JOVACET
Journal of Vocational, Adult and Continuing Education and Training

To cite: Zwane, D. P., Ramorola, M. Z., & Sibanda, S. (2026). Visual literacy through multimodal learning: Implications for EFAL teaching in the South African TVET sector. *Journal of Vocational, Adult and Continuing Education and Training*, *9*(1), 83–102.
http://doi.org/10.14426/jovacet.v9i1.637

Visual literacy through multimodal learning: Implications for English First Additional Language teaching in the South African TVET sector

DUDUZILE PENELOPE ZWANE (ezwanedp@unisa.ac.za) Research Unit: Self-Directed Learning, Faculty of Education, North-West University, Potchefstroom, South Africa
ORCID link: https://orcid.org/0000-0002-3068-9799

MMANKOKO ZIPHORAH RAMOROLA (Ziphorah.Ramorola@nwu.ac.za) School of Mathematics, Science and Technology, Faculty of Education, North-West University, Mahikeng, South Africa
ORCID link: https://orcid.org/0000-0003-4746-7049

SIFISO SIBANDA (Sifiso.Sibanda@nwu.ac.za) School of Languages Education, Faculty of Education, North-West University, Mahikeng, South Africa
ORCID link: https://orcid.org/0000-0002-3924-5414

ABSTRACT

This study responds to the growing demand for 21st-century skills in technical and vocational education and training (TVET), where traditional literacy is insufficient for engaging with multimodal communication environments. It foregrounds the importance of visual literacy (VL) as essential competence for students entering dynamic and complex workplaces. Despite this need, traditional textbook-based pedagogies remain prevalent in TVET colleges, limiting students' opportunities to develop critical VL analysis skills. The study explores how multimodal VL instruction can provide opportunities for learning among National Certificate (Vocational) NC(V) students in English First Additional Language (EFAL) classrooms. A qualitative and design-based research (DBR) approach was employed to mine data through focus groups, interviews, lesson observations and document analysis. Purposive sampling assisted in identifying 63 participants comprising NC(V) Level 2 EFAL students and their lecturers. The data from all the instruments were thematically analysed. Findings indicate both challenges and opportunities in developing VL in the EFAL curriculum at South African TVET colleges. The study concludes that students' difficulties in interpreting body language, facial expressions, cartoon frames, speech bubbles and typographic cues reveal that VL cannot be assumed as an innate skill but is socially and pedagogically mediated.

Keywords: Multimodal learning, visual literacy (VL), English First Additional Language (EFAL), technical and vocational education and training (TVET), National Certificate (Vocational)

This work is licensed under the Creative Commons Attribution Share-alike 4.0 International Licence.

Introduction

Students' capacity to draw meanings from visual texts has become a critical skill (Romero & Bobkina, 2021), emphasising the importance of visual literacy (VL) education in developing students' ability to engage with visual materials for meaning. The concept of VL is not new, having been explored in literature as early as the 19th century (see, for example, Debes, 1969), and it refers to the ability to discriminate and interpret the visual actions, objects and symbols, whether natural or man-made, that individuals encounter in their environment (Debes, 1969, 185). In contemporary education, VL has gained considerable prominence in both teaching and learning contexts (Lundy & Stephens, 2015; Martínez-Cardama & Caridad-Sebastián, 2019; Matusiak et al., 2019; Romero & Bobkina, 2021). There is an emerging pedagogical shift in how VL texts are approached, with students expected to become viewers capable of interacting with both written and visual texts to produce multimodal meanings (Romero & Bobkina, 2021). Reflecting on this evolving perspective, Pena and Dobson (2021) regard VL as both a teachable skill and an innate semiotic activity. Similarly, the board of directors of the Association of College and Research Libraries (ACRL) (2011) highlights that scholarly engagement with images requires interpretation and analytical skills that should be integrated into the curriculum.

Proficiency in VL is widely recognised as a fundamental skill that empowers individuals and contributes to South Africa's broader educational and socio-economic development. However, many students enrolled in South African technical and vocational education and training (TVET) colleges lack the requisite VL skills needed to effectively interpret visual texts prescribed in the English First Additional Language (EFAL) curriculum for the National Certificate (Vocational) NC(V) programme. In the South African education system, English is studied as an additional language by the majority of students enrolled in TVET colleges, whose home language is not English. The intention is to develop proficiency in English as an additional language rather than as a first language. Thus, the NC(V) programme offered at TVET colleges with an EFAL curriculum comprises four language modalities: listening and speaking, reading and viewing, writing and presenting, and language and communication (DHET, 2014, p. 2). Among these, the reading and viewing modality, in which VL is taught, carries the highest weighting, accounting for 40% of the curriculum and approximately 30 notional teaching hours (DHET, 2014). This substantial allocation reflects the complexity of the knowledge and skills required for students to engage effectively with written and visual texts (DHET, 2014). One of the key subject outcomes of this modality requires students to 'read/view and assess visual texts and non-verbal forms of communication and formulate appropriate responses to the intended message' (DHET, 2014, p. 10). This outcome underscores the importance of developing students' VL competencies to enable them to interpret, analyse and respond critically to a wide range of multimodal texts encountered in both academic and everyday contexts.

To support the development of these competencies, DHET (2014) recommends exposing students to a diverse range of multimodal text types. These include, but are not limited to,

(a) visual texts such as comic strips and television advertisements; (b) film elements, including camera angles, shots, movement, framing and composition; (c) visual representations such as mind maps, flow charts and diagrams; and (d) non-verbal forms of communication, including body language and gestures (DHET, 2014, p. 25). These text types are inherently multimodal because they integrate multiple semiotic modes such as audio, visual, gestural, tactile, spatial, spoken and written texts for meaning-making (Magnusson & Godhe, 2019; Niu, 2022). Consequently, effective VL instruction requires students to interpret meaning across multiple modes rather than relying solely on written language.

The teaching of VL in the English curriculum has evolved considerably over time. Mbelani (2019) traces this development, arguing that VL has become increasingly significant in contemporary educational discourse, particularly in relation to debates surrounding the enduring legacies of apartheid and educational inequality in South Africa. Although VL was incorporated into the English curriculum during the 1970s, its implementation was shaped by the apartheid education system, which maintained racially segregated educational departments and prescribed different curricula for different racial groups (Msila, 2007). Consequently, students' access to VL instruction varied substantially according to the curriculum which was prescribed for them. VL was incorporated into the English First Language (EFL) curriculum, which was primarily designed for white students, but was excluded from the English Second Language (ESL) curriculum, which was predominantly offered to black students (Mbelani, 2019). These historical inequalities continue to influence curriculum implementation of VL in post-apartheid South Africa, particularly in under-resourced TVET colleges where many lecturers may have had limited opportunities for sustained professional development in VL, potentially affecting their ability to teach multimodal texts effectively.

Although VL is recognised as a fundamental component of the NC(V) EFAL curriculum, little is known about the pedagogical approaches employed by lecturers to develop students' VL competencies or about the challenges they encounter in teaching this aspect of the curriculum at TVET colleges in South Africa. This lack of empirical evidence highlights a significant knowledge gap and provides the rationale for the present study, which seeks to explore lecturers' pedagogical practices in the teaching of VL in the NC(V) EFAL programme. Therefore, this study responds to the following questions:

- What challenges arise when VL is taught and learnt in EFAL classrooms at TVET colleges?
- What opportunities exist for teaching VL using multimodal approaches in EFAL classrooms at TVET colleges?

The ability to deduce meanings from texts is a critical skill that students in TVET colleges have to master in the 21st century. Rather than just being text readers and writers, students of the 21st century are expected to become viewer–students who are capable of interacting with both written and visual texts to develop multimodal meanings (Romero & Bobkina, 2021). Students in South Africa and many other nations are continuously exposed to images in their daily lives

but they struggle to locate, interpret, assess, use and produce images and visual media (ACRL, 2011). Since English is the medium of teaching in the majority of schools, colleges and universities in South Africa, students who lack VL skills find it difficult to learn EFAL in TVET courses. In fact, there are a lot of images in English teaching materials, and the combination of images and text is evidence of new publishing systems and developing technologies worldwide (Ataie & Mohammadi, 2021). As a result, someone who lacks visual competences may feel alienated and powerless when navigating, discussing and interpreting visual texts.

In order to explain the fundamental pedagogical approaches that offer a framework for VL in the teaching of languages like EFAL content, researchers like Debes (1969) and others studied and developed theories that explain the foundational pedagogical approaches which provide a framework for VL in the teaching of languages such as English in second-language contexts (see Avgerinou & Pettersson, 2011; Golding & Verrier, 2020; Grape, 2020; Lim et al., 2021; Sibanda & Sibanda, 2013; Teane & Gombwe, 2023). VL in vocational education has not received much attention in South Africa. Students need to be viewers and interpreters of texts in addition to readers and writers; hence an investigation of the implications of multimodal learning of VL for the subject EFAL in TVET appears to be crucial. Similarly, Johnson (1977), who explored Debe's decision to promote photographic equipment as a tool for visual communication, raised a concern about the necessity of understanding the nature of VL, noting:

> Through Debes, Eastman Kodak Company became the sponsor of early visual literacy development. Their 'educational' periodical, Visuals are a Language, began using the term in 1967 and continued using it through 1973, usually as a marketing device to encourage teachers to sponsor 'visual communication' projects in which photography could be taught as a language arts and English activity (Johnson, 1977, p. 9).

Although VL is significant to both language arts and English, its meaning and implications have been debated for years because of this expressed concern. This has resulted in an increasing lack of clarity regarding the approaches that should be outlined in an EFAL subject specifically intended for a TVET context. In order to locate VL in the context of multimodal approaches in the EFAL curriculum, this article first presents a review of the literature. It critically engages with key scholarly perspectives on VL, multimodal learning, and teaching and learning in TVET environments by synthesising and aligning these insights with the objectives of the study. That is followed by an outline of the research methodology of the enquiry, a thematic discussion of its findings and the implications for EFAL development, and, finally, concluding observations.

Visual literacy in English First Additional Language TVET classrooms

Students are exposed to visual stimuli, making the ability to derive meaning from images an essential contemporary literacy skill. According to Romero and Bobkina (2021), traditional notions of literacy are no longer sufficient, as students must function as 'viewer–readers' capable

of constructing meaning across multiple modes. This shift has significant implications for English education, requiring an expanded pedagogical focus that integrates reading and viewing through the development of VL. The present study responds to this need by foregrounding the development of VL skills among TVET NC(V) students through structured analysis of cartoons and advertisements. The conceptual foundations of VL can be traced to John Debes (1969), who defined VL as a set of visual competencies developed through sensory integration. Subsequent scholarship has expanded this view, positioning VL as a cognitive and interpretive process involving the decoding, evaluation and creation of visual meaning (Golding & Verrier, 2020; Sibanda & Sibanda, 2013). The ACRL (2011) further conceptualises VL as a set of abilities that enables individuals to find, interpret, evaluate and create visual media, emphasising its importance in academic contexts where visual materials are increasingly pervasive. This broader conceptualisation reinforces the view of VL as an essential component of contemporary literacy, particularly in multimodal learning environments.

The growing prominence of VL is closely linked to the proliferation of visual media (Bautista & Gutierrez, 2022). However, existing literature is largely grounded in first-language contexts, limiting its applicability to EFAL settings, such as South African TVET colleges (Bautista & Gutierrez, 2022; Magnusson & Godhe, 2019). Although visual education is established in early schooling, it tends to diminish in higher education (Kędra & Žakevičiūtė, 2019), where text-based approaches dominate. This creates a disconnect between students' everyday engagement with multimodal content and their ability to interpret visuals in academic contexts. Among NC(V) Level 2 students, high exposure to visual content does not necessarily translate into well-developed VL competencies, particularly in the analysis of cartoons and advertisements, thereby necessitating targeted pedagogical intervention.

Romero and Bobkina (2021) emphasise that students must navigate meaning across both textual and visual modes, yet multimodal literacies remain underemphasised in classroom practice. While VL is recognised as essential for constructing multimodal meaning, it is still insufficiently taught, thereby justifying the need for structured instructional approaches. Avgerinou and Pettersson (2011) conceptualise VL as a holistic, learnable competence encompassing visual perception, language, learning, thinking and communication. The authors argue that, similar to verbal language, visual language possesses its own grammar, vocabulary and syntax, requiring students to decode meaning across both cognitive and affective dimensions (Avgerinou & Pettersson, 2011).

In this study, VL frameworks are used to support students' interpretation of cartoons and advertisements. The focus for cartoons is on elements such as facial expressions, body language, framing, typography and symbolism, in alignment with the ACRL (2011) standards. Prompts for the advertisements were aligned to the attention, interest, desire and action (AIDA) principle. The AIDA principle is a widely recognised framework in advertising theory that aligns with hierarchical models of consumer behaviour which conceptualise progression from awareness to purchase (Baber, 2022; Chaojie, 2015). In the context of TVET colleges, it is incorporated into the NC(V) Level 2 EFAL curriculum, including prescribed textbooks such as *Hands-on training: First additional language* (Wade et al., 2004),

where it functions as both an analytical and pedagogical tool to guide students in critically interpreting and producing advertisements.

Grape (2020) highlights the increasing dominance of visual media alongside the continued marginalisation of VL in formal education, despite its central role in cognition and interpretation. Visual processing plays a critical role in meaning-making, memory and interaction with the environment (Raiyn, 2016). These insights underscore the need for explicit VL instruction. Accordingly, our study foregrounds structured VL analysis frameworks to support students' engagement with cartoons and advertisements as complex multimodal texts.

This body of literature highlights both the importance of VL and the gaps in its implementation in EFAL classrooms, particularly in TVET contexts. While scholars emphasise the need for students to engage with multimodal texts, there remains limited evidence on how VL is effectively taught and what challenges arise in practice. These gaps directly inform the present study, which investigates the challenges associated with teaching and learning VL in EFAL classrooms, as well as the potential of multimodal approaches to enhance VL development.

Multimodal learning in TVET classrooms

The 21st century has witnessed a shift from monomodal to multimodal forms of communication, necessitating broader conceptions of literacy. Contemporary scholarship recognises literacy as extending beyond reading and writing to include the ability to interpret and produce meaning across multiple modes, including visual, auditory and digital forms (Lim et al., 2021; Sibanda & Sibanda, 2013). In this context, multimodal literacy has become essential for students navigating increasingly complex communication environments. However, VL remains underdeveloped in the South African TVET sector, particularly in NC(V) EFAL classrooms. 'Multimodality' refers to the use of multiple communicative modes such as images, gestures, speech and digital media, often engaging several sensory modalities simultaneously (Niu, 2022; Pan & Zhang, 2020). Initiatives such as the University of the Witwatersrand's multimodality project (Archer & Newfield, 2014) demonstrate the relevance of multimodal approaches in linguistically and culturally diverse South African classrooms, promoting pedagogies that accommodate multilingual realities.

As many students do not speak English as their first language, its predominance as the language of teaching and learning presents additional challenges. By incorporating visual and experiential elements into instruction, multimodal approaches offer alternative pathways to meaning-making and can enhance access and inclusivity. Such pedagogies have been shown to increase student engagement and improve reading competency while encouraging active participation (Pan & Zhang, 2020). Despite these benefits, significant challenges remain. Effective implementation is often hindered by teacher resistance, limited pedagogical knowledge and resource constraints (Niu, 2022). Concerns have also been raised regarding disparities in students' cultural capital, which may affect their ability to interpret multimodal texts (Stein & Newfield, 2014). Moreover, tensions persist between innovative instructional approaches and

traditional assessment practices that privilege written modes (Magnusson & Godhe, 2019). In response to these complexities, this study adopts a qualitative and design-based research (DBR) approach to explore how multimodal VL instruction can be effectively implemented in resource-constrained TVET contexts while addressing persistent pedagogical gaps.

Theories underpinning multimodality

Lim (2024) proposes the concept of intersemiosis, referring to meaning that emerges through the integration of multiple semiotic resources. While acknowledging the richness of combining modes, Lim (2024) also emphasises the importance of examining how these resources interact to produce meaning. The present study explores intersemiosis in NC(V) EFAL classrooms, where the integration of diverse semiotic resources for teaching VL remains underexplored in TVET colleges. Magnusson and Godhe (2019) argue that the analysis of multimodal texts should also consider mediation and affordances. 'Mediation' refers to the meaning potential of the different modes, while 'affordances' relates to what can be achieved through a particular mode based on its material and semiotic properties. Semiotic affordances are shaped by historical and social contexts, influencing how meaning is constructed. In English language teaching, teachers can still utilise images, audio clips and graphical representations to engage multiple communicative channels and stimulate students' senses, thereby enhancing engagement and participation (Pan & Zhang, 2020).

Expanding on this perspective, Stein and Newfield (2014) identify multimodal social semiotics as the dominant theoretical framework underpinning multimodal education. They define it as the study of how individuals use various modes, such as speech, writing, images, gestures and sound, as resources for meaning-making. Teaching and learning are inherently multimodal, as meaning is constructed through the interaction of these modes. However, this highlights a gap in TVET teaching practice, where it remains unclear how TVET lecturers fully utilise these semiotic resources to support learning through VL instruction. By examining how multimodal resources are integrated into VL lessons, this study investigates how different modes create layered and meaningful learning experiences, thereby supporting a deeper appreciation of VL.

Teaching and learning in the TVET environment

The shift from lecturer-centred to student-centred pedagogies in the TVET sector is increasingly recognised as a global imperative. International frameworks emphasise autonomous learning through innovative pedagogical models. For instance, Malaysia's competency-based Model of Pedagogical Reasoning and Action and Singapore's 'hands-on, minds-on, hearts-on' approach both foreground student-centred, task-oriented learning designed to meet 21st-century demands (Mohamed & Mohamoud, 2019; Tan & Seet, 2020). Similarly, in China, digital pedagogies such as flipped classrooms and game-based learning reposition lecturers as facilitators of learning rather than mere knowledge transmitters (Chen & Chan, 2024). Despite these advancements, challenges persist due to insufficient lecturer training and low student motivation.

In South Africa, policy interventions by the Department of Higher Education and Training (DHET) aim to enhance teaching quality in TVET colleges; however, entrenched lecturer-centred approaches, resource constraints and exclusion of lecturers from policy development limit effective implementation (Matabane et al., 2022). Additional challenges include diverse student academic backgrounds in NC(V) programmes and the use of English as the medium of instruction, which can exacerbate comprehension barriers and reduce classroom participation (Teane & Gombwe, 2023). Student-centred, multimodal pedagogies, grounded in social constructivist and heutagogical principles, offer potential to enhance engagement and accommodate diverse learning needs, thus aligning TVET education with contemporary educational and workplace demands.

The literature reviewed across VL, multimodal learning and TVET teaching practices highlights both opportunities and constraints in EFAL classrooms. While multimodal approaches offer inclusive and engaging pathways for meaning-making, persistent challenges such as lecturer-centred practices, limited resources and language barriers pose obstacles to effective learning. These insights directly inform the first research question, which investigates the challenges that arise when VL is taught and learnt in EFAL classrooms. Simultaneously, the potential of student-centred, multimodal pedagogies to enhance engagement and comprehension addresses the second research question, which examines the possibilities for teaching VL through multimodal approaches in EFAL classrooms. By exploring both the barriers and opportunities in the TVET context, this study seeks to provide practical strategies for improving VL teaching and learning.

Methodology

The present study was situated in an interpretivist world view and employed an inductive approach integrated with DBR in order to explore the challenges and opportunities of teaching and learning VL through the EFAL curriculum in South African TVET colleges. DBR, characterised by its iterative and cyclical nature (Tinoca et al., 2022), enabled collaboration between lecturers and students in the co-design and development of multimodal strategies to enhance VL among EFAL students. The DBR process was implemented in two cycles. In the first cycle, students completed cartoon analysis frameworks, interpreting visual elements such as facial expressions to infer characters' emotions. These artefacts served as data to evaluate the extent to which students could transfer VL skills beyond guided instruction. In the second cycle, students applied similar analytical frameworks to advertisements they selected. Prompts guided students to interpret visual features such as font, layout and symbolic elements, drawing on the AIDA principles (Chaojie, 2015).

An inductive approach to data collection and analysis was adopted to identify methods of data collection and analysis (Thomas, 2003). Qualitative data were collected through semi-structured interviews with three lecturers, three focus-group discussions comprising six participants each and lesson observations (Baker, 2006; Yulianti & Sulistyawati, 2021). All interviews and the focus-group discussions were conducted face to face on campus using a self-developed interview schedule (Adams, 2015). The sessions were audio-recorded and

transcribed verbatim for analysis. A purposive sampling strategy was used to select NC(V) Level 2 EFAL lecturers and students from one TVET college in South Africa (Campbell et al., 2018; Onwuegbuzie & Collin, 2007). Ethical clearance was obtained from the relevant institutional ethics committee as well as institutional gatekeepers. All participants provided informed consent and were assured of anonymity and confidentiality.

Data were analysed using thematic analysis following the six-phase framework proposed by Braun and Clarke (2006). Codes and categories were systematically developed to organise the data into meaningful patterns. The analysis was supported by ATLAS.ti software, which facilitated the management and organisation of large qualitative datasets (Soratto et al., 2020). Feedback on the analysis is reported in the findings below, in response to the research questions.

Findings

What challenges arise when VL is taught and learnt in EFAL classrooms?

Key themes that emerged from the qualitative data, as expanded on below, were related to the challenge of interpreting verbal and non-verbal cues, to technical understanding of cartoon analysis and to the lack of pedagogical practice. Students expressed significant difficulty in how to interpret 'body language' (1:10 ¶ 92; 2:19 ¶ 68; 2:59 ¶ 93) and 'facial expression' (2:7 ¶ 11; 3:24 ¶ 95; 4:9 ¶ 11; 5:29 ¶ 46) and in acquiring knowledge on 'how to answer questions about verbal and body language' (2:16 ¶ 236) as key semiotic resources in the analysis of cartoon and comic strips. Students acknowledged their challenges of interpreting cartoons and remarked, 'Sometimes you can't interpret the cartoons; it's hard' (2:19 ¶ 68), and 'We tend not to see maybe the body languages or interpret some of the cartoons' (3:24 ¶ 95). Lecturers shared this sentiment: 'They [the students] do not know how to work with cartoons, you know like facial expressions, body language, those things they don't understand most of them' (5:29 ¶ 46; 4:9 ¶ 11). Students' difficulties in interpreting cartoons corroborate Romero and Bobkina's (2021) emphasis of the critical importance of VL education.

The incorporation of VL education into the EFAL curriculum (ACRL, 2011; Mbelani, 2019) would better equip students with the skills required to engage critically with visual materials and construct meaning (Romero & Bobkina, 2021). In this study, students were often unable to determine whether a character was expressing happiness or sadness or to infer the intended message conveyed by the cartoonist through visual cues. One student shared the following:

> Facial expression on how to identify emotions of cartoons and how can we tell when a cartoon character is not happy, how to spot body language when the character is not interested in what is being said; also their interest or posture towards the conversation (3:13 ¶ 13).

From the observations, many students incorrectly responded to the question on how the cartoonist used the speech bubbles to indicate if the characters were shouting, thinking or

speaking. Students lacked the technical understanding of speech bubbles and their function in different contexts of the cartoon or comic strip. For instance, some students were unable to identify the number of frames in a comic strip or cartoon, while others struggled to interpret the meaning of speech bubbles when drawn in varied stylistic forms. Similarly, a smaller number of students could identify how the movement was depicted in a cartoon or comic strips they were analysing. The use of font to emphasise intensity or importance in characters' speech was also problematic. While some students recognised surface-level features such as bold or italicised font, they were unable to articulate the significance of the typographic choices in conveying tone, emphasis or characters intent. This finding resonates with scholarship that positions images as central to meaning-making (Avgerinou & Pettersson, 2011; Grabe, 2020). These authors argue that images function holistically, simultaneously conveying meaning and evoking emotional responses to interpret their environment. Instead, the findings suggest that VL operates as a socially and pedagogically mediated competence, rather than as an inherent skill. Students' interpretations were not only shaped by individual perception, but were also constrained by contextual and material factors, including the quality of the visual artefacts. This highlights how the design and selection of visual materials can either enable or hinder meaning-making processes.

This limited engagement must also be understood in relation to pedagogical practice. The lack of a structured approach to teaching cartoons was identified by students as a significant challenge, pointing to inconsistencies in how visual texts are mediated in the classroom. This suggests that the difficulty is not located solely within the students' interpretive abilities but is also shaped by how VL is taught. This aligns with literature (Mbelani, 2019; Niu, 2022) that points to pedagogical shortcomings in VL instruction. The absence of formal training in VL constrains some lecturers' capacity to teach VL content effectively, thereby limiting students' opportunities to develop critical interpretive skills. Consequently, the teaching of visual texts risks remaining fragmented and superficial, rather than being approached as a systematic and theoretically informed component of EFAL pedagogy in the TVET context.

However, Peña and Dobson (2021:1) posit that 'the first perception of VL is a set of skills and competencies which can be taught formally in the classroom', thereby foregrounding the role of explicit instruction in developing students' visual interpretive abilities. This perspective is reinforced by the findings of the present study, which reveal that students' challenges extend beyond emotional interpretation to include structural and semiotic features of visual texts. Participants explicitly attributed their misinterpretations to the use of unclear, black-and-white cartoons (1:32 ¶ 86), noting that such representations made it difficult to discern the emotions and the expressions. They believed that colourful cartoons might provide them with an idea of 'how a particular cartoon looks or how it reacts' (3:29 ¶ 114; 2:7 ¶ 11). Students further expressed the importance of colourful cartoons:

> Be able to tell if a person is laughing out of happiness or sadness, understand what a cartoonist tries to say using body language (1:10 ¶ 92; 3:34 ¶ 12).

> To know how to answer questions of verbal, body language and visual cues (2:59 ¶ 93; 2:16 ¶ 236).
>
> Facial expression on how to identify emotions of cartoons and how can we tell when a cartoon character is not happy, how to spot body language when the character is not interested in what is being said; also their interest or posture towards the conversation (3:13 ¶ 13).

Students' interpretations were not only shaped by individual perception, but were also constrained by contextual and material factors, including the quality of the visual artefacts. This highlights how the design and selection of visual materials can either enable or hinder meaning-making processes (Magnusson & Godhe, 2019; Niu, 2022).

What are the opportunities of implementing multimodal approaches for teaching Visual Literacy (VL) in English First Additional Language (EFAL)?

The findings were derived from the second cycle of the DBR approach aimed at evaluating the designed intervention. The VL analysis frameworks were perceived to have made a positive impact on the learning of VL by both lecturers and students. The incorporation of multimodal approaches enhanced students' comprehension of VL texts, as well as active participation (Pan & Zhang, 2020). As evidenced by the classroom observations, students engaged with audiovisual texts (2:17 10m 34s; 3:3 6m 6s; 1:10 48m 35s) and developed skills to interpret VL texts for meaning and identify gaps for further support (Lim, 2019; Magnusson & Godhe, 2019). Both students (1:32 ¶ 85; 2:26 ¶ 74) and lecturers (2:18 ¶ 44; 3:16 ¶ 48) noted improved VL examination performance post-intervention from the students who participated in the intervention. The implementation of VL analysis frameworks addressed the VL pedagogy gaps (Bautista & Gutierrez, 2022; Yeh & Lohr, 2010) by providing students with an essential structure to analyse VL text that would otherwise have been difficult to navigate without guidance (2:15 ¶ 50; 3:14 ¶ 41). Students shared that the frameworks provided them with structure, as they did not know where to begin analysing the VL texts without prompts and that the frames made it easier for them to interpret cartoon elements. The following extracts highlight the students' positive reflections:

> I wouldn't know where to start, the analysing … . (2:15 ¶ 50).
>
> Frames aided in advertisement analysis which provided a structured framework to evaluate element (1:29 ¶ 22).
>
> Frames provided a structured lens (1:31 ¶ 77),
>
> Help me to create my advertisement (3:60 ¶ 110).

Other students commented that the frameworks 'broadened the knowledge about advertisements' (1:30 ¶ 44), provided them with 'solid understanding of display, structure and principle of design' (1:5 ¶ 68) and also assisted 'by enlightening me with information' (2:20 ¶ 33).

Additionally, both lecturers and students indicated that the frameworks were beneficial in scaffolding VL analysis. They both confirmed that students performed better in the internal examination written immediately after the implementation of multimodal approaches. One lecturer attested, 'The group that we were doing the intervention with, they performed much better than all the other groups' (2:18 ¶ 44). Two other lecturers shared the sentiment, 'Mostly the "A" group that participated, this time their marks were better than the "B" group so maybe it influenced that' (3:16 ¶ 48), and 'it helped us during our exams as well…' (1:32 ¶ 85). This confirmation was further emphasised by the illustrations in Figures 1 to 3 as guided by the AIDA principle described in the following section. Permission to use the students' activities was granted.

The AIDA principle

The AIDA principle can be explained as follows:

Attention

Figure 1
Student-designed advertisement to address the attention principle

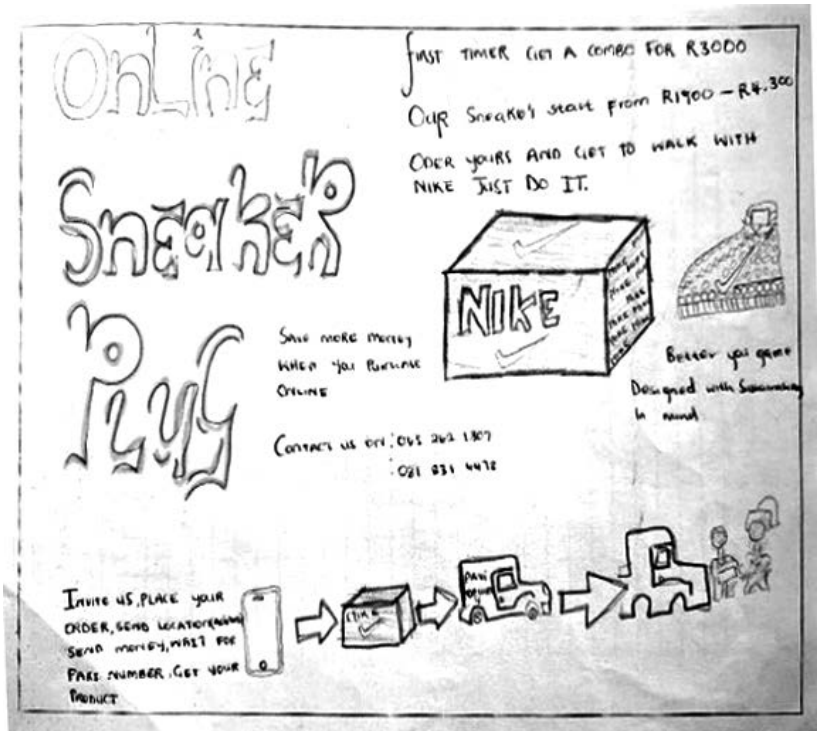


Figure 2
Student-designed advertisement to address the attention principle

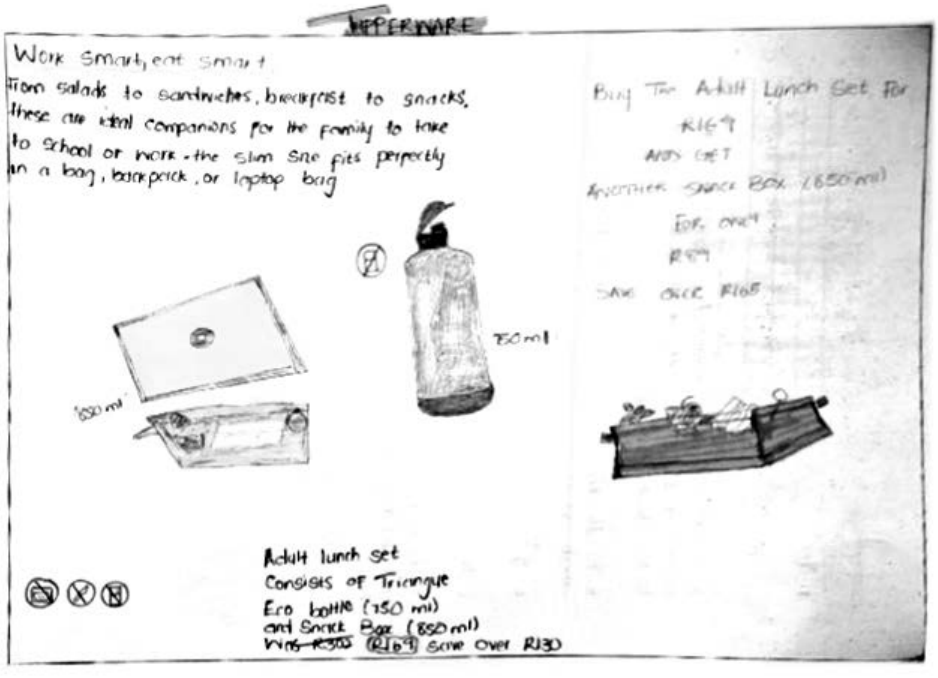


Figures 1 and 2 show how the student implemented the attention principle by employing colour and diverse font styles. In Figure 2, the student addressed the principle of a call to action through offering discounted prices on specific purchases and providing instructions on how to access the advertised goods. The student implemented the attention principle by employing colour and diverse font styles. The inclusion of images is expected to generate interest. The findings reveal that most students could respond correctly, guided by the prompt. For instance, students were prompted to identify how the font was used as a visual

aid to reinforce the message of the advertisement. Some students, although not discussing it in detail, identified that font in the analysed advertisements was used to 'attract customers' (3:17 ¶ 233), and 'used base font to make it stand out' (1:17 ¶ 194). Keeping the focus of visual reinforcement, students were prompted to comment on how the advertisements attracted the attention of the potential consumer. The same prompt was used to guide students towards presenting the same comment on the audiovisual advertisements that were displayed. Similar to the previous prompt, most students could identify the visual elements that attracted the attention of the reader or viewer. They successfully addressed the prompt on visual elements used to attract attention, as they explained: 'used visual elements like bright colours and bold' (1:19 ¶ 405), 'visual elements like font and pictures' (2:20 ¶ 368) and 'showing the car on the ad' (3:20 ¶ 629). This suggests that the use of the advertisement analysis prompts assists students to respond correctly to how the advertisement attracts the attention of the potential consumer for both the print and audiovisual advertisements.

Interest

Figure 3
Student-designed advertisement to address the interest principle



In Figure 3, the student addressed the principle of a call to action through offering discounted prices on specific purchases and providing instructions on how to access the advertised goods. The student employed colour and diverse font styles in the images to generate interest. Students

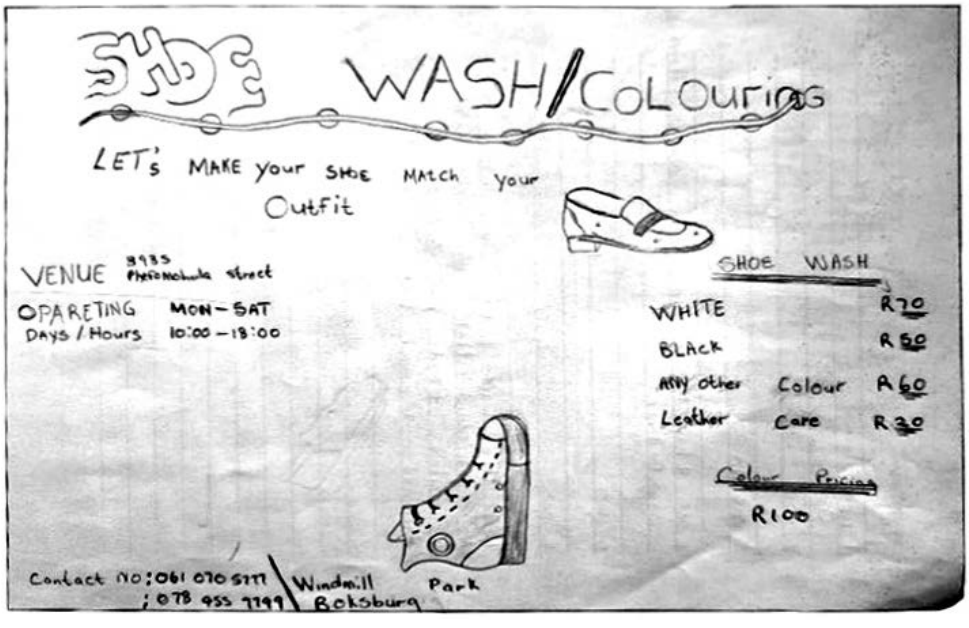
were prompted to discuss how the advertisement managed to sustain the attention of potential consumers. Some students were assisted by the prompt and provided correct descriptions of how the attention was sustained in both print and audiovisual advertisements. There were responses such as: 'attention sustained by mentioning the benefits of the product' (1:20 ¶ 452), 'through latest music and bright colours' (3:23 ¶ 702) and 'by using people that people like to see on TV' (2:21 ¶ 408). What has also possibly maintained the interest of the reader or viewer is whether the advertisement's content is relatable. Students were prompted to comment on the relatability of the advertisements in terms of facial expressions, body language, framing, typography and symbolism in line with the AIDA principle. Some students were able to indicate how the advertisements were relatable by stating, 'the make-up is expensive' (1:14 ¶ 580), while others could not indicate reasons for their interpretation (3:14 ¶ 880; 1:14; 2:15; 3:14).

Desire

Students had to use prompts to interpret how advertisements use various elements to evoke a sense of aspiration in potential consumers. Some students could link the images used to create desire by the advertisement: 'the car parked outside and ready to go and it is written the best or nothing' (1:23 ¶ 234), and 'the lady in the advert is smiling, showing that she is enjoying the aroma' (3:27 ¶ 360).

Action

Figure 4
Student-designed advertisement to address the action principle



The action principle addresses elements in the advertisement that encourage the reader or viewer to purchase the product or service. Figure 4 is relatable to students and the main message is clear. Relatability addresses the principle of sustaining the attention of the potential consumer. While students did not make use of colour, they were encouraged to draw on visual and textual resources that were readily available within their contexts, provided that these supported the application of the AIDA principle. The varied fonts and images attract the attention of consumers. Students also implemented the principle of a call to action by providing details of how to access the product and the event and by also providing prices for the service and the event being advertised. Students were able to identify how the advertisers encouraged a call to action such as through, 'get a big mac when you visit and register' (1:26 ¶ 322), 'call them to schedule a tour' (3:30 ¶ 490) and 'dial the USSD code to win your share of 100 000' (2:10 ¶ 322).

There were palpable improvements in the teaching and learning approaches by both lecturers and students upon the inclusion of multimodal texts. Lecturers explained that 'participation of the students were better in the [uhm] multimodal session' (1:6 ¶ 7). One of the lecturers expressed this thought: 'If I had to rate, I think it was good, because they were engaging, the engagement was there'. (2:2 ¶ 12) Lecturers explored more collaborative approaches and students' participation levels improved. Students' improved participation levels could be attributed to students enjoying the multimodal teaching approach (Niu, 2022; Pan & Zhang, 2020). Additionally, the inclusion of multimodal modes allowed students to express their positive feedback: 'I learned through discussions' (1:19 ¶ 49) 'with friends' (2:27 ¶ 31). Students were actively engaged and exercised critical thinking in the discussion of keywords (3:6 3m 54s, 9:3, 1).

Conclusion

The aim of the study was to explore how VL instruction can provide opportunities for learning among NC(V) Level 2 students in EFAL classrooms at South African TVET colleges. Drawing on the qualitative research approach supported by DBR, the study produced VL analysis frameworks such as the comic strips and cartoons analysis frame and the display advertisement analysis frame to guide students in the analysis and interpretation of visual texts. Within the context of TVET, the findings reveal a clear disconnect between students' everyday exposure to visual media and their ability to critically interpret and analyse visual texts in the EFAL curriculum. This relates to interpreting body language, facial expressions, cartoon frames, speech bubbles and typographic cues. These challenges are compounded by a lack of structured teaching approaches and limited formal training in VL, which constrain lecturers' capacity to scaffold students' interpretive abilities effectively.

In conclusion, this study has demonstrated that VL is no longer a peripheral component of language education but a core competency for students navigating multimodal communication environments in the 21st century. This gap is compounded by persistent lecturer-centred

pedagogies, limited pedagogical training in VL and resource constraints, all of which restrict meaningful engagement with multimodal texts.

Importantly, the study provides empirical evidence that multimodal approaches to VL instruction can significantly enhance students' comprehension and participation. The use of structured VL analysis frameworks enabled students to engage more confidently with complex visual texts, supporting both their interpretive skills and their ability to produce meaning across modes. These approaches not only scaffolded learning, but also fostered collaboration, critical thinking and increased motivation.

However, the effectiveness of these interventions is contingent upon systemic transformation. The findings underscore the need for sustained professional development for lecturers, deliberate integration of VL into the EFAL curriculum and the provision of appropriate multimodal resources. Without such support, VL instruction risks remaining fragmented and inconsistently implemented.

Ultimately, the study affirms that lecturer practices are central to cultivating students' multimodal meaning-making abilities. By embracing student-centred, multimodal pedagogies, TVET colleges can better prepare students for lifelong learning and the evolving demands of the contemporary workplace. Future research should extend this work across diverse TVET contexts to further refine pedagogical models and explore scalable strategies for embedding VL in vocational education.

Acknowledgements

We gratefully acknowledge the TVET college that participated in this study, as well as the NC(V) Level 2 EFAL lecturers and students for their valuable participation and contribution.

REFERENCES

Adams, W. C. (2015). Conducting semi-structured interviews. In Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of practical program evaluation* (Vol. 19). John Wiley & Sons, pp. 492–505.

Archer, A., & Newfield, D. (2014). *Multimodal approaches to research and pedagogy*. Routledge, pp. 1–18.

Association of College and Research Libraries (ACRL). (2011). *Visual literacy competency standards for higher education*. http://www.ala.org/acrl/standards/visualliteracy

Ataie, A., & Mohammadi, M. (2021). Evaluation of visual literacy of high school English textbooks in Iran: An intercultural perspective. *International Journal of Foreign Language Teaching and Research*, *10*(40), pp. 41–52.

Avgerinou, M. D., & Pettersson, R. (2011). Toward a cohesive theory of visual literacy. *Journal of Visual Literacy, 30*(2), 1–19. https://doi.org/10.1080/23796529.2011.11674687

Baber, H. (2022). Application of the AIDA model of advertising in crowdfunding. *International Journal of Technoentrepreneurship*, *4*(3), pp. 167–179.

Baker, L. (2006). Observation: A complex research method. *Library Trends*, *55*(1), pp. 171–189.

Bautista, J. C., & Gutierrez, M. R. M. (2022). A grounded theory on the comprehension processing of teachers as ESL readers of multimodal still visuals. *RELC Journal*, *53*(3), pp. 582–596.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), pp. 77–101.

Campbell, R., Goodman-Williams, R., Feeney, H., & Fehler-Cabral, G. (2018). Assessing triangulation across methodologies, methods, and stakeholder groups: The joys, woes, and politics of interpreting convergent and divergent data. *American Journal of Evaluation*. https://doi.org/10.1177/1098214018804195

Chaojie, L. (2015). The application of 'Aida' model to economic classes. Вестник Южно-Уральского государственного университета. Серия: Образование. Педагогические науки, *7*(4), pp. 77–83.

Chen, Z., & Chan, J. (2024). Large language model in creative work: The role of collaboration modality and user expertise. *Management Science*, *70*(12), pp. 9101–9117.

Debes, J. L. (1969). The loom of visual literacy. *Audiovisual Instruction*, *14*(8), pp. 25–27.

Department of Higher Education and Training (DHET). (2014). *English First Additional Language (Level 3) final subject guidelines*. Government Printers.

Golding, S., & Verrier, S. (2020). Teaching people to read comics: The impact of a visual literacy intervention on comprehension of educational comics. *Journal of Graphic Novels and Comics*. https://doi.org/10.1080/21504857.2020.1786419

Grabe, M. E. (2020). Visual cognition. In *Handbook of visual communication.* Routledge, pp. 51–70.

Johnson, B. D. (1977). *Visual literacy, media literacy and mass communications for English instruction*. Northwestern University.

Kędra, J., & Žakevičiūtė, R. (2019). Visual literacy practices in higher education: What, why and how? *Journal of Visual Literacy*, *38*(1–2), pp. 1–7.

Lim, F. V. (2019). Investigating intersemiosis: A systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse. *Discourse Studies*, *20*(1), pp. 34–58.

Lim, F. V. (2024). The multimodal turn in higher education. In *Current trends in EMI and multimodality in higher education*. Routledge, pp. 9–25.

Lim, F. V., Towndrow, P. A., & Tan, J. M. (2021). Unpacking teachers' multimodal pedagogies in the primary English language classroom in Singapore. *RELC Journal*, pp. 1–15. https://doi.org/10.1177/00336882211011783

Lundy, A. D., & Stephens, A. E. (2015). Beyond the literal: Teaching visual literacy in the 21st century classroom. *Procedia – Social and Behavioral Sciences*, *174*, pp. 1057–1060.

Magnusson, P., & Godhe, A.-L. (2019). Multimodality in language education: Implications for teaching. *Designs for Learning*, *11*(1), pp. 127–137. https://doi.org/10.16993/dfl.127

Martínez-Cardama, S., & Caridad-Sebastián, M. (2019). Social media and new visual literacies: Proposal based on an innovative teaching project. *Education for Information*, *35*, pp. 337–352. https://doi.org/10.3233/EFI-180214

Matabane, M. E., Matabane, R. B., Moloi, T. J., & Sibaya, K. T. (2022). Exploring factors that teachers view as hindering quality in teaching and learning at a TVET college. *Research in Educational Policy and Management*, *4*(1), pp. 51–64.

Matusiak, K. K., Heinbach, C., Harper, A., & Bovee, M. (2019). Visual literacy in practice: Use of images in students' academic work. *College & Research Libraries*, *80*(1), pp. 123–137.

Mbelani, M. (2019). Exploring the impact of discourses in advancing visual literacy: A case of interpreting advertisements. *English Academy Review*, *36*(1), 46–57. https://doi.org/10.1080/10131752.2019.1589143

Mohamed, A. T. A., & Mohamoud, A. A. (2019). The role of visual literacy in teaching English vocabulary: A case study in Khartoum State secondary schools' English language teachers. *SUST Journal of Educational Sciences*, *20*(1), pp. 179–188.

Msila, V. (2007). From apartheid education to the Revised National Curriculum Statement: Pedagogy for identity formation and nation building in South Africa. *Nordic Journal of African Studies*, *16*(2). https://doi.org/10.53228/njas.v16i2.63

Niu, L. (2022). Analysis of multimodal teaching of college English under the background of artificial intelligence. *Security and Communication Networks*. https://doi.org/10.1155/2022/3833106

Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, *12*(2), pp. 281–416.

Pan, X., & Zhang, Z. (2020). An empirical study of the application of a multimodal approach to teaching reading in EFL in senior high school. *International Journal of Emerging Technologies in Learning*, *15*(2), pp. 98–111.

Peña, E., & Dobson, T. M. (2021). The lost years of visual literacy. *Journal of Visual Literacy*, *40*(1), pp. 1–14.

Raiyn, J. (2016). The role of visual learning in improving students' high-order thinking skills. *Journal of Education and Practice*, *7*(24), pp. 115–121.

Romero, E. D., & Bobkina, J. (2021). Exploring critical and visual literacy needs in digital

learning environments: The use of memes in the EFL/ESL university classroom. *Thinking Skills and Creativity*, *40*. https://doi.org/10.1016/j.tsc.2020.100783

Sibanda, J., & Sibanda, L. (2013). Visual literacy development through the mediation of Grade 4 English textbooks. *Journal of Visual Literacy*, *32*(2), pp. 39–66.

Soratto, J., Pires, D. E. P. D., & Friese, S. (2020). Thematic content analysis using ATLAS.ti software: Potentialities for research in health. *Revista Brasileira de Enfermagem*, *73*.

Stein, P., & Newfield, D. (2014). Multiliteracies and multimodality in English education in Africa: Mapping the terrain. *English Studies in Africa*, *49*(1), pp. 1–21.

Tan, S. H., & Seet, I. (2020). Infrastructure and pedagogy innovation – a differentiating factor in TVET. In *Anticipating and preparing for emerging skills and jobs: Key issues, concerns, and prospects*. Singapore: Springer, pp. 109–115.

Teane, F. M., & Gombwe, R. (2023). Self-efficacy as a tool to enhance students' classroom participation: TVET/FET college students' views on the causes of student passivity. *Community College Journal of Research and Practice*, *47*(11), pp. 696–708.

Thomas, D. R. (2003). *A general inductive approach for qualitative data analysis*.

Tinoca, L., Piedade, J., Santos, S., Pedro, A., & Gomes, S. (2022). Design-based research in the educational field: A systematic literature review. *Education Sciences*, *12*(6), p. 410.

Wade, F., Steenkamp, L., & Muirhead, E. (2004). *Hands-on training: First additional language NC(V) Level 2 student book* (3rd ed.). Pearson Education South Africa.

Yeh, H. T., & Cheng, Y. C. (2010). The influence of the instruction of visual design principles on improving pre-service teachers' visual literacy. *Computers & Education*, *54*(1), pp. 244–252.

Yeh, H., & Lohr, L. (2010). Towards evidence of visual literacy: Assessing pre-service teachers' perceptions of instructional visuals. *Journal of Visual Literacy*, *29*(2), pp. 183–197.

Yulianti, T., & Sulistyawati, A. (2021). Enhancing public speaking ability through focus group discussion. *JURNAL PAJAR (Pendidikan Dan Pengajaran)*, *5*(2), pp. 287–295.