



American Journal of Economics and Business Innovation (AJEBI)

ISSN: 2831-5588 (ONLINE), 2832-4862 (PRINT)

VOLUME 5 ISSUE 3 (2026)

**PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA**

Self-Directed Learning and Critical Thinking in Higher Education: A Critical Literature Review

Hasan Ahmad^{1*}, Md Anwar Hossain²

Article Information

Received: June 25, 2026

Accepted: August 09, 2026

Published: September 05, 2026

Keywords

Critical Thinking, Higher Education, Literature Review, Problem-Based Learning, Self-Directed Learning

ABSTRACT

This review integrates 41 peer-reviewed sources of literature related to self-directed learning (SDL) and critical thinking (CT) in higher education with a critical analysis of the novelty of each source's contribution, as well as the theoretical underpinnings, measurement, and empirical evidence of each. The papers were selected by performing a structured search with the search terms SDL and CT in three databases (Google Scholar, Scopus and Web of Science), and filtering for relevance to higher education, as well as for peer-review. To make it relevant to higher education, the search was limited to the past 15 years and to the most cited papers (including seminal papers) and, to make it peer-reviewed, limited to peer-review journals. No pooling of papers took place; papers were synthesized thematically. The review reveals that SDL and CT are consistently conceptualized as interdependent, learnable competencies that are influenced by pedagogical design and social context more than by personal traits, and that instructional strategies like PBL have small to moderate positive effects on the development of CT, with effects of SDL interventions remaining modest and varied across settings. The common limitations are the use of self-report questionnaires (usually the Self-Directed Learning Readiness Scale [SDIRS]), correlational and cross-sectional research designs, poorly formulated definitions of SDL, CT, and self-regulated learning, and samples confined to the North American, European, East Asian, and health-professionals. Most recent publications are based on current frameworks instead of offering new theory; they adapt these to new delivery methods, including: online, blended, and AI assisted learning. In the review, it is concluded that there is a great need to develop further objectivity based CT and realistic SDL practice-based designs, which are validated in the field.

INTRODUCTION

A literature review is not just a summary of the publications about a topic being researched, it is also a critical evaluation of the contribution and novelty of evidence, the way in which evidence was gathered and the methodological weaknesses and limitations, it is also a demonstration of gaps in the evidence, (Vaičiūnienė & Kazlauskienė, 2023). In the policy and pedagogical world, “self-directed learning (SDL) skill” and “critical thinking (CT) skill” have become mainstream concepts, with both being identified as important skills for graduates and both increasingly being expected to be developed within online and blended learning models, and now in the new context of AI-supported learning, where control of pacing, selection of resources, and self-monitoring are increasingly shifting to the learner (Rui *et al.*, 2024; Melisa *et al.*, 2025). This institutional shift is complemented by a legacy of research; the ideas of SDL and CT have been underpinned by 40 years of foundational theories and empirical studies since Knowles' (1978) work on andragogy versus pedagogy, to the recent systematic reviews on the effects of generative AI on students' analytical skills (Melisa *et al.*, 2025).

The purpose of this review is to summarize and synthesize 41 peer-reviewed articles with theoretical accounts and

background information about the evolution of SDL and CT in higher education and to review the empirical data supporting the constructs: instruments for measuring these constructs, and critically assess the contribution these articles make to the field, differentiating between actual conceptual advances and literature that is mainly a mere translation of existing models into new contexts. Although they are drawn from different theoretical fields, SDL from adult or continuing education (Loeng, 2020; Knowles, 1978) and CT from cognitive and educational psychology, a significant number of empirical literature covered in this review draw links between the two, suggesting that they reinforce each other or overlap (Turan & Koç, 2018; Rui *et al.*, 2024).

There are three objectives for this review. Firstly, a theoretical vocabulary is plotted in literature and secondly it is established whether recent literature is based on the previous models or repeats them. Second, to evaluate the empirical evidence base, particularly the research design, sample characteristics and instruments measuring SDL and CT. Thirdly, to identify areas of literature that has yielded reasonably certain conclusions and gaps that can be useful in conducting further primary research. The Literature Review is short as it is not included here due to the convention of this Journal.

¹ Lecturer, London Academy for Applied Technology, United Kingdom

² Senior Lecturer, Bath Spa University, United Kingdom

* Corresponding author's e-mail: hasan.ahmad@laat.ac.uk

LITERATURE REVIEW

Theoretical Foundations

The theoretical vocabulary of SDL is based on the adult learning theory. Knowles (1978) and later others (Loeng, 2020) proposed the seminal difference between pedagogy and andragogy: from the assumption that the teacher is responsible for what and how a learner comes to know to the learner's self-concept moving from dependence towards increased self-direction. This early conceptual work was further developed by Brookfield (1985), who considers self-direction as a methodology that is different from andragogy and that SDL should be studied as a process that takes place within, rather than outside, social and institutional settings (which later empirical studies of the social embeddedness of SDL, such as Kim&Yang (2020), confirmed).

Morris (2019) gives a more general definition of SDL as one of the fundamental competencies, and suggests that formal education (teacher-centred) needs to be complemented by more learner-centred teaching methods, though he doesn't elaborate on these in any great detail, this is more of a conceptual contribution with limited operational specificity. Zimmerman's (1990, 2002) social-cognitive model of self-regulated learning as a three-stage process of forethought, performance and self-reflection has been used repeatedly for more than three decades, reflecting its status as a true conceptual contribution, with the ability to provide a common testable vocabulary for a phenomenon previously described impressionistically. Grow (1991) however offered a less comprehensive model, the Staged Self-Directed Learning model, which proposed that learners pass through four stages of dependence to self-direction and that teaching should be geared to the learner's stage; the model was replicated in teaching guides, but there is limited evidence of its actual operation and the four stages have not been compared to longitudinal data.

As mentioned earlier, the majority of quantitative studies reviewed here employed a version of the Self-Directed Learning Readiness Scale (SDLRS) or a similar self-report instrument; the first attempts to measure SDL as a construct instead of an ideal were methodological and not theoretical, and a SDLRS was validated by Long and Agyekum (1983); a second study was undertaken in a work environment, reported by Guglielmino, Guglielmino and Long (1987), is explored below. While both Robinson (2020) and Loeng (2020) restate this older theory in the context of a current audience in the fields of pharmacy and adult-education respectively, they do not offer any new mechanisms.

Instrumentation and Measurement Development

In the original SDLRS, however, there is no interest in theorising SDL, but rather developing and testing instruments of measurement of SDL, and this sub-literature has been allowed to grow beyond the scope of the original SDLRS. Teo *et al.* (2010) created and tested the Self-Directed Learning with Technology

Scale (SDLT) for younger learners, which builds upon the tradition of measuring SDL and was designed to apply to technology-mediated learning settings, a setting increasingly emphasized with the advent of online and blended learning. The constructs of readiness identified on the original scale by Guglielmino were not specific to one profession and Fisher *et al.* (2001) developed a separate scale for the readiness in nursing education. Kumar *et al.* (2021) validated the SDLRS internationally on medical students in India, thus showing that the instrument has been validated in new cultural and disciplinary contexts but not by a different measurement approach. More recently, Castillo (2025) demonstrated with a sample of 300 nursing students in the Philippines that students' perceptions of their clinical-education environment had a significant impact on their SDLRS measured readiness, further reinforcing the reading of the construct as both situational and dispositional, and which is found throughout the sub-literature.

Recently, systematic reviews have focused on the instrument itself and scrutinized it more rigorously. As mentioned above, there were concerns regarding the definition of SDL and Taylor *et al.* (2023) did a systematic review of SDL assessment practices in undergraduate health-professions education and found significant variation in how SDL is operated and assessed within programmes. With regards to the introduction of SDL practices, Berger-Estilita *et al.* (2025), based on a mixed methods systematic review of thirty-four studies and over 5700 health professionals, reported that besides individual learner disposition, 'cultural context of the organisation is also a significant factor in the success of SDL practice. The study by Wilkes *et al.* (2026) was the most methodologically advanced that we found in this review: a three-level random-effects meta-analysis of the effects of SDL in health-professions education, this included 74 effect sizes from 48 studies with an overall effect (Cohen's $d = 0.34$) that was small to medium, and with lots of heterogeneity; they note that the effects of SDL as an intentional learning intervention were significantly larger than as an unstructured learner attribute. It was one of the more important findings of the literature that attempted to quantify SDL, which will turn up again later in the Results and Discussion portions of this paper.

Empirical Evidence in Higher Education

Recent empirical research in the last decade has moved the definitional centre of gravity to closer align with the contextual factors that enable it to flourish. While the earlier syntheses were single factor, one of the most detailed syntheses to date was the systematic review by Vaičiūnienė and Kazlauskienė (2023) which systematically reviewed 27 studies, distinguishing three types of oppressive and two types of liberating factors: the learning environment, knowledge possessed by the learner, and the proactivity of the learner, and considers oppressive factors analytically distinct from

liberating ones, not merely their absence. The only truly longitudinal study in this literature is the study by Kim and Yang (2020) who studied medical students over time and concluded that the concept of group cohesion was more predictive of the medical students' preparedness to use new learning strategies than individual attitudes. This is a unique outcome, as it considers SDL not just as an individual attitude but as a socially constructed attitude, unlike the attitude that is assumed when using the SDLRS approach.

The literature review carried out by Rui *et al.* (2024) revealed that SDL's emphasis on learner autonomy and self-monitoring is well aligned with deep learning processes as the latter feature learners' self-reflection and integrative approach to content, with most research conducted on the processes relying on self-report measures or qualitative sources, rather than behavioral measures of outcomes. Aulakh *et al.* (2025) conducted a systematic review and meta-analysis of 19 studies to compare the effectiveness of SDL with traditional didactic learning in the specific setting of undergraduate medical education, reporting moderate, inconsistent, benefits of SDL and cautioning that the diversity of the SDL interventions in these studies hinders the ability to draw confident conclusions from any one of the reported effect size estimates.

The role of Internet and technology mediated environment is becoming more prominent. Doo *et al.* (2023) conducted a meta-analysis of the impact of SDL on learning outcomes in MOOCs and reported a small but positive overall effect, much like the earlier studies, most of the SDL measured included in the meta-analysis were self-reported. In another study, Xu *et al.* (2023) reviewed existing literature on the construct related to SRL (self-regulated learning) and academic achievement in online and blended learning contexts by conducting a scoping review of the literature, which revealed that there has been an increase in research over the years, but a scattered distribution of research across disciplines and measurement types. There is a considerable overlap between the two reviews: Broadbent and Poon (2015) and Luo and Zhou (2024), and it is not unreasonable to ask if there is any new explanation to be added, in which case much of the value of Luo and Zhou is in summarizing what Broadbent and Poon had already provided.

The critical-thinking aspect of literature is more of a battleground. The results of the current study indicated that there was a positive relationship between SDL readiness and critical thinking disposition among sports-science students, which may be viewed as a reinforcing relationship between SDL readiness and critical thinking disposition, however, because of the correlational design of the study, it was not possible to determine direction of causality, and because of the sample of a single discipline and a single country, generalisability of the results is limited. In a more recent correlational study conducted in a higher education context in the Philippines, students' grit and willingness to communicate in English were

all again shown to be good predictors of their critical-thinking scores (Lagunday *et al.*, 2025), thus reiterating the more general pattern in which non-cognitive and affective factors – not just instructional design – consistently were implicated in CT outcomes, albeit in this case again precluding any causal claim. There is a more mixed review of the CT teaching in higher education in the studies done by Behar-Horenstein and Niu (2011), whose results indicated that the average effect size for the majority of the interventions reviewed was statistically significant but varied widely across studies, which means that certain interventions cannot be used with confidence. It is difficult not to agree with the OECD-commissioned report by Van Damme and Zahner (2022) on the scale of the report and that they didn't go for any single pedagogical tradition, but that a standardised test was applied to a variety of countries and that the conclusion was that, throughout a full degree programme, critical thinking skills were little to no better around one-fifth of the students in the study. The result is not contradictory to two smaller intervention studies, such as the study by Witarsa and Muhammad (2023) which showed significant increase in CT achievement of social-science students as the result of project-based learning, where the results of small-scale intervention studies are generally less generalizable than those of large-scale intervention studies.

Critical Thinking Pedagogy and Intervention Studies

A body of research that specifically evaluates various forms of CT instruction, such as problem-based learning (PBL) is growing. Abrami *et al.* (2015) found a weighted mean effect size of 0.30 using standardised CT measures in a meta-analysis of 341 effect sizes from 39 quasi- and true-experimental studies, and recommended both generic and content-specific CT instruction is effective when it is not incidental but is explicit. This is one of the methodologically soundest sources reviewed in part because it is not strictly a narrative synthesis, but rather deals with a number of primary studies.

The most empirical research has occurred in health-professions education education using PBL. A systematic review by Ge *et al.* (2025) included 13 previous systematic reviews and meta-analyses on the effect of PBL on clinical skills and CT; the evidence of these effects is strong, but they recommend that this evidence needs to be strengthened due to the differences in quality of the included studies across the reviews. Su *et al.* (2025) conducted a more narrow-focused systematic review and meta-analysis of the effectiveness of PBL on CT skills, and while their overall conclusion was positive, their findings were broadly similar to those of Ge *et al.* (2025). Outside the medical field, Yu and Zin (2023) conducted a systematic review of CT oriented adaptations of PBL models in general, and Anggraeni *et al.* (2023) conducted a systematic review of PBL research that promotes CT skills, both of which concluded that the effectiveness of PBL is not solely the use of the PBL label, but it is

also greatly influenced by how the problem component is implemented in the ill-structured of PBL in the inquiry component.

A parallel line of enquiry is related to the mechanisms or processes of cognition as opposed to the pedagogical model that is being used to have students learn CT. Rivas *et al.* (2022) used the ARDESOS-DIAPROVE metacognitive and problem-based instructional programme to first year students in psychology and concluded that explicit instruction of critical thinking not only led to a significant increase in students' CT scores, but also in metacognitive regulation, which suggested a complementary relationship between CT and metacognition. On a higher level of synthesis, Andreucci-Annunziata *et al.* (2023) reviewed five previous systematic reviews (SLRs) of SLRs regarding CT in higher education to synthesize the relevant literature using the PRISMA 2020 guidelines, and found that CT is still not well conceived in the literature but that four pedagogical approaches — general method, infusion, immersion, mixed method — are recurring pedagogical strategies that can be mapped onto most instructional interventions. Their work is a valuable methodological contribution because their overall structure and contribution to a broken evidence base rather than providing a further isolate of a single study.

Discipline-specific literatures enhance the writing and provide some tension. Calma and Davies (2021) offer a practical, measurable theoretical framework of CT skills in business education, partly as a response to the lack of agreement on the definition of CT and CT assessment found by Calma and Cotronei-Baird (2021) and this is a rare case in the literature where the focus of a methodological critique is addressed in a later paper by the same authors, rather than simply restated by a subsequent reviewer. In his 2013 systematic review of CT in nursing education, Chan has come up with broadly similar findings to Calma and Cotronei-Baird's in business-education, indicating that the issues with the definition and assessment of CT are not unique to a particular field. In terms of leadership development, Flores *et al.* (2012) conducted a study of the CT skills of college graduates 10 years prior to the findings reported by the OECD, and discovered that "many college graduates do not leave their college with evidence of CT skills development." For their recent systematic review on the impact of generative AI (GAI) tools like ChatGPT on students' CT development, Melisa *et al.* (2025) retrieved the literature in the Scopus and ERIC databases for the period 2023-2024, finding that

although there has been a lot of interest in students' own production of AI-generated content, there is also a concurrent and growing concern that this may have an impact on their independent evaluative engagement with information, but that the evidence base is rather recent and largely perception-based as opposed to behaviourally measured. This is the most common weakness. This is the most frequent weakness.

Most sources examined here have a shared limitation of the use of self-report measures (in the form of SDLRS variants or CT disposition scales). The SDLRS is currently still the preferred measurement instrument in clinical research, probably because they found it to be an effective instrument or because there is no other instrument that is as interesting for researchers. Indeed, it has been adapted in various ways, with only slight changes in the wording or content of the scale, but not in the self-report logic used by the scale (Fisher *et al.*, 2001; Kumar *et al.*, 2021; Berger-Estilita *et al.*, 2025). It also has social-desirability bias and the only qualitative study included in this group, (Douglass & Morris, 2014), found that students' self-reported SD did not necessarily correspond with the observed SD in the classroom, which is not captured in quantitative, SDLRS-based designs. This is not just a technical footnote; in fact, practically every analysis of the effectiveness of intervention found in this literature uses a self-perception measure for intervention, and Wilkes *et al.* (2026) only partially addresses this by dividing intervention into SDL-as-designed and SDL-as-assumed-trait; it does not distinguish between these two forms of intervention and the two ways in which they might be behaviourally measured.

MATERIALS AND METHODS

Search Strategy and Keyword Assembly

The search of sources was conducted with a structured search strategy on Google Scholar, Scopus, and Web of science with a combination of terms from Concept Block A, and Concept Block B terms with Boolean operators, as in table 2, which is in line with other literature in this field (e.g., Luo and Zhou, 2024). It was not a fixed date cut-off but rather highly cited seminal theoretical and instrument-validation research papers that continued to be influential in the empirical research conducted today were sought for, as were more recent systematic reviews and meta-analyses (as early as 20 or so years old) in order to capture the conceptual and historical roots and current state of the field.

Table 2: Keyword assembly structure.

Concept block A	Boolean operator	Concept block B
"self-directed learning"	AND	"higher education"
"self-regulated learning"	OR / AND	"university"
"critical thinking"	OR / AND	"college"

Inclusion and Exclusion Criteria

The sources retrieved with the search above were screened using the inclusion/exclusion criteria listed in Table 1. This initial search yielded 356 records, from which 112 duplicates were removed, and 244 records were then subject to title and abstract screening, with 153 of these records not meeting the topical relevance and peer-reviewed status criteria and leaving 91 full-text articles for eligibility screening against the criteria listed in Table 1. Fifty other full-text articles were excluded most often for being discussed only in the context of SDL or CT, or for not being in an HE or adult-education context,

or for a lack of a substantive research design or synthesis, leaving 41 sources for thematic synthesis. Because it is a critical but not systematic review, screening and eligibility were done by only one reviewer, and there was no step for interrater reliability, the numbers in Figure 1 should be interpreted as a clear documentation of the selection process, not as proof of interrater reliability. The sources were chosen by design, and, by design, all of them were read by one reviewer; they included a wide range of types of sources and contributions found in the field, not just a single “study design.

Table 1: Inclusion and exclusion criteria for article selection.

Inclusion criteria	Exclusion criteria
Directly addresses self-directed learning (SDL) and/or critical thinking (CT)	Addresses SDL/CT only in passing, with no substantive theoretical or empirical treatment
Set within a higher-education, professional-education, or adult-education context	Focused solely on primary or secondary (K–12) education with no higher-education relevance
Peer-reviewed journal article, systematic review, meta-analysis, or peer-reviewed conference paper	Non-peer-reviewed sources (e.g., blog posts, unrefered working papers, trade publications)
Published in English	Not available in English and no reliable translation accessible
Theoretical, instrument-validation, empirical, or synthesis (review/meta-analytic) contribution	Purely descriptive institutional reports with no research design or synthesis
Full text accessible for reading and thematic coding	Full text not accessible through institutional or open-access channels

Selection Process and Thematic Coding

The whole selection process is displayed in Figure 1. This screening process led to the selection of each source

based on (i) the theoretical model or definition of SDL and/or CT included, (ii) the research design used and the characteristics of the sample, (iii) the instrument used

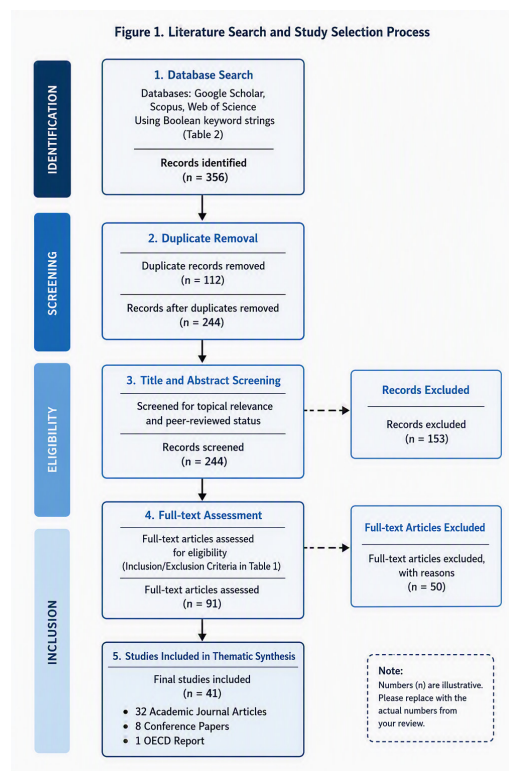


Figure 1: Literature search and study selection process.

for measurement (if one was included), and (iv) the main findings. This mixed design sampling strategy is suitable for a critical evaluative review that aims to assess the novelty and contribution of existing syntheses of primary studies, but not to combine primary studies and report a new quantitative synthesis (i.e. a new effect size) (for example, Abrami *et al.*, 2015; Wilkes *et al.*, 2026). Where quantitative effect sizes are reported below (e.g. Abrami *et al.*, 2015; Wilkes *et al.*, 2026), these are cited from the meta-analyses and do not represent new calculations. For purposes of comparative evaluation of novelty, which is the primary goal of the purpose of this literature review, the sources were themed into the subsections above (in the Literature Review) and categorized as conceptual, methodological or contextual/translational (in the Results and Discussion section below (Table 3).

RESULTS AND DISCUSSION

In the forty-one sources examined here, there are three general contributions—and it is well to differentiate among them, if not to be excused for an exaggeration for saying that the field has made great strides. The former is the conceptual contribution, a new, real, explanatory mechanism. The models most evidently depicting such movement are Zimmerman’s (1990) cyclical self-regulation model, as well as Kim and Yang’s (2020) social-cohesion finding. The second type is

methodological contribution: the development and/or validation of a measurement instrument such as Long and Agyekum (1983), Guglielmino *et al.* (1987) and more recently instrument focused studies like Teo *et al.* (2010), Fisher *et al.* (2001) and Kumar *et al.* (2021) that are shaping empirical studies many years – and in the most recent cases, just years – after their publication. The third one that was most prevalent in the last 5 years is the contextual or translational contribution, i.e., the model is used in a new context, discipline or delivery, i.e., pharmacy education, blended-learning delivery, and MOOC-based learning, as introduced in Robinson and Persky (2020), Loeng (2020), Luo and Zhou (2024), Doo *et al.* (2023), and Rui *et al.* (2024). None of these three types of contributions are less valuable, but the review cannot be considered truly original, if it proposes them as equal types of contributions, especially the context-based and the translational ones, which are now rather frequent, and often involve the use of existing SDL models and CT approaches with reference to online, blended and current AI-mediated learning.

This typology is summarised in table 3 against the main sources consulted, to ensure that the distribution of types of contribution are not implicit in the narrative above.

If the sources are read sequentially, without interpretation, a fairly coherent picture emerges, based solely on the evidence of the intervention. The positive small-to

Table 3: Typology of contribution across the reviewed literature.

Contribution type	Representative sources	Nature of contribution
Conceptual	Zimmerman (1990, 2002); Kim&Yang (2020)	Introduces a genuinely new explanatory mechanism (cyclical self-regulation; social cohesion as a driver of SDL)
Methodological	Long&Agyekum (1983); Guglielmino <i>et al.</i> (1987); Teo <i>et al.</i> (2010); Fisher <i>et al.</i> (2001); Kumar <i>et al.</i> (2021)	Develops or validates a measurement instrument that subsequently shapes empirical studies
Contextual / translational	Robinson&Persky (2020); Loeng (2020); Luo&Zhou (2024); Doo <i>et al.</i> (2023); Rui <i>et al.</i> (2024); Melisa <i>et al.</i> (2025)	Applies an existing model or instrument to a new discipline, delivery mode, or technology without altering its core structure
Large-scale synthesis	Abrami <i>et al.</i> (2015); Van Damme&Zahner (2022); Ge <i>et al.</i> (2025); Wilkes <i>et al.</i> (2026)	Aggregates effect sizes or standardised assessment data across many primary studies or countries

moderate effect on CT outcomes evidenced by multiple independent meta-analyses (Abrami *et al.*, 2015; Ge *et al.*, 2025; Su *et al.*, 2025), is more convincing than a single study, as PBL is the most widely studied approach to structured, explicit CT instruction. In the SDL literature, however, the small-to-moderate effect is broadly comparable to that seen in the CT-instruction literature, but there is a greater degree of heterogeneity, and the difference between an SDL instructional intervention and an SDL treated as an assumed learner characteristic that Wilkes *et al.* (2026) explain is a plausible reason why the SDL literature has historically been more variable

in terms of effect estimates than the CT-instruction literature, many of the earlier SDL studies reviewed here included. This definition was largely not mentioned in the previous study on SDL carried out by Vaičiūnienė and Kazlauskienė (2023) and others, but is one of the more useful clarifications that have come from the more recent and methodologically more robust sources that are included in this review.

The second pattern relates to the disciplinary and geographical distribution of the evidence base. The reviews of literature show that health-professions education is over-represented in both the SDL literature

(Chan, 2013; Kim & Yang, 2020; Kumar *et al.*, 2021; Aulakh *et al.*, 2025; Berger-Estilita *et al.*, 2025; Ge *et al.*, 2025; Su *et al.*, 2025; Wilkes *et al.*, 2026) and, perhaps, the PBL literature (Chan, 2013; Kim & Yang, 2020; Aulakh *et al.*, 2025; Wilkes *et al.*, 2026) because clinical education has long been associated with PBL as a signature pedagogy, and there are formal health-professions accreditation requirements to document both CT and SDL outcomes. Comparatively under-represented are the areas of business education (Calma & Davies, 2021; Calma & Cotronei-Baird, 2021), general higher education (Behar-Horenstein & Niu, 2011; Abrami *et al.*, 2015), and a few social-science and humanities studies (Witarsa & Muhammad, 2023); there are very few syntheses across disciplines. When interpreting effect sizes it is important to consider that many of the highest quality quantitative studies (Abrami *et al.*, 2015; Ge *et al.*, 2025; Wilkes *et al.*, 2026) have been carried out in populations of health professionals, with the importance of generalisability to other disciplines like Engineering and the Arts not assumed.

A third pattern is definitional inconsistency, as it arises as a criticism in the literature, and not just in this particular sub-field. At the business-education level Calma and Davies (2021) state explicitly that they see no consensus on the definition; at the level of the general systematic reviews of CT there seems to be an equivalent lack of consensus definition, according to Andreucci-Annunziata *et al.* (2023), although decades of research have been conducted, and there appears to be an even greater lack of definition of SDL at the level of health-professions education, as reported by Taylor *et al.* (2023). The continued criticism in these three disciplines, with their different methodologies, makes it more of a structural problem in the field and is unlikely to be resolved by a single new instrument or definitional paper, but rather will require coordinated action of the journals or the profession to agree on common definitions and reporting standards.

Finally, the influence that the research field is subject to and the weaknesses it has are reflected in the presence of the new research area of generative AI in the reviewed literature (Melisa *et al.*, 2025). Conversely, a systematic review in a rapidly evolving application field such as AI's impact on CT is possible in just approximately two years of the mainstream use of ChatGPT, demonstrating the ability to rapidly move to a new application context when it is a new one. However, the Melisa *et al.* (2025) review is entirely perception-based and early-adoption evidence, and lacks behaviourally validated CT outcome measures, as has been documented to be the most common methodological weakness in both 40-year-long SDL and CT research.

The other implication is that the field should not be the average of the sources simply but also that the disagreement between the sources should be taken into account. In other findings the differences between the two independent reviews (PBL, generally positive, as in Ge *et al.*, 2025 and clinical skills, more qualified

and discipline-specific, as in Yu and Zin, 2023 and Anggraeni *et al.*, 2023) seem more like information than noise: the differences could be due to the quality of implementation and not to the pedagogical logic of PBL. The ramifications for institutions that provide foundation education and undergraduate business education are clear: unless the effects of problem-based learning are present in the institutions' module through the following design elements, then the effects are unlikely to be replicated: explicit instruction, structured facilitation and ill-structured problems, the three elements in the strongest syntheses (Abrami *et al.*, 2015; Ge *et al.*, 2025). Another related question, also largely under debate, is transfer: Do gains in critical thinking that were made in a particular domain of content or instruction transfer to the critical thinking in a context outside of this domain? Or are they skills taught in a particular discipline that are superficially similar to critical thinking but picked up in the context where they are being used? (Rivas *et al.*, 2022; Su *et al.*, 2025) This review does not include any studies that directly assess transfer, but rather retest the same or a related skill in the same domain in which it was taught (Rivas *et al.*, 2022; Turan & Koç, 2018). At the level of scale, Van Damme and Zahner (2022) present a cross-national design that is standardised in their assessments, with the least overall findings, as would be expected if smaller, single-institution intervention studies overestimate their results because the outcome measure is more closely tied to the instructional content than a general purpose standardised assessment.

Together, these patterns suggest that the future of the field will be driven more by the research that can easily and quickly compare intervention fidelity across sites and test transfer to new domains with behaviourally validated measures, and report effect sizes in a format that can be easily pooled into clean future metaanalyses (e.g., Abrami *et al.*, 2015; Ge *et al.*, 2025; Wilkes *et al.*, 2026). It is a more general observation than the more specific measurement gap described in the Conclusion below, but related: Cross-study comparisons, as now exists in the field will not be possible without a better way to measure the outcomes.

Contribution of this Review

The chief value of this review, then, is comparative and evaluative, rather than empirical, as it does not report an original primary study, but rather does that which the individual sources it draws on does not do for itself: it puts forty-one sources side by side, and classifies each of them according to the kind of contribution(s) it actually makes. Table 3, which classifies novel contributions as conceptual, methodological and contextual/translational, is a transferable typology that can be used to draw conclusions about novel contributions in the current literature, and the analysis of the novel contributions in this literature using this typology shows that there are fewer novel contributions in terms of theory or measurement than may seem to be suggested by the number of recent publications.

A second contribution is to make the SDL and CT literatures come together and not to regard them as two distinct sets of texts as most prior reviews have done. CT has been shown to produce more consistent results across independent meta-analyses than do effects of intervention research in SDL, as shown when analysed separately as this review does in its Literature Review subsections; this difference may be due to methodological variations across the independent meta-analysis studies in the way the two constructs are operationalised, but not due to any fundamental difference between the learnabilities of these two skills. This is not evident in either reading body stand alone, but emerges when the evidence bases for the two are compared side-by-side in the Discussion.

A third contribution is to acknowledge, as has been pointed out by three separate reviews in the three distinct domains of business education (Calma & Davies, 2021); general higher education (Andreucci-Annunziata *et al.*, 2023); and health professions (Taylor *et al.*, 2023), that the definitional inconsistencies in this field are not a matter of one discipline, but are structural. The practical significance of this is that a problem arising in any one of the three literatures is unlikely to be resolved by a single definitional paper in that one literature, but more likely by a coordinated action of journal editors or of professional bodies that was not suggested when any one of the three source reviews is read in isolation.

Future Research Directions

Based on that synthesis, and the gaps and opportunities that have been identified in the Discussion, four directions for future primary research are most closely suggested by this synthesis. The first is the need for a study which uses a longitudinally designed and sustained SDL intervention, and which includes a behaviourally validated measure of CT (not a self-report disposition scale) so that any improvements observed in critical thinking skills could not be confused with correlational association, since no study has been found that combines these elements. Second, transfer studies are needed to verify that the improvements in critical thinking evidenced in one context transfer to another context (domain), and to verify that a different measure of critical thinking is used as the dependent measure as opposed to the use of the same measure of critical thinking or the same task in a different context where they had received instruction. Third, it is important to build the evidence base beyond health professions and business education into the humanities, social sciences, and creative disciplines, where claims of SDL and CT about graduate outcomes can be found, albeit comparatively, but where the evidence base is much weaker (Witarsa & Muhammad, 2023). Fourth, there is a need for research that brings a behavioural impact measure of generative AI on CT behaviour over time into the picture, and not only based on initial perceptual evidence, as thus far provided (Melisa *et al.*, 2025).

Practical And Societal Significance

The implications of this review for the design and justification of higher education curricula that assert the development of self-directed learning and critical thinking as outcomes of higher education are also important. To build new provision which is based on ‘problem-based learning’ or ‘self-directed learning’, without taking into account a specific set of design features, explicit instruction, structured facilitation and problems that are truly ill structured’ is an unattainable target and one that is not backed by the best available evidence and research. Where a programme team needs to show the development of graduate attributes (e.g., independent learning or analysis) through a module as part of a quality-assurance or accreditation process, the typology in the review might result in a range of curriculum claims being identified, some of which could be supported by good, convergent evidence or a single small study or assumed association. The review has a broader social dimension, however. Although critical thinking has emerged as a key employability skill that, together with self-direction, is increasingly seen as essential in the workplace (and as something that higher education is expected to provide), the OECD’s own cross-national assessment reviewed here (Van Damme & Zahner, 2022) concludes that measurable improvements in critical thinking over a full degree programme tend to be relatively small. It is not just an academic curiosity; it is an accurate understanding of the current evidence for what teaching practice can and cannot deliver versus an uncritical rehash of the possibly more promising individual study findings of the field for institutions, funders, and prospective students. The review’s conclusion that the impact of AI on critical thinking is measured almost exclusively by perception not behaviour in the moment is also timely and of public interest, as there is an absence of a robust evidence base to inform decisions about what is considered acceptable use of AI in assessment.

CONCLUSION

From the literature reviewed in this section, the following points are noted about SDL/critical thinking as a learned skill: The relationship between SDL and critical thinking as learned skills is mutually reinforcing and is not a necessary outcome of the characteristics of the learner; They are more likely to be determined by pedagogical design and social context than by learner characteristics; The conclusions are complex because the literature is characterized by self-report instruments, correlational or cross-sectional designs, inconsistent definitions, and a limited research base in some disciplines and world regions. Independent meta-analyses most consistently show positive effects of structured, explicit critical-thinking instruction (especially in the context of problem-based learning) and similar average effects for self-directed learning interventions, while there appears to be significant variation in effects, at least in part

because of the persistent confusion in the field regarding the ambiguity between self-directed learning as a designed intervention and self-directed learning as a personal trait. In fact, the vast majority of publications are extensions of previous theoretical models and measurement tools to be used in new contexts (web, blended, AI mediated learning). There are very few new original concepts or methods and most of the recent publications are simply new applications of existing concepts or measurement devices. Critical thinking does not appear as a methodological priority, but as a practical one, as the biggest need for future research is a long-term and cross-institutional study which can relate a behaviourally validated measure of critical thinking (not a self-reported instrument) with an authentic, sustained and self-directed learning practice, to which the gains in critical thinking can then be assigned, based on correlational relationships among the forty-one sources reviewed here.

REFERENCES

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research*, 85(2), 275–314. <https://doi.org/10.3102/0034654314551063>
- Andreucci-Annunziata, P., Riedemann, A., Cortés, S., Mellado, A., del Río, M. T., & Vega-Muñoz, A. (2023). Conceptualizations and instructional strategies on critical thinking in higher education: A systematic review of systematic reviews. *Frontiers in Education*, 8, Article 1141686. <https://doi.org/10.3389/educ.2023.1141686>
- Anggraeni, D. M., Prahani, B. K., Suprpto, N., Shofiyah, N., & Jatmiko, B. (2023). Systematic review of problem-based learning research in fostering critical thinking skills. *Thinking Skills and Creativity*, 49, Article 101334. <https://doi.org/10.1016/j.tsc.2023.101334>
- Aulakh, J., Wahab, H., Richards, C., Bidaisee, S., & Ramdass, P. V. A. K. (2025). Self-directed learning versus traditional didactic learning in undergraduate medical education: A systematic review and meta-analysis. *BMC Medical Education*, 25, Article 55. <https://doi.org/10.1186/s12909-024-06449-0>
- Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning*, 8(2), 25–42. <https://doi.org/10.19030/tlc.v8i2.3554>
- Berger-Estilita, J., Krista, L., Gogollari, A., Schmitz, F., Elfering, A., & Guttormsen, S. (2025). Self-directed learning in health professions: A mixed-methods systematic review of the literature. *PLOS ONE*, 20(5), Article e0320530. <https://doi.org/10.1371/journal.pone.0320530>
- 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>
- Brookfield, S. D. (1985). Self-directed learning: A conceptual and methodological exploration. *Studies in the Education of Adults*, 17(1), 19–32. <https://doi.org/10.1080/02660830.1985.11730445>
- Calma, A., & Cotronei-Baird, V. (2021). Assessing critical thinking in business education: Key issues and practical solutions. *The International Journal of Management Education*, 19(3), Article 100531. <https://doi.org/10.1016/j.ijme.2021.100531>
- Calma, A., & Davies, M. (2021). Critical thinking in business education: Current outlook and future prospects. *Studies in Higher Education*, 46(11), 2279–2295. <https://doi.org/10.1080/03075079.2020.1716324>
- Castillo, M. M. M. (2025). A parsimonious model of nursing students' clinical learning environment and self-directed learning: Basis for a self-directed learning development program. *American Journal of Physical Education and Health Science*, 3(1), 57–70. <https://doi.org/10.54536/ajpehs.v3i1.4807>
- Chan, Z. C. Y. (2013). A systematic review of critical thinking in nursing education. *Nurse Education Today*, 33(3), 236–240. <https://doi.org/10.1016/j.nedt.2013.01.007>
- Doo, M. Y., Zhu, M., & Bonk, C. J. (2023). Influence of self-directed learning on learning outcomes in MOOCs: A meta-analysis. *Distance Education*, 44(1), 86–105. <https://doi.org/10.1080/01587919.2022.2155618>
- Douglass, C., & Morris, S. R. (2014). Student perspectives on self-directed learning. *Journal of the Scholarship of Teaching and Learning*, 14(1), 13–25. <https://doi.org/10.14434/josotl.v14i1.3202>
- Fisher, M., King, J., & Tague, G. (2001). Development of a self-directed learning readiness scale for nursing education. *Nurse Education Today*, 21(7), 516–525. <https://doi.org/10.1054/nedt.2001.0589>
- Flores, K. L., Matkin, G. S., Burbach, M. E., Quinn, C. E., & Harding, H. E. (2012). Deficient critical thinking skills among college graduates: Implications for leadership. *Educational Philosophy and Theory*, 44(2), 212–230. <https://doi.org/10.1111/j.1469-5812.2010.00672.x>
- Ge, W. L., Zhu, X. Y., Lin, J. B., Jiang, J. J., Li, T., Lu, Y. F., Mi, Y. F., & Tung, T. H. (2025). Critical thinking and clinical skills by problem-based learning educational methods: An umbrella systematic review. *BMC Medical Education*, 25, Article 455. <https://doi.org/10.1186/s12909-025-06951-z>
- Grow, G. O. (1991). Teaching learners to be self-directed. *Adult Education Quarterly*, 41(3), 125–149. <https://doi.org/10.1177/0001848191041003001>
- Guglielmino, P. J., Guglielmino, L. M., & Long, H. B. (1987). Self-directed learning readiness and performance in the workplace. *Higher Education*, 16(3), 303–317. <https://doi.org/10.1007/BF00148972>
- Kim, S., & Yang, E. B. (2020). Does group cohesion foster self-directed learning for medical students? A

- longitudinal study. *BMC Medical Education*, 20, Article 55. <https://doi.org/10.1186/s12909-020-1962-7>
- Knowles, M. S. (1978). Andragogy: Adult learning theory in perspective. *Community College Review*, 5(3), 9–20. <https://doi.org/10.1177/009155217800500302>
- Kumar, A. P., Omprakash, A., Mani, P. K. C., Swaminathan, N., Maheshkumar, K., Maruthy, K. N., Sathiyasekaran, B. W. C., Vijayaraghavan, P. V., & Padmavathi, R. (2021). Validation of internal structure of the self-directed learning readiness scale among Indian medical students using factor analysis and the structural equation modelling approach. *BMC Medical Education*, 21, Article 641. <https://doi.org/10.1186/s12909-021-03035-6>
- Lagunday, B. J. D., Restor, M. A. D., Sasam, A. L., & Oliveros, J. C. (2025). Students' grit and willingness to communicate in English as predictors of critical thinking skills. *Journal of Tertiary Education and Learning*, 3(3), 23–36. <https://doi.org/10.54536/jtel.v3i3.5176>
- Loeng, S. (2020). Self-directed learning: A core concept in adult education. *Education Research International*, 2020, Article 3816132. <https://doi.org/10.1155/2020/3816132>
- Long, H. B., & Agyekum, S. K. (1983). Guglielmino's self-directed learning readiness scale: A validation study. *Higher Education*, 12(1), 77–87. <https://doi.org/10.1007/BF00140273>
- Luo, R. Z., & Zhou, Y. L. (2024). The effectiveness of self-regulated learning strategies in higher education blended learning: A five years systematic review. *Journal of Computer Assisted Learning*, 40(6), 3005–3029. <https://doi.org/10.1111/jcal.13052>
- Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Al Fuad, Z. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Educational Process: International Journal*, 14, Article e2025031. <https://doi.org/10.22521/edupij.2025.14.31>
- Morris, T. H. (2019). Self-directed learning: A fundamental competence in a rapidly changing world. *International Review of Education*, 65, 633–653. <https://doi.org/10.1007/s11159-019-09793-2>
- Rivas, S. F., Saiz, C., & Ossa, C. (2022). Metacognitive strategies and development of critical thinking in higher education. *Frontiers in Psychology*, 13, Article 913219. <https://doi.org/10.3389/fpsyg.2022.913219>
- Robinson, J. D., & Persky, A. M. (2020). Developing self-directed learners. *American Journal of Pharmaceutical Education*, 84(3), 292–296. <https://doi.org/10.5688/ajpe847512>
- Rui, L., Mohamad Nasri, N., & Mahmud, S. N. D. (2024). The role of self-directed learning in promoting deep learning processes: A systematic literature review. *F1000Research*, 13, Article 761. <https://doi.org/10.12688/f1000research.150612.1>
- Su, T., Liu, J., Meng, L., Luo, Y., Ke, Q., & Xie, L. (2025). The effectiveness of problem-based learning (PBL) in enhancing critical thinking skills in medical education: A systematic review and meta-analysis. *Frontiers in Education*, 10, Article 1565556. <https://doi.org/10.3389/feduc.2025.1565556>
- Taylor, T. A. H., Kemp, K., Mi, M., & Lerchenfeldt, S. (2023). Self-directed learning assessment practices in undergraduate health professions education: A systematic review. *Medical Education Online*, 28(1), Article 2189553. <https://doi.org/10.1080/10872981.2023.2189553>
- Teo, T., Tan, S. C., Lee, C. B., Chai, C. S., Koh, J. H. L., Chen, W. L., & Cheah, H. M. (2010). The self-directed learning with technology scale (SDLTS) for young students: An initial development and validation. *Computers & Education*, 55(4), 1764–1771. <https://doi.org/10.1016/j.compedu.2010.08.001>
- Turan, M. B., & Koç, K. (2018). The impact of self-directed learning readiness on critical thinking and self-efficacy among the students of the school of physical education and sports. *International Journal of Higher Education*, 7(6), 98–105. <https://doi.org/10.5430/ijhe.v7n6p98>
- Vaičiūnienė, A., & Kazlauskienė, A. (2023). Liberating and oppressive factors for self-directed learning: A systematic literature review. *Education Sciences*, 13(10), Article 1020. <https://doi.org/10.3390/educsci13101020>
- Van Damme, D., & Zahner, D. (2022). *Does higher education teach students to think critically?* OECD Publishing. <https://doi.org/10.1787/cc9fa6aa-en>
- Wilkes, S., Maggio, L. A., Martin, P. C., Melton, J., & Zheng, B. (2026). Self-directed learning in health professions education: A systematic review and meta-analysis. *Perspectives on Medical Education*, 15(1). <https://doi.org/10.5334/pme.2128>
- Witarsa, & Muhammad, S. (2023). Critical thinking as a necessity for social science students' capacity development: How it can be strengthened through project-based learning at university. *Frontiers in Education*, 7, Article 983292. <https://doi.org/10.3389/feduc.2022.983292>
- Xu, Z., Zhao, Y., Liew, J., Zhou, X., & Kogut, A. (2023). Synthesizing research evidence on self-regulated learning and academic achievement in online and blended learning environments: A scoping review. *Educational Research Review*, 39, Article 100510. <https://doi.org/10.1016/j.edurev.2023.100510>
- Yu, L., & Zin, Z. M. (2023). The critical thinking-oriented adaptations of problem-based learning models: A systematic review. *Frontiers in Education*, 8, Article 1139987. <https://doi.org/10.3389/feduc.2023.1139987>
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2