflection (Evaluation)	0.08	0.03	0.25	0.31	1.00	0.29	0.27	0.22	0.09
Motivational Beliefs	0.11	0.06	0.18	0.22	0.29	1.00	0.36	0.23	0.04
Help-Seeking & Resource Use	0.09	0.12	0.21	0.24	0.27	0.36	1.00	0.26	0.05
Time Management	0.06	-0.07	0.47	0.29	0.22	0.23	0.26	1.00	0.12
Environmental Structuring	-0.02	-0.10	0.10	0.07	0.09	0.04	0.05	0.12	1.00

First, a moderate positive correlation was observed between Forethought (Planning) and Time Management ($r = 0.47$), indicating that learners who engage in goal-setting, sequencing, and

anticipatory planning are also more likely to allocate and manage time effectively. This supports prior SRL models (Zimmerman, 2000) that frame time management as a subcomponent of the forethought phase, reinforcing the notion that these behaviours co-occur as part of proactive learning regulation.

A second moderate correlation was found between Help-Seeking and Resource Use and Motivational Beliefs ($r = 0.36$), suggesting that learners who demonstrate higher levels of confidence, engagement, or interest are more likely to request assistance or utilize supplementary resources. Qualitative comments substantiated this association, with several learners linking resource availability and support with increased confidence and emotional investment in learning.

Additionally, Performance Monitoring was moderately correlated with Self-Reflection (Evaluation) ($r = 0.31$), aligning with cyclical SRL frameworks (Winne & Hadwin, 2008), where ongoing monitoring facilitates later evaluative judgments. Learners who tracked their performance, for instance, also reflected on areas for improvement, underscoring the sequential and recursive nature of SRL processes.

In contrast, Environmental Structuring demonstrated low or near-zero correlations with most other SRL components, suggesting that learners' efforts to manage their learning environment, such as overcoming platform usability issues or navigating course logistics, operated relatively independently from deeper cognitive or metacognitive strategies. This decoupling highlights a potentially under-recognized barrier in digital learning design: that logistical or technical friction may impair regulatory engagement without being cognitively integrated.

Correlation with satisfaction scores revealed additional insights. A weak but positive correlation was found between sentiment polarity and satisfaction ($r = 0.15$), confirming that learners expressing more positive emotional tone tended to report higher satisfaction. Similarly, Help-Seeking and Resource Use was positively associated with satisfaction ($r = 0.12$), reinforcing the value of accessible resources and instructional support in shaping learner perceptions.

Conversely, Forethought ($r = -0.08$), Time Management ($r = -0.07$), and Environmental Structuring ($r = -0.10$) were negatively correlated with satisfaction, albeit weakly. These inverse relationships suggest that learners who experienced challenges in planning, pacing, or course navigation may have been more likely to express dissatisfaction, potentially due to a perceived mismatch between course design and their regulatory needs.

Together, these findings highlight that learner satisfaction in self-directed online environments is influenced not only by content quality but also by the emotional tone of feedback and the presence (or absence) of regulatory supports. While SRL behaviours tend to manifest in compound and interrelated forms, particularly around planning, monitoring, and motivation, certain components such as environmental structuring appear to operate in isolation, signalling opportunities for targeted design interventions. The data reinforce the importance of embedding motivational scaffolds and resource support mechanisms to enhance both learner satisfaction and regulatory success.

Findings from the correlation analysis revealed that planning, monitoring, and self-reflection, core phases in Zimmerman's (2000) cyclical model of self-regulated learning (SRL), are not discrete or sequential in practice, but mutually reinforcing processes that often co-occur as compound regulatory strategies. This was evident in the moderate positive correlation observed between Performance Monitoring and Self-Reflection ($r = 0.31$), as well as the close association between Forethought (Planning) and Time Management ($r = 0.47$), a planning-related behavioural regulation.

These relationships suggest that learners who set clear goals and sequence their tasks (Forethought) are more likely to manage their time deliberately, track their progress while engaging with learning materials (Monitoring), and evaluate their performance against internal goals or external criteria (Self-Reflection). Importantly, the expression of one behaviour often appears to activate or amplify the others.

For instance, goal-oriented planning may create the mental structure necessary for ongoing monitoring, which in turn provides the data and insight that makes reflective evaluation more meaningful.

This synergistic interplay aligns with theoretical propositions by Winne and Hadwin (2008), who posit that regulation is recursive and information-rich; each phase informs the next, with feedback loops supporting adaptive refinement of learning strategies. In this study, such feedback loops were visible in learner comments that combined references to pacing, clarity of expectations, and quiz performance—indicating that learners were not only thinking ahead but continuously adjusting and interpreting their progress in light of outcomes.

Moreover, the co-activation of planning, monitoring, and reflection also suggests that interventions designed to support one dimension (e.g., planning tools or checklists) may have positive spillover effects on others (e.g., improved monitoring or reflective accuracy). Thus, the design of online learning environments should avoid treating SRL components as isolated features and instead promote integrated scaffolds that simultaneously engage these interdependent phases—for example, by combining goal-setting prompts with progress dashboards and self-evaluation rubrics.

By recognizing these components as compound and co-regulated, rather than linear or modular, educators and instructional designers can better anticipate learner needs and tailor supports that reflect the dynamic, situated nature of SRL.

RQ 5: What distinct learner typologies emerge from SRL-themed feedback, and how do these archetypes reflect differing self-regulation challenges and support needs in self-directed online learning environments?

To address RQ5, a qualitative synthesis was conducted on SRL-coded and sentiment-tagged feedback to develop a learner typology framework. This process aimed to identify recurring patterns of regulatory orientation, support needs, and affective tone across learners. The goal was to move beyond isolated SRL components and sentiment scores to construct integrated learner archetypes that reflect how SRL behaviours coalesce within real-world online learning contexts.

The process involved three stages:

1. **Thematic Coding:** Comments were first coded into the seven core SRL components (e.g., forethought, help-seeking) and tagged for sentiment polarity (positive, neutral, negative). This coding served as the theoretical foundation.
2. **Cluster Affinity Mapping:** Using the output from unsupervised KMeans clustering, thematically similar comments were grouped and cross-referenced with SRL subcategories. Each cluster was reviewed for dominant concerns, expressed regulatory behaviours, and emotional tone.
3. **Pattern Consolidation:** Through iterative comparison and synthesis, six recurring learner profiles emerged. These profiles were identified based on the dominant regulatory challenges and support-oriented feedback themes that defined each cluster. Each archetype represents a cluster of behaviours and dispositions rather than a single trait or sentiment.

The six resulting archetypes—Planner, Clarifier, Frustrated Navigator, Strategic Seeker, Reflective Improver, and Isolated Achiever—are summarized below.

1. The Planner (20.77%): Planners are characterized by their strong emphasis on preparation, sequencing, and overall course structure. They often express a need for clarity in objectives, logical flow of content, and manageable pacing. A representative comment includes, *"Would be better if the course was structured in a more logical flow,"* which indicates a desire for a more roadmap-driven learning experience. These learners tend to thrive when provided with clear goals and structured milestones.

2. The Clarifier (18.95%): Clarifiers focus on conceptual clarity, assessment transparency, and explanation access. Their feedback often reflects frustration with vague instructions or ambiguous case studies. One learner noted, *"Need to refine case studies and questions to be more precise,"* signalling the importance of guided reasoning and feedback loops. Clarifiers benefit most from real-time explanation features and embedded scaffolding that support understanding.

3. The Frustrated Navigator (16.73%): This group frequently encounters issues related to the platform's usability, navigation, and content accessibility. Feedback such as, *"The course always shows incomplete, even though I completed all tasks,"* illustrates the technical and interface barriers that disrupt their learning flow. Addressing this group's needs involves enhancing UI/UX design and providing just-in-time technical guidance.

4. The Strategic Seeker (15.73%): Strategic Seekers are proactive in their learning process, often requesting tools that help them externalize cognition, such as note-taking, downloadable references, or cross-referencing features. A typical comment, *"Add a space where we can take notes,"* reflects their desire to organize and personalize learning materials. These learners respond well to environments that provide auxiliary learning tools and resource dashboards.

5. The Reflective Improver (14.52%): Reflective Improvers demonstrate a high level of metacognitive awareness, using the course as a tool to understand and improve their performance. One learner wrote, *"This training will help in my entire career,"* suggesting that they integrate learning experiences into long-term personal or professional development. These learners benefit from reflection prompts and portfolio-based assessments that track growth over time.

6. The Isolated Achiever (13.31%): Isolated Achievers are typically independent learners who progress autonomously but may experience a lack of connection or community. A learner's comment, *"Felt very alone at times. Maybe a discussion board could help,"* underscores the emotional and social void they sometimes encounter. While they value flexibility, their experience improves with options for low-friction peer engagement and mentorship.

This typology reveals the diversity of learner profiles in self-directed online learning and highlights the importance of designing with learner variability in mind. By understanding these archetypes, educators and platform developers can better tailor the learning experience—offering targeted scaffolds, adaptive features, and emotional or social support that align with distinct SRL needs.

RQ6: To what extent do learner typologies correlate with specific self-regulated learning (SRL) components, and how do these associations validate the behavioural underpinnings of each archetype?

To validate the behavioural foundations of the learner archetypes identified in RQ5, we examined statistical correlations between each typology and the seven core SRL components. This step extends the findings of RQ2, which established interrelationships among SRL strategies, by testing whether the six emergent learner profiles are quantitatively grounded in distinct patterns of regulatory behaviour.

To enhance interpretability, both tables and the figure are presented together. Tables 8. and 9. provide precise numerical values and highlight the strongest associations for each learner typology, ensuring statistical transparency. Figure 1., in contrast, offers a visual representation that allows readers to quickly identify patterns of strength and weakness across typologies and SRL components. The combination of tables and figure thus balances analytical precision with intuitive visualization, making the results more accessible for both researchers and practitioners.

Table 8. Correlation Matrix between Typologies and SRL Components

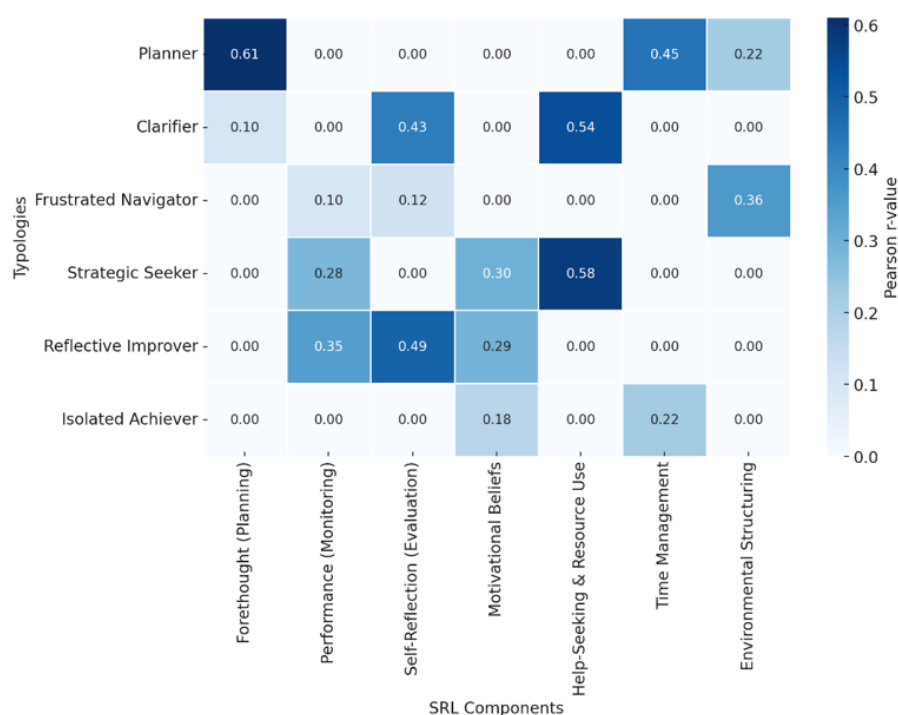
Typology	Forethought (Planning)	Performance (Monitoring)	Self-Reflection (Evaluation)	Motivational Beliefs	Help-Seeking & Resource Use	Time Management	Environmental Structuring
Planner	0.61	—	—	—	—	0.45	0.22
Clarifier	~0.10 (low)	—	0.43	—	0.54	—	—
Frustrated Navigator	—	~0.10 (weak)	~0.12 (weak)	—	—	—	0.36
Strategic Seeker	—	0.28	—	0.30	0.58	—	—
Reflective Improver	—	0.35	0.49	0.29	—	—	—
Isolated Achiever	—	—	—	0.18	—	0.22	—

Note: Values represent Pearson's correlation coefficients (r). "—" indicates negligible or non-reported correlation.

Table 9. Strongest Typology–SRL Associations

Typology	Strongest SRL Associations (r-values)
Planner	Forethought (0.61), Time Management (0.45), Environmental Structuring (0.22)
Clarifier	Help-Seeking & Resource Use (0.54), Self-Reflection (0.43)
Frustrated Navigator	Environmental Structuring (0.36), weak links to Performance/Self-Reflection
Strategic Seeker	Help-Seeking & Resource Use (0.58), Motivational Beliefs (0.30), Performance (0.28)
Reflective Improver	Self-Reflection (0.49), Performance (0.35), Motivational Beliefs (0.29)
Isolated Achiever	Time Management (0.22), Motivational Beliefs (0.18)

Figure 1. Heatmap of correlations between learner typologies and SRL components



Note: Darker shading indicates stronger positive Pearson correlations (r-values) between learner typologies and SRL components.

Quantitative Validation of Typologies

Pearson correlation coefficients were computed between each learner typology and SRL component, based on frequency-normalized occurrence of coded behaviours across the dataset ($n = 826$). As shown in Table 8., the resulting correlation matrix provides strong empirical support for the construct validity of the typologies, confirming that they reflect distinct SRL orientations rather than merely descriptive or anecdotal categories.

Key results include:

- The Planner exhibited a strong positive correlation with Forethought (Planning) ($r = 0.61$) and a moderate association with Time Management ($r = 0.45$), confirming their proactive engagement in goal setting, sequencing, and pacing. A smaller correlation with Environmental Structuring ($r = 0.22$) also suggests a peripheral concern with interface layout and learning space organization.
- The Clarifier was most strongly aligned with Help-Seeking and Resource Use ($r = 0.54$) and Self-Reflection (Evaluation) ($r = 0.43$), reinforcing their reliance on feedback loops and clarification mechanisms. A low correlation with Forethought ($r \approx 0.10$) indicates that while Clarifiers value understanding, they may not engage deeply in pre-task planning.
- The Frustrated Navigator showed a moderate correlation with Environmental Structuring ($r = 0.36$), and only weak links to Performance Monitoring and Reflection, validating their typology as primarily shaped by platform and interface challenges rather than cognitive strategy deficits.
- The Strategic Seeker demonstrated strong alignment with Help-Seeking and Resource Use ($r = 0.58$), and moderate correlations with Motivational Beliefs ($r = 0.30$) and Performance Monitoring ($r = 0.28$). This combination reflects their proactive, resource-aware approach to learning, driven by confidence and self-direction.
- The Reflective Improver was most closely associated with Self-Reflection ($r = 0.49$), followed by Performance Monitoring ($r = 0.35$) and Motivational Beliefs ($r = 0.29$). This supports the profile of a metacognitively engaged learner who integrates feedback and performance awareness into long-term development.
- The Isolated Achiever, while less strongly associated with any particular SRL component, demonstrated small but positive correlations with Time Management ($r = 0.22$) and Motivational Beliefs ($r = 0.18$), indicating a preference for autonomous regulation coupled with affective disengagement from the learning community.

These findings confirm that the learner archetypes are not merely narrative labels but are empirically grounded in observable SRL behaviours. The typologies emerged from open-ended qualitative feedback, were confirmed through unsupervised clustering, and are now validated through correlational analysis—demonstrating a convergent triangulation of methods.

Each archetype maps onto a distinct regulatory profile, suggesting that SRL is best understood not as a universal set of traits, but as a context-sensitive pattern of interrelated strategies. For example:

- The Planner represents a front-loaded regulatory mode, focused on organizing and sequencing learning in advance.
- The Reflective Improver exemplifies a feedback-driven, developmental orientation that integrates performance monitoring and affective engagement.
- The Strategic Seeker highlights regulatory fluidity, combining motivation, help-seeking, and strategy adaptation in response to learning demands.

These differentiated profiles carry clear instructional implications: adaptive systems and course designers should scaffold SRL in diverse ways that reflect the variability in learner regulation patterns.

The results also provide a solid empirical foundation for developing adaptive learner personas, informing both automated personalization and human-centred support strategies in self-directed online learning.

RQ7: What latent themes can be uncovered through unsupervised machine learning of learner feedback, and how do these clusters reinforce or challenge the SRL components and learner typologies identified in previous analyses?

To triangulate the findings established in RQ1 through RQ6, an unsupervised machine learning analysis was conducted on the full corpus of learner feedback. This phase employed a Term Frequency–Inverse Document Frequency (TF-IDF) vectorization technique in combination with KMeans clustering to uncover latent semantic patterns. The aim was to validate, challenge, or extend the manually coded SRL components and emergent learner typologies by deriving patterns directly from the lexical structure of the data—thereby minimizing interpretive bias and enhancing the robustness of the findings.

Latent Clusters and Thematic Mapping

The optimal number of clusters was determined using silhouette analysis, and a five-cluster solution was selected for its thematic coherence and statistical stability. Each cluster was defined by its most salient keywords (based on TF-IDF weights) and interpreted through both an SRL lens and the archetypal learner typologies.

As summarized in Table 10, the five emergent themes—Content Clarity, Technical Usability, Resource Access, Structure and Flow, and Engagement—mapped consistently onto previously defined SRL components and typologies, offering independent confirmation of the earlier frameworks.

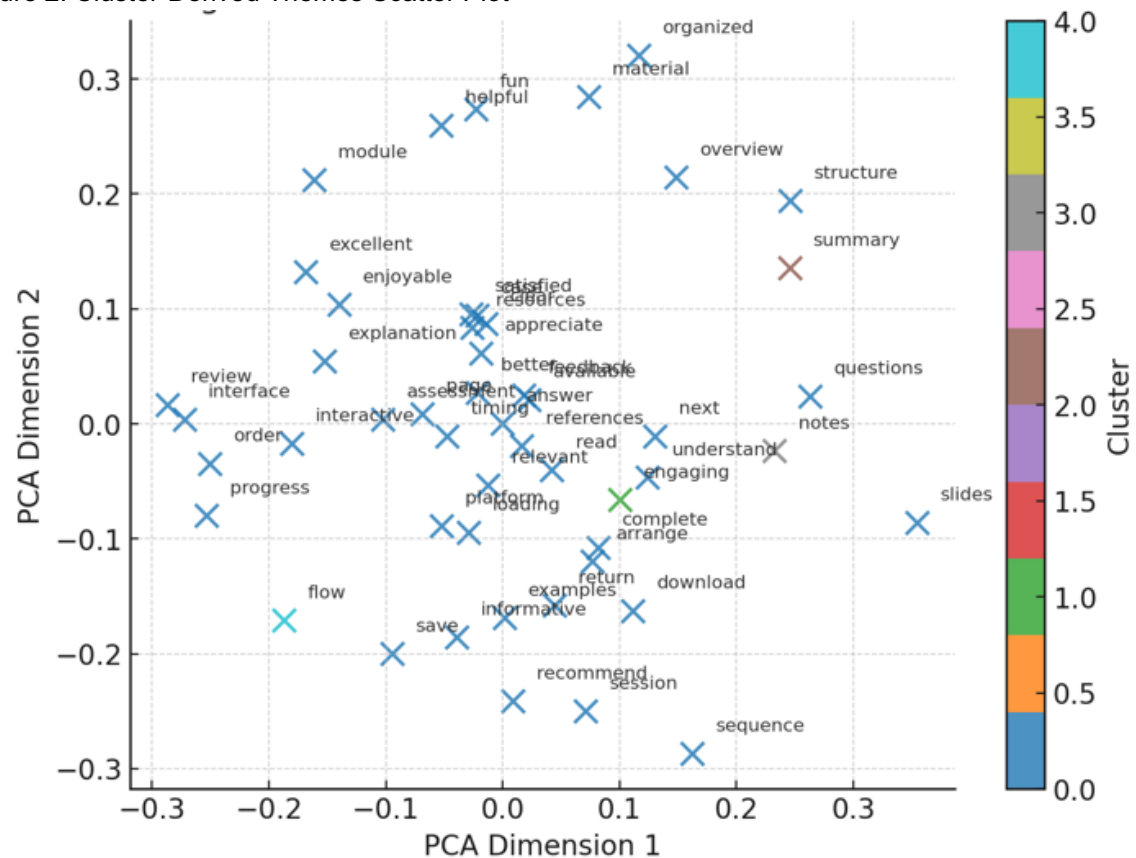
Table 10. Summary of Cluster-Derived Learner Themes

Cluster	Top Keywords	Interpretation / Linked SRL Components & Typologies
Cluster 0: Content Clarity & Practical Relevance	explanation, case, examples, questions, feedback, assessment, understand, answer, clear, relevant	Focus on clarity, alignment of assessments, practical examples. Linked to Help-Seeking, Self-Reflection, Motivational Beliefs; typologies include Clarifier, Reflective Improver
Cluster 1: Technical Usability & Navigation	page, complete, platform, session, save, progress, loading, next, return, interface	Concerns with usability, progress tracking, session saving. Linked to Environmental Structuring, Time Management; typology: Frustrated Navigator
Cluster 2: Resource Access & Depth of Content	material, notes, resources, references, available, review, download, read, summary, slides	Requests for more resources, notes, references. Linked to Help-Seeking, Resource Use, Performance; typologies: Strategic Seeker, Planner
Cluster 3: Structure, Flow, and Timing	module, flow, structure, order, timing, arrange, sequence, better, organized, overview	Need for structured flow and pacing. Linked to Forethought, Time Management; typology: Planner
Cluster 4: Engagement and Emotional Resonance	engaging, helpful, excellent, informative, interactive, enjoyable, satisfied, fun, appreciate, recommend	Positive sentiment, course value, enjoyment. Linked to Motivational Beliefs, Self-Reflection; typologies: Reflective Improver, Strategic Seeker

Visualization of Semantic Clusters

To complement the lexical interpretation, a scatter plot was generated using dimensionality reduction (e.g., t-SNE or PCA) applied to the TF-IDF feature space. Each point in Figure 2 represents a high-weighted keyword, coloured by cluster membership. The spatial separation of terms provides a visual affirmation of cluster coherence, with clear groupings (e.g., “fun,” “helpful,” and “engaging” for Engagement; “platform,” “progress,” and “page” for Usability) highlighting distinct latent themes.

Figure 2. Cluster-Derived Themes Scatter Plot



Convergence of Machine Learning and Manual Coding

The alignment between the machine-derived clusters and manually coded SRL components strengthens the interpretive validity of both approaches. Each cluster was not only statistically robust but conceptually coherent, reflecting core SRL dimensions such as planning (Cluster 3), help-seeking (Clusters 0 and 2), or emotional motivation (Cluster 4). Furthermore, the clusters mirrored the learner archetypes established in RQ5 and validated in RQ6, for example:

- Cluster 0's focus on explanation and feedback resonates with the Clarifier and Reflective Improver archetypes.
- Cluster 1's interface challenges reaffirm the presence of Frustrated Navigators in the sample.
- Cluster 2's resource needs directly support the behavioural profile of Strategic Seekers.

This triangulation indicates that SRL behaviours, learner sentiment, and engagement styles are intricately interwoven, and can be independently verified through both human coding and computational approaches.

Interpretive Contribution and Design Implications

The convergence of unsupervised clustering with SRL theory and learner typologies underscores a central conclusion: Learner feedback contains rich behavioural signals that reflect underlying self-regulatory processes. Unsupervised methods not only confirm existing frameworks but also help to refine them, revealing subtle distinctions (e.g., between Strategic Seekers and Planners) and surfacing under-addressed needs (e.g., emotional isolation or interface friction).

By integrating machine learning into SRL research pipelines, educators and platform designers can move beyond anecdotal analysis to scalable, pattern-based insights, allowing for adaptive system design that better supports diverse learner profiles in self-directed environments.

RQ8: How can SRL-based learner typologies and cluster-derived feedback themes inform practical instructional design strategies in self-directed online learning environments?

This research question operationalizes the analytical insights derived from RQ1 through RQ7 by translating them into actionable instructional design strategies for self-directed online learning. Rather than presenting abstract principles, the recommendations offered here are grounded in a triangulated empirical foundation: (1) the thematic SRL coding of 826 learner comments, (2) the development of six inductively derived learner typologies, and (3) the identification of five latent feedback clusters using unsupervised machine learning. Together, these sources illuminate both the individual regulation patterns of learners and the systemic design frictions that shape their experiences—allowing for the construction of differentiated, SRL-aligned design interventions.

From the SRL coding, learners were found to most frequently express needs tied to help-seeking, forethought, and self-reflection. Comments requesting answer explanations, clearer assessment feedback, and better content relevance highlight concrete behavioural markers of self-regulation. The six learner typologies developed in RQ5, namely, the Planner, Clarifier, Frustrated Navigator, Strategic Seeker, Reflective Improver, and Isolated Achiever, further revealed that learners cluster around specific regulatory orientations, each associated with distinct frustrations and support needs. Complementing this, the five data-driven feedback clusters (Content Clarity, Technical Usability, Resource Access, Structure and Flow, Engagement) uncovered through TF-IDF and KMeans clustering mirror both SRL patterns and typology themes. This convergence confirms the validity of the learner archetypes and reveals recurrent friction points in platform design and content delivery.

These three data layers together form the foundation of a design strategy model that aligns pedagogical supports with observable patterns of self-regulation.

Planners, who emphasize anticipation and course structuring, benefit from design features such as visual learning paths, milestone markers, and clearly indicated progress checkpoints. These affordances help them sequence their learning, manage time proactively, and maintain a sense of direction. Implementing these features involves modularizing course content into digestible units with visible pacing expectations and clear completion indicators.

Clarifiers, by contrast, seek transparency in assessments and clarity in instruction. They require design elements such as expandable feedback explanations, inline pop-up clarifications, and on-demand access to worked examples or instructor commentary. These learners respond well to environments where ambiguity is minimized and where every task comes with cognitive scaffolds that clarify reasoning and correctness.

Frustrated Navigators often encounter friction due to poor user interface design or technical inconsistency. Their feedback suggests the need for smoother navigation systems, autosave functions, breadcrumb trails, and “resume where you left off” features. Improving platform usability is essential for this group, as technical frustrations can significantly inhibit both motivation and the ability to sustain regulatory focus.

Strategic Seekers are proactive learners who thrive in environments that support self-direction and organization. They benefit from downloadable resources, centralized hubs for reference materials, and integrated tools for note-taking and bookmarking. For these learners, externalizing cognition through resource interaction and personal annotation is key to managing learning load and retaining information.

Reflective Improvers are metacognitively inclined learners who value self-evaluation and longitudinal growth. Instructional designs that support reflection, such as end-of-module journaling prompts, portfolio submissions, and self-assessment checklists, allow them to track and make meaning of their learning trajectories. These features also deepen their emotional investment in the course, as they see their progress tied to personal and professional development goals.

Isolated Achievers, while generally self-sufficient, often experience a lack of social presence in the learning environment. They respond well to optional, low-pressure social structures—such as opt-in discussion boards, peer mentoring features, or light community nudges—that preserve autonomy while reducing feelings of isolation.

When viewed from a system-level perspective, the unsupervised clusters reinforce these design needs by identifying broader structural and emotional drivers of learner experience. For instance, the “Content Clarity” cluster points to the need for well-framed case studies and explicit rubrics, which address both Clarifiers’ and Reflective Improvers’ desire for conceptual precision. The “Technical Usability” cluster validates the difficulties experienced by Frustrated Navigators and emphasizes the importance of robust, intuitive LMS interfaces. Meanwhile, the “Resource Access” cluster highlights the recurring demand for accessible, flexible content formats that support Strategic Seekers. The “Structure and Flow” cluster links back to the Planner typology and underscores the role of pacing and sequencing in sustaining motivation. Finally, the “Engagement” cluster, characterized by terms like “interactive,” “enjoyable,” and “satisfied”, suggests that emotional resonance, gamification, and personalized encouragement can boost motivation, particularly for Reflective Improvers and Isolated Achievers.

Taken together, these findings form the basis of a **differentiated instructional model** rooted in learner behaviour, emotional response, and regulatory patterning. Rather than attempting to meet all learner preferences simultaneously, a strategy that is neither scalable nor pedagogically sound, this model promotes the strategic alignment of design features with high-frequency needs and high-impact barriers. Such alignment not only supports SRL processes but also fosters learner satisfaction, emotional resilience, and course completion.

In practical terms, this means designing learning environments where cognitive, behavioural, and emotional supports are scaffolded in ways that are responsive to learner typologies and informed by empirical data. For course developers and instructional designers, the model serves as a decision-making tool—offering a research-validated basis for prioritizing features that are most likely to yield regulatory, affective, and motivational gains.

Ultimately, the answer to RQ8 confirms that SRL-informed learner typologies and latent feedback clusters can be effectively mobilized to inform instructional design. The convergence of theory, user voice, and machine learning reinforces the reliability of the proposed model and its relevance to the evolving demands of online professional education.

Integrating Typologies and Cluster Themes in Instructional Design: A Unified Framework

The integration of SRL-based learner typologies and cluster-derived feedback themes offers a robust, empirically grounded framework for the design of self-directed online learning environments. Rather than treating these two analytic outputs as discrete entities, this study demonstrates how they can be strategically synthesized to support differentiated instructional design, enhance learner engagement, and address structural barriers to effective self-regulation.

Learner typologies, as developed in this study, capture intra-individual variation in regulatory preferences, challenges, and strengths. These archetypes, such as the Planner, Clarifier, Frustrated Navigator, Strategic Seeker, Reflective Improver, and Isolated Achiever, provide insight into the dominant SRL strategies or breakdowns characterizing different subsets of learners. Each typology reflects behavioural dispositions aligned with specific SRL components, and, as shown in prior analyses, these were quantitatively validated through significant correlations with planning, help-seeking, reflection, motivational beliefs, and other key SRL dimensions.

Cluster-derived themes, on the other hand, emerge from unsupervised machine learning techniques and represent systemic patterns across the entire learner population. The five identified clusters—Content Clarity and Practical Relevance, Technical Usability and Navigation, Resource Access and Content Depth, Structure and Flow, and Engagement and Emotional Resonance—highlight platform-level design issues and recurring content-related concerns. These clusters reinforce the SRL categories at a macro level, surfacing patterns of breakdown or support in learners' attempts to regulate cognition, emotion, and behaviour.

By integrating these two levels of analysis, instructional designers can move beyond generic or monolithic course structures toward more responsive and scalable learning environments. The combined framework enables two key design practices:

Each learner typology indicates a set of support needs that can be directly addressed by interventions aligned with cluster-identified themes. For instance:

- The *Planner* typology's need for logical progression and scheduling aligns closely with the *Structure and Flow* cluster, suggesting that timeline visualizations, modular progression, and pacing controls should be prioritized in course design.
- The *Clarifier*'s emphasis on answer transparency and feedback mirrors the *Content Clarity* cluster, indicating the value of embedded explanations, annotated rubrics, and formative feedback mechanisms.
- The *Frustrated Navigator* typology shares structural pain points with the *Technical Usability* cluster, reinforcing the urgency of improving interface design and learner orientation.

This alignment ensures that design changes are not only data-driven but also theoretically justified, grounded in validated SRL constructs.

While feedback clusters reflect concerns broadly shared across the learner cohort, typologies highlight how different learners experience the same design elements in different ways. For example, both Planners and Reflective Improvers may engage with progress dashboards, but for distinct regulatory reasons—Planners to manage pacing, Reflective Improvers to evaluate learning outcomes. Recognizing these differences allows instructional designers to layer support features with multiple entry points and optional scaffolds, promoting both structure and autonomy.

Moreover, emotional engagement and motivational reinforcement, central to the *Isolated Achiever* and *Engagement* cluster, can be addressed through opt-in discussion boards, milestone celebrations, and

reflective prompts. These design features cater to the motivational beliefs and social regulation dimensions of SRL, as identified across both analytic layers.

This unified framework provides several implications for practice:

1. Design with Regulatory Diversity in Mind: Avoid assuming a homogeneous learner profile. Instead, accommodate varying SRL strategies through optional tools, flexible pacing mechanisms, and scaffolded pathways that support both novice and experienced self-regulators.
2. Prioritize Convergence Zones: Focus limited design resources on high-frequency pain points where typology needs and cluster themes converge (e.g., explanation access, course structure, technical navigation), as these are likely to yield the most significant improvements in learner experience and satisfaction.
3. Scaffold SRL Processes, Not Just Content Mastery: Embed supports that explicitly target planning, monitoring, help-seeking, and reflection, not only to enhance learning outcomes, but also to build learners' long-term self-regulatory capacity.
4. Adopt Iterative, Data-Informed Redesign: Treat learner typologies and clusters as diagnostic tools that evolve over time. Use ongoing feedback and performance analytics to refine supports and adjust design strategies in response to emerging learner patterns.

By synthesizing SRL-based learner typologies with machine-learned feedback themes, this study advances a cohesive instructional design framework that is both theoretically grounded and empirically actionable. The resulting model equips designers to address individual variability and systemic friction simultaneously, promoting more inclusive, engaging, and self-regulation-sensitive online learning environments. This dual-layered approach not only enhances the pedagogical alignment of course structures with learner needs but also provides a replicable model for future SRL-driven course design in digital education contexts.

Discussion

This study advances the understanding of self-regulated learning (SRL) in self-directed online environments by offering a nuanced, data-rich analysis of how learners express, experience, and navigate regulatory behaviours. While the significance of SRL in online learning contexts is well established (Broadbent & Poon, 2015; Zimmerman, 2000), the present study contributes novel insights by triangulating sentiment, typology, behavioural coding, and machine learning outputs to generate a multidimensional portrait of regulatory practice.

A key finding is the *heterogeneity* of SRL behaviours across learners. Despite the established notion that SRL skills are learnable and developmental (Zimmerman & Schunk, 2011; Panadero, 2017), this analysis revealed persistent gaps in foundational strategies such as planning, help-seeking, and outcome evaluation. Help-Seeking and Resource Use emerged as the most prevalent component across 826 comments, often expressed through learner frustrations regarding inaccessible materials, limited feedback, or inadequate support. These patterns were not only evident in qualitative themes but also statistically substantiated: Help-Seeking was positively correlated with both learner satisfaction ($r = 0.12$) and sentiment polarity ($r = 0.15$), reinforcing its instrumental role in learner experience.

Particularly salient were the emotional valences attached to specific SRL breakdowns. Negative sentiments were disproportionately associated with subcategories such as Goal Specification, Evaluation of Outcome, and Guidance Requests—core elements of the Forethought and Self-Reflection phases in SRL models (Winne & Hadwin, 2008). The fact that Guidance Requests were uniformly negative (100%) highlights the emotional and cognitive toll experienced by learners when instructional environments lack sufficient scaffolds. This affirms Panadero's (2017) critique that SRL research has

often under-theorized the affective dimension of regulation, and positions emotional tone as a critical barometer for identifying regulatory friction.

The identification of six learner typologies, Planner, Clarifier, Frustrated Navigator, Strategic Seeker, Reflective Improver, and Isolated Achiever, provides further granularity to these behavioural trends. These archetypes are not merely descriptive; they were empirically validated through significant correlations with SRL components. For example, Planners showed strong associations with Forethought ($r = 0.61$) and Time Management ($r = 0.45$), yet also reported lower satisfaction levels when course structures failed to align with their preparatory efforts. This supports Khat's (2017) argument that self-direction is a necessary but insufficient condition for SRL; learners may exhibit high motivation but still require environmental affordances to successfully regulate.

Similarly, Reflective Improvers were closely linked with Self-Reflection ($r = 0.49$), and Clarifiers with Help-Seeking ($r = 0.54$), suggesting that different learners prioritize distinct regulatory dimensions. The Strategic Seekers' preference for auxiliary tools (e.g., notes, resources, references) corresponds with Performance Monitoring and Motivational Beliefs, indicating a proactive stance that benefits from design flexibility and content depth. The Isolated Achiever, by contrast, demonstrated relatively weak SRL correlations, reflecting their independence but also emotional vulnerability in the absence of community, echoing findings on the socio-emotional aspects of online learning (Yu, 2023; D'Mello et al., 2020).

The unsupervised clustering analysis (TF-IDF + KMeans) further corroborated these findings by surfacing five latent learner feedback themes—Content Clarity, Technical Usability, Resource Access, Structure and Flow, and Engagement. These clusters aligned strongly with both SRL categories and typologies. For example, the Technical Usability cluster reinforced the Frustrated Navigator typology, while Content Clarity mirrored the concerns of Clarifiers and Reflective Improvers. Notably, the Engagement cluster overlapped with high-satisfaction learners and positive sentiment, underscoring the motivational and emotional dimensions of regulatory success.

Moreover, the interplay between SRL, emotion, and satisfaction was reinforced through statistical patterns: motivational beliefs were associated with the highest satisfaction scores ($M = 8.20$), and subcomponents with negative sentiment corresponded to lower satisfaction (e.g., Goal Specification, $M = 6.89$). These findings affirm that SRL is not purely cognitive; it is intrinsically affective, dynamically influenced by learner perception, platform responsiveness, and emotional engagement.

Theoretically, this study supports the cyclical model of SRL (Zimmerman, 2000) and computational model (Winne & Hadwin, 2008), highlighting the recursive nature of planning, monitoring, and reflection. Compound regulatory strategies, such as those exhibited by Strategic Seekers and Reflective Improvers, demonstrated how multiple SRL components co-activate to support learning. Yet, weak correlations involving Environmental Structuring suggest that contextual navigation (e.g., platform usability) may operate semi-independently from internal regulatory cycles. This echoes recent findings by Lobos et al. (2024), which emphasize the role of interface design and digital friction in modulating regulatory behaviour.

Importantly, these insights translate into actionable implications for instructional design. Rather than designing uniformly for an idealized SRL learner, course environments must support regulatory diversity and adaptivity. Typology-informed scaffolds, such as visual progress indicators for Planners or reflection journals for Improvers, can be coupled with system-level design responses to cluster-derived themes (e.g., reducing technical friction, clarifying assessment rubrics). These design interventions should not only address cognitive demands but also foster emotional and motivational resonance, creating a holistic environment conducive to sustainable SRL.

In summary, this discussion converges on three key insights. First, SRL behaviours are diverse, patterned, and context-sensitive—not uniformly distributed across learners. Second, emotional tone and satisfaction are closely intertwined with regulatory functioning, revealing both strengths and

vulnerabilities in design. Third, effective support for SRL requires multi-level interventions: individualised scaffolds anchored in learner typologies, and structural reforms aligned with system-wide cluster themes. Together, these findings advance both SRL theory and practice, providing a foundation for designing emotionally intelligent, regulation-sensitive online learning environments that are responsive to the evolving needs of diverse learners.

Conclusion and Suggestions

This study has offered a multi-layered analysis of how self-regulated learning (SRL) manifests within learner feedback in self-directed online environments, providing both theoretical extension and empirical critique of how SRL operates in practice. Through a convergence of pedagogical coding, sentiment analysis, correlation testing, and unsupervised clustering, the findings challenge the assumption that SRL is uniformly enacted or inherently effective across digital learning contexts.

The evidence demonstrates that SRL is not simply a set of internal learner attributes but is significantly shaped, and at times constrained, by system design and contextual affordances. Regulatory components such as help-seeking, resource use, and motivational beliefs were consistently associated with higher satisfaction and positive sentiment, affirming their role as productive and well-supported behaviours within the platform. However, the finding that forethought and time management—typically regarded as essential SRL strategies, were linked to dissatisfaction undercuts the traditional assumption of their universal benefit. These components became sources of frustration when course pacing, navigation, or sequencing failed to align with learners' planning efforts, thereby exposing a critical mismatch between regulatory intent and systemic responsiveness.

This misalignment also surfaced in the identified learner typologies. Planners, despite their strategic orientation, reported lower satisfaction when their regulatory efforts were undermined by unclear objectives or inflexible course flows. Clarifiers, who depended on interpretive scaffolds, struggled in environments lacking feedback loops or conceptual transparency. These profiles, validated through correlations with specific SRL components, confirm that regulatory tendencies alone do not guarantee success; rather, their effectiveness is contingent upon the learning environment's ability to meet corresponding needs. Notably, the Frustrated Navigator typology illustrated how platform-level usability failures can obstruct even the most motivated learners, suggesting that environmental structuring remains an under-theorised but crucial dimension of SRL.

The cluster-derived themes provided an independent yet convergent lens on these dynamics, reinforcing that many learner struggles stem not from lack of effort but from systemic design weaknesses. Clusters related to content clarity, technical usability, and pacing exposed recurrent structural issues that directly interfere with the enactment of SRL. For example, when learners encounter interface glitches or ambiguous instructions, the opportunity to self-monitor or reflect is not just diminished, it becomes irrelevant. These findings challenge prevailing SRL models that centre regulation within the learner and instead argue for models that recognise regulation as distributed, mediated, and co-constructed with the learning environment.

At a theoretical level, the study supports the cyclical and computational frameworks of SRL but calls for sharper attention to their limitations. The weak correlations of environmental structuring with other SRL components, despite frequent learner complaints, suggest that current models insufficiently account for the material and technical conditions under which regulation occurs. Moreover, the consistent alignment between positive affect, satisfaction, and SRL behaviours supports the argument that emotion is not merely a backdrop to regulation but a constitutive element. This lends empirical weight to D'Mello et al.'s (2020) and Panadero's (2017) calls for an affect-integrated theory of SRL and affirms the need to reconceptualise regulation not as a solitary cognitive act but as an emotionally and contextually embedded process.

In addressing the constraints of this study, several targeted suggestions emerge. To overcome the retrospective nature of post-course feedback, future research should incorporate real-time, multimodal data collection, including interaction logs, sentiment trajectories, and biometric signals, to trace regulation as it unfolds. This would address the temporal and self-selection biases inherent in self-report data and offer richer behavioural insights. Similarly, expanding beyond a single professional learner demographic is necessary to test the generalisability of typologies and design implications. Cross-context studies in undergraduate, vocational, and informal learning settings would clarify whether the observed patterns are universal or domain-specific. The limitations of lexicon-based sentiment tools can also be addressed by deploying fine-tuned large language models trained on educational discourse to capture subtle emotional variations and cultural nuances.

Methodologically, while triangulation improved interpretive reliability, future research should explore greater automation and inter-rater calibration in coding and cluster validation, to reduce the influence of subjective bias. Finally, the study's focus on individual regulation leaves room for deeper exploration of co-regulated and socially shared regulation, especially in environments where collaborative or community-based learning is prominent.

From a practical standpoint, the findings demand a critical reassessment of online instructional design. The tendency to treat SRL as an internal learner problem has allowed platforms to outsource responsibility for regulation to the user, resulting in systems that reward resilience rather than foster it. A more ethically sound approach would involve embedding adaptive supports that respond to learners' regulatory profiles, such as providing structure and pacing tools for Planners, feedback visualisations for Clarifiers, and interface improvements for Frustrated Navigators, while also addressing shared needs for clarity, engagement, and emotional support identified in cluster analysis. These are not optional enhancements; they are prerequisites for equitable participation in digital learning.

In conclusion, this study contributes to a growing body of evidence that SRL in online environments is not merely a psychological construct but a situated, affective, and socially mediated phenomenon. It cautions against over-reliance on idealised models of regulation divorced from their implementation context and instead advocates for a more critical, design-informed perspective. Future research and design should embrace the complexity of SRL, not to control it, but to scaffold it intelligently, responsively, and justly.

References

- Alghamdi, S. S., Bull, C., & Kharrufa, A. (2025). Designing for self-regulation in informal programming learning: Insights from a storytelling-centric approach [Preprint]. *arXiv*. <https://doi.org/10.48550/arXiv.2507.22671>
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/j.iheduc.2015.04.007>
- Chatti, M. A., Yücepur, V., Muslim, A., Guesmi, M., & Joarder, S. (2023). Designing theory-driven analytics-enhanced self-regulated learning applications. *arXiv*. <https://arxiv.org/abs/2303.12388>
- Cho, M.-H., & Heron, M. L. (2015). Self-regulated learning: The role of motivation, emotion, and use of learning strategies in students' learning experiences in a self-paced online mathematics course. *Distance Education*, 36(1), 80–99. <https://doi.org/10.1080/01587919.2015.1019963>

- Dai, Y., Luan, L., & Lin, X. (2023). The Effects of Online Learning Readiness on Self-Regulated Learning for the First-Time Online Learning Students. *Asian Journal of Distance Education*, 18(2), 42-62.
- D'Mello, S., Dieterle, E., & Duckworth, A. (2020). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist*, 55(1), 54–70.
<https://doi.org/10.1080/00461520.2020.1727529>
- Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, 46(1), 6–25.
<https://doi.org/10.1080/00461520.2011.538645>
- Greene, J. A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist*, 45(4), 203–209. <https://doi.org/10.1080/00461520.2010.515935>
- Hadwin, A. F., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and socially shared regulation of learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 83–106). Routledge.
<https://doi.org/10.4324/9781315697048-6>
- Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2019). What multimodal data can tell us about self-regulated learning? *Learning and Instruction*, 101268.
<https://doi.org/10.1016/j.learninstruc.2019.04.004>
- Khat, H. (2017). Academic performance and the practice of self-directed learning: The adult learner perspective. *Journal of Further and Higher Education*, 41(1), 44–59.
<https://doi.org/10.1080/0309877X.2015.1062840>
- Lämsä, J., de Mooij, S., Aksela, O., Athavale, S., Bistolfi, I., Azevedo, R., Bannert, M., Gasevic, D., Molenaar, I., & Järvelä, S. (2025). Measuring secondary education students' self-regulated learning processes with digital trace data. *Learning and Individual Differences*, 118, Article 102625. <https://doi.org/10.1016/j.lindif.2024.102625>
- Lobos, K., Bruna Jofré, D., & Santana, J. (2024). New challenges for higher education: Self-regulated learning in blended learning contexts. *Frontiers in Education*, 9, 1457367.
<https://doi.org/10.3389/feduc.2024.1457367>
- Martin, F., Long, S., Haywood, K., & Xie, K. (2025). Digital distractions in education: A systematic review of research on causes, consequences and prevention strategies. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-025-10550-6>
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422. <https://doi.org/10.3389/fpsyg.2017.00422>
- Pintrich, P. R. (2000). "The role of goal orientation in self-regulated learning" in *Handbook of Self-Regulation*. eds. M. Boekaerts, P. R. Pintrich, and M. Zeidner (Amsterdam: Elsevier), 451–502.
- Prasse, D., Sailer, M., & Fischer, F. (2024). Designing for self-regulated learning with technology: An umbrella review of systematic reviews and meta-analyses. *Educational Technology Research and Development*, 72, 57–83. <https://doi.org/10.1007/s11423-023-10301-2>
- Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, 2(1), 7–12.
<https://doi.org/10.18608/jla.2015.21.2>
- Steinert, S., Avila, K. E., Ruzika, S., Kuhn, J., & Küchemann, S. (2023). Harnessing large language models to enhance self-regulated learning via formative feedback. *arXiv preprint arXiv:2311.13984*.

- Wang, C.-H., Salisbury-Glennon, J. D., Dai, Y., Lee, S., & Dong, J. (2022). Empowering college students to decrease digital distraction through the use of self-regulated learning strategies. *Contemporary Educational Technology*, 14(4), ep388. <https://doi.org/10.30935/cedtech/12456>
- Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (pp. 297–314). Lawrence Erlbaum.
- Yu, B. (2023). Self-regulated learning: A key factor in the effectiveness of online learning for second language learners. *Frontiers in Psychology*, 13, 1051349. <https://doi.org/10.3389/fpsyg.2022.1051349>
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50031-7>
- Zimmerman, B. J., & Schunk, D. H. (2011). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Routledge. <https://doi.org/10.4324/9781410601032>

About the Author(s):

Henry, Khat; International Chamber of Commerce, France, henrykhat@gmail.com, [ORCID](https://orcid.org/0000-0002-8679-8695) number: <https://orcid.org/0000-0002-8679-8695>

Author's Contributions ([CRediT](#))

Henry Khat: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

The author has read and agreed to the published version of the manuscript.

Sustainable Development Goals ([SDGs](#))

This study is linked to the following SDG(s): Quality education (SDG 4)

Authors' Disclosures

This paper was proofread and edited with the assistance of ChatGPT 5 (as of Oct 2025), complementing the human editorial process. The human author critically assessed and validated the content to maintain academic rigor. The author also assessed and addressed potential biases inherent in the AI-generated content. The final version of the paper is the sole responsibility of the human author.

Data Accessibility Statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics and Consent

All procedures complied with institutional data protection guidelines; formal ethics approval was not required due to the use of archival, anonymized data.

Funding Information

Not applicable.

Acknowledgements

Not applicable.

Competing Interests

The author has no competing interests to declare.

Article History

Received: 03/11/2025 – Accepted: 03/09/2026

Suggested citation:

Khat, H. (2026). Rethinking Self-Regulated Learning in Online Education: Typologies, Affective Dynamics, and Design Implications from Learner Feedback. *Asian Journal of Distance Education*, 21(2), 78–111. <https://doi.org/10.5281/zenodo.22286048>



Authors retain copyright. Articles published under a Creative Commons Attribution 4.0 (CC-BY) International License. This licence allows this work to be copied, distributed, remixed, transformed, and built upon for any purpose provided that appropriate attribution is given, a link is provided to the license, and changes made were indicated.