

VIETNAM NATIONAL UNIVERSITY, HANOI
UNIVERSITY OF LANGUAGES & INTERNATIONAL STUDIES

PHẠM BÍCH NGỌC

DOCTORAL THESIS

**VISUAL IMAGES IN ENGLISH MEDICAL TEXTBOOKS
IN VIETNAM FROM MULTIMODAL
DISCOURSE ANALYSIS PERSPECTIVE**

*(Ngôn ngữ hình ảnh trong giáo trình tiếng Anh chuyên ngành Y
tại Việt Nam từ lý thuyết diễn ngôn đa thức)*

Major: English Linguistics

Code: 9220201.01

Hanoi, 2026

VIETNAM NATIONAL UNIVERSITY, HANOI
UNIVERSITY OF LANGUAGES & INTERNATIONAL STUDIES

PHẠM BÍCH NGỌC

DOCTORAL THESIS

**VISUAL IMAGES IN ENGLISH MEDICAL TEXTBOOKS
IN VIETNAM FROM MULTIMODAL
DISCOURSE ANALYSIS PERSPECTIVE**

*(Ngôn ngữ hình ảnh trong giáo trình tiếng Anh chuyên ngành Y
tại Việt Nam từ lý thuyết diễn ngôn đa thức)*

Major: English Linguistics

Code: 9220201.01

Supervisor: Assoc.Prof.Dr. Lam Quang Dong

Hanoi, 2026

DECLARATION

I hereby declare that this doctoral thesis is entirely my own research work, completed as part of my PhD program at the University of Languages and International Studies, Vietnam National University, Hanoi. I certify that all sources of ideas, data, and quotations used in this thesis - whether published or unpublished - have been fully acknowledged and referenced. I also declare that this doctoral thesis contains no material that has been submitted, in whole or in part, for any other degree or diploma at this or any other institution. This work represents my independent research conducted under the supervision of Assoc.Prof.Dr. Lam Quang Dong at University of Languages and International Studies, Vietnam National University, Hanoi.

Signature of PhD student:

Pham Bich Ngoc

Tables of Contents

Tables of Contents.....	ii
List of Abbreviations.....	xi
Abstract	xii
Acknowledgements	xiii
List of Figures and Tables.....	xiv
CHAPTER I: INTRODUCTION.....	1
1.1 Background and Rationale for the Study	1
1.2 Identification of the Problem	3
1.3 Research Aims and Questions.....	5
1.3.1 Overall Aim of the Study	6
1.3.2 Specific Research Objectives.....	6
1.3.3 Research Questions	7
1.3.4 Justification of Research Focus.....	8
1.4 Scope of the Study	9
1.5 Significance of the Study	11
1.5.1 Theoretical Significance	11
1.5.2 Methodological Significance	13
1.5.3 Pedagogical and Practical Significance	13
1.5.4 Contextual and Societal Significance	14
1.5.5 Overall Contribution	15
1.6 Structure of the Thesis	15

CHAPTER II: LITERATURE REVIEW	18
2.1 Theoretical Background.....	18
2.1.1 English for Specific Purposes	18
2.1.2 Medical ESP Textbooks and Their Pedagogical Roles	21
2.1.3 Social Semiotics and Systemic Functional Linguistics	27
2.1.4 Multimodal Discourse Analysis.....	33
2.2 Grammar of Visual Design (Kress & van Leeuwen, 2021).....	35
2.2.1 Representational Meaning.....	35
2.2.2 Interactive Meaning	37
2.2.3 Compositional Meaning.....	39
2.3 Multimodal Metafunctions: Extending Halliday through Unsworth....	40
2.3.1 Multimodal Metafunctions.....	40
2.3.2 Intersemiosis and Image-Text Relations.....	44
2.4 Previous Research	47
2.4.1 Global Developments in Research on Visual and Multimodal Texts	47
2.4.2 Empirical Findings from Multimodal Educational Research	48
2.4.3 ESP/EMP-related Multimodal Studies	49
2.4.4 Vietnam-based Research on Multimodality and ESP/EMP Textbooks.....	51
2.4.5 Synthesis and Implications.....	52

2.5 Constructing a Social-Semiotic Multimodal Framework: A Systemic Integration of Visual Grammar and Multimodal Metafunctions	53
2.5.1 Shared social-semiotic foundations	53
2.5.2 Systemic complementarity: from intra-modal to inter-modal analysis	54
2.5.3 Critical integration for analysing English medical textbooks.....	56
2.5.4 A stratified social-semiotic model: from modes to multimodal ensembles	57
2.6 Research Synthesis.....	60
2.6.1 Research Gap	61
2.6.2 Conceptual Framework for the Study	67
2.7 Concluding Remarks.....	71
CHAPTER III: RESEARCH METHODOLOGY	75
3.1 Research Paradigm.....	75
3.2 Research Design.....	79
3.3 Data Corpus.....	82
3.3.1 Rationale for Book-level Selection.....	82
3.3.2 Textbook 1: Medical Terminology: A Living Language (Fremgen & Frucht, 2019).....	83
3.3.3 Textbook 2: Career Paths: Medical (Evans, Dooley, & Tran, 2012)	85
3.3.4 Unit of Analysis and Corpus Delimitation.....	86
3.4 Data Selection Criteria	87

3.4.1 Construct-to-Criterion Alignment with the Research Questions...	87
3.4.2 Unit of Analysis: Page Ensemble.....	88
3.4.3 Inclusion and Exclusion Rules.....	89
3.4.4 Textbook-Specific Selection Trajectories.....	90
3.4.5 Corpus Size, Saturation Logic, and Documentation.....	91
3.4.6 Analytical Payoff	91
3.5 Data Collection Procedure	92
3.6 Data Analysis	95
3.6.1 Data Analysis Procedure	95
3.6.2 Pilot Study: Trial Qualitative Analysis of Five Random Page Ensembles.....	98
3.6.3 Integrated Multimodal Analytical Framework	102
3.6.4 Coding and Analytical Phases.....	106
3.6.5 Rubricized Multimodal Codebook.....	113
3.7 Reliability, Validity, and Ethical Considerations	121
3.7.1 Interpretive Orientation and Epistemic Coherence.....	122
3.7.2 Reliability as Reflexive Consistency	122
3.7.3 Validity as Interpretive Sufficiency and Theoretical Saturation .	123
3.7.4 Ethical Integrity and Cultural Sensitivity	123
3.7.5 Consolidation: Trustworthiness as an Emergent Property of Design	124
3.8 Concluding Remarks.....	126

CHAPTER IV: VISUAL MEANING IN MEDICAL TEXTBOOK IMAGERY	128
4.1 Visual Meaning as Disciplinary Vision: Analytical Orientation and Scope	128
4.1.1 Chapter focus and alignment with RQ1	128
4.1.2 Unit of analysis: the page ensemble.....	129
4.1.3 Corpus and evidential base: 40 page ensembles across two textbook ecologies.....	129
4.1.4 Analytical lens: Grammar of Visual Design operationalised within IMAF - Stratum 1.....	130
4.1.5 Coding source and procedural discipline: Visual Codebook + Appendix B-D + interpretive salience calibration	131
4.1.6 Chapter trajectory and discourse logic.....	131
4.2 Distribution of Visual Semiotic Resources across the Corpus	132
4.3 Representational Meaning in Medical Textbook Imagery	134
4.3.1 Conceptual Representation of Medical Knowledge	136
4.3.2 Narrative Representation of Medical Processes and Practices	139
4.4 Interactive Meaning: Viewer Positioning and Epistemic Relations ...	143
4.4.1 Gaze and Visual Address	144
4.4.2 Social Distance, Angle, and Epistemic Relations.....	152
4.4.3 Angle and Authority.....	154
4.4.4 Modality Cues and the Construction of Visual Credibility	155

4.4.5 Interim Synthesis: Interactive Meaning Profiles of MT and CPM	155
4.5 Compositional Meaning: Visual Information Organisation and Pedagogic Control	156
4.5.1 Information Value and Layout Structures.....	157
4.5.2 Salience Patterns and Attention Management	162
4.5.3 Framing, Segmentation, and Integration.....	164
4.5.4 Reading Paths and Multimodal Navigation	165
4.5.5 Interim Synthesis.....	165
4.6 Integrated Visual Meaning Profiles: Consolidation of Findings for RQ1	166
4.6.1 Visual Meaning Profile of MT: “Terminological-Conceptual Regime”.....	166
4.6.2 Visual Meaning Profile of CPM: “Contextual-Professional Regime”.....	168
4.6.3 Cross-textbook Comparison: Disciplinary Vision and Pedagogic Orientation.....	170
4.7 Concluding Remarks	171
CHAPTER V: INTERSEMIOTIC MEANING IN MEDICAL TEXTBOOKS	173
5.1 Analytical Orientation and Intersemiotic Scope	173
5.2 Distribution of Intersemiotic Relations across the Corpus	174
5.3 Intersemiotic Elaboration and Conceptual Stabilisation.....	176

5.3.1 Systematic Elaboration and Conceptual Stabilisation in Medical Terminology	176
5.3.2 Selective Elaboration in Career Paths: Medical.....	179
5.4 Intersemiotic Enhancement and Procedural Meaning-Making	182
5.4.1 Procedural enhancement in Medical Terminology: constrained but epistemically functional	183
5.4.2 Procedural enhancement in Career Paths: Medical: action-oriented and performative	185
5.4.3 Comparative interpretation	186
5.5 Projection and the Enactment of Professional Roles	187
5.5.1 Strong projection in Career Paths: Medical	187
5.5.2 Weaker/constrained projection in Medical Terminology	189
5.5.3 Synthesis: projection as the bridge to pedagogy.....	191
5.6 Intersemiotic Density and Cross-Textbook Synthesis	192
5.7 Emergent Multimodal Pedagogy in Medical English Instruction	195
5.7.1 Epistemic multimodal pedagogy in Medical Terminology	195
5.7.2 Enactment-oriented multimodal pedagogy in Career Paths: Medical	196
5.7.3 Integrative implications.....	197
5.7.4 Implications for Vietnamese EMP classroom practice.....	198
5.8 Concluding Remarks	201
CHAPTER VI: CONCLUSIONS	203
6.1 Restatement of Research Purpose and Methodological Orientation ..	203

6.2 Summary of Key Findings in Relation to Research Questions	206
6.2.1 Key Findings for RQ1: Visual Meaning in EMP Textbooks	206
6.2.2 Key Findings for RQ2: Intersemiotic Meaning	208
6.2.3 Integrated Answer to the overarching research question.....	210
6.3 Theoretical Contributions	211
6.3.1 Advancing Multimodal Theory in EMP Research	212
6.3.2 Operationalising IMAF as an Integrated Explanatory Framework	213
6.3.3 Re-conceptualising Intersemiotic Density as an Epistemic Indicator	215
6.4 Pedagogical Contributions	217
6.4.1 Multimodal Disciplinary Literacy in Medical English.....	217
6.4.2 Two Complementary Multimodal Pedagogies	219
6.5 Practical and Institutional Implications for the Vietnamese Context .	221
6.5.1 Implications for EMP Curriculum Design.....	222
6.5.2 Implications for Teaching Practice	223
6.5.3 Implications for Materials Selection and Adaptation	225
6.6 Limitations of the Study.....	227
6.6.1 Corpus Scope and Analytic Generalisability	227
6.6.2 Focus on Semiotic Affordances Rather than Learning Outcomes	227
6.6.3 Domain-Specificity of EMP.....	228

6.6.4 Theoretical Constraints of Multimodal Frameworks	228
6.6.5 Interpretive Mediation in Multimodal Analysis	229
6.7 Directions for Future Research	230
6.7.1 Classroom Enactment of Textbook Multimodality	230
6.7.2 Assessing Multimodal Medical Literacy	231
6.7.3 Comparing Imported Textbooks with Vietnamese-Produced EMP Materials.....	231
6.7.4 Longitudinal Studies of EMP Curriculum Design and Sequencing	232
6.8 Concluding Remarks.....	233
REFERENCES.....	235
APPENDICES.....	247

List of Abbreviations

CPM	Career Paths: Medical
CFS	Conceptual Framework of the Study
ESP	English for Specific Purposes
EMP	English for Medical Purposes
GVD	Grammar of Visual Design
OSCEs	Objective Structured Clinical Examinations
ICF	Integrated Conceptual Framework
IMAF	Analytical Framework: Integrated Multimodal Analytical Framework
MDA	Multimodal Discourse Analysis
MT	Medical Terminology: A Living Language
SFL	Systemic Functional Linguistics

Abstract

This dissertation investigates how medical meaning and pedagogic orientations are co-constructed through visual and linguistic resources in English for Medical Purposes textbooks used in Vietnamese medical universities. Situated in multimodal discourse analysis and grounded in social-semiotic and intersemiotic theory, the study conceptualises textbooks as designed multimodal artefacts in which disciplinary knowledge and professional positioning emerge through systematic visual-verbal orchestration rather than linguistic content alone. Using a theory-driven, interpretive-explanatory qualitative design, the research analyses forty purposively selected page ensembles from two widely used series: *Medical Terminology: A Living Language* and *Career Paths: Medical*. It develops and operationalises an Integrated Multimodal Analytical Framework, drawing on the Grammar of Visual Design (Kress & van Leeuwen, 2021) and intersemiotic relations and metafunctional alignment (Unsworth, 2008) across three strata: visual meaning, image-text relations, and pedagogic interpretation. Findings show that the textbooks realise distinct multimodal regimes. *Medical Terminology* foregrounds dense diagrammatic elaboration, classificatory organisation, and observer-oriented design, stabilising biomedical concepts and privileging epistemic control. By contrast, *Career Paths: Medical* prioritises contextualised imagery, procedural enhancement, and interactional projection, supporting professional action and role-based communication. Overall, the study demonstrates that medical meaning emerges from a systematic division of semiotic labour across modes, challenging the view of images as mere illustrations. The dissertation validates the Integrated Multimodal Analytical Framework as an explanatory approach and clarifies how multimodal textbook design shapes medical knowing, professional positioning, and pedagogic orientation in the Vietnamese EMP context.

Key words: English for Medical Purposes, Integrated Multimodal Analytical Framework, Grammar of Visual Design, intersemiotic complementarity, multimodal textbook design; multimodal disciplinary literacy, Vietnam.

Acknowledgements

This dissertation marks the culmination of a long and intellectually rewarding journey, which would not have been possible without the guidance, encouragement, and generosity of many individuals and institutions to whom I owe my deepest gratitude.

First and foremost, I would like to express my heartfelt appreciation to my supervisor, Assoc.Prof.Dr. Lam Quang Dong, at the University of Languages and International Studies, Vietnam National University, Hanoi. His insightful guidance, scholarly expertise, and patient mentorship have been invaluable throughout every stage of my research. His intellectual rigor and encouragement have not only shaped this thesis but also profoundly influenced my academic growth and professional identity.

My sincere thanks also go to professors for their constructive feedback, critical insights, and thoughtful comments during seminars and the doctoral committee, which have significantly enriched the analytical and theoretical dimensions of this thesis.

I am equally grateful to my colleagues and fellow doctoral candidates at the Faculty of Postgraduate Studies, ULIS - VNU, whose stimulating discussions, collaboration, and mutual support have made this academic journey truly inspiring.

Special thanks are extended to the faculty members and administrators at Hai Phong University of Medicine and Pharmacy, and other medical institutions, for granting me access to the English medical textbooks used in this research. Their cooperation and assistance during the data collection process were essential to the successful completion of this research.

Finally, my deepest gratitude goes to my beloved family and friends, whose unwavering love, patience, and encouragement have been my greatest motivation and strength throughout the challenges of this doctoral pursuit.

To all who have supported me academically, professionally, and personally I offer my sincere and heartfelt thanks.

List of Figures and Tables

Figures	Page
Figure 2.1 Cyclic Stratified Model of Meaning in SFL	31
Figure 2.2 Conceptual Framework of the Study	71
Figure 3.1 Data Collection Procedure	92
Figure 3.2 Data Analysis Procedure	97
Figure 3.3 Integrated Multimodal Analytical Framework	103
Figure 3.4 Reflexive Trustworthiness Framework	124

Tables	Page
Table 3.1 Interpretive Salience Calibration Scale (0-3)	111
Table 3.2 Mapping Research Questions to Analytical Focus and Coding Phases	112
Table 3.3 Module A: Visual Meaning Rubric	115
Table 3.4 Module B: Linguistic Resources in Intersemiotic Alignment	117
Table 3.5 Module C: Intersemiotic Coupling Rubric	118
Table 3.6 Interpretive Salience Calibration Scale (0-3)	119

CHAPTER I: INTRODUCTION

This chapter introduces the study and establishes its conceptual and contextual foundations. It begins by outlining the background and rationale for examining visual images in English medical textbooks used in Vietnam. It then identifies the research problem and presents the aims, objectives, and research questions guiding the investigation. The chapter also discusses the significance of the study and delineates its scope. Collectively, these sections lay the groundwork for the theoretical and methodological framework developed in the subsequent chapters.

1.1 Background and Rationale for the Study

In an era where knowledge circulation is increasingly mediated through visual and digital modes, understanding how images contribute to meaning-making has become essential in educational discourse. Contemporary texts in science and medicine no longer rely solely on verbal language; rather, they function as multimodal ensembles in which images, text, layout, and graphical symbolism jointly construct meaning (Bezemer & Kress, 2016). Within the broader field of English for Specific Purposes (ESP), English for Medical Purposes (EMP) represents a specialised branch concerned with the communication of medical knowledge and professional practice. EMP textbooks commonly employ anatomical diagrams, procedural flowcharts, clinical photographs, radiological images, and infographics to support learners' conceptual understanding of complex biomedical phenomena.

These visual resources, however, are frequently treated as supplementary or decorative rather than as integral meaning-making components in the Vietnamese ESP/EMP context. Pedagogical practices and textbook evaluation tend to foreground linguistic elements – terminology, grammar, and reading comprehension – while implicitly reducing visuals to

secondary status. Such an orientation potentially underestimates the epistemic role of images in shaping disciplinary literacy, clinical reasoning, and the acquisition of professional knowledge, particularly for learners of English as a foreign language. For these learners, the synergy between visual and verbal modes may significantly scaffold or constrain their comprehension of medical concepts.

The theoretical motivation for examining this synergy derives from the multimodal extension of Systemic Functional Linguistics (SFL). Halliday's (1978) metafunctional model conceptualises meaning-making through ideational, interpersonal, and textual functions. Building on this foundation, Kress and van Leeuwen (1996/2006/2021) developed the Grammar of Visual Design (GVD), articulating how representational structures depict participants and processes, how interactive structures position viewers through gaze and distance, and how compositional structures organise visual information. GVD offers a powerful toolkit for reading images as semiotic resources.

Nonetheless, visual grammar alone is insufficient for understanding how images and text integrate to construct coherent disciplinary meanings. Unsworth's (2006) work on multimodal metafunctions and intersemiosis highlights the need to examine how linguistic and visual modes complement, extend, or reinforce one another within multimodal texts. His metafunctional mapping extends SFL across modes, enabling a fuller account of multimodal meaning-making.

Despite the availability of these theoretical frameworks, empirical studies applying them to English medical textbooks – particularly in non-Anglophone contexts such as Vietnam – remain scarce. Research in multimodal discourse analysis has examined science textbooks (e.g., Tang, 2013) and EFL/ESL language textbooks (Hutchinson and Torres, 1994), yet

little is known about how EMP textbooks used in Vietnam visually and verbally construct medical knowledge. Local ESP research tends to focus on vocabulary, metaphor, reading strategies, or translation, with minimal attention to the multimodal design of instructional materials.

EMP textbooks widely used in Vietnam such as *Medical Terminology: A Living Language* (Fremgen & Frucht, 2019) and *Career Paths: Medical* (Evans, Dooley & Tran, 2012) serve as core materials in medical universities. However, their visual systems, semiotic affordances, and intermodal dynamics remain underexplored. This study is therefore motivated by the need to bridge theoretical gaps in multimodal discourse studies, pedagogical gaps in ESP/EMP material design, and contextual gaps in Vietnam's medical English education. Investigating how visual images and verbal language co-construct medical discourse can provide new insights into multimodal meaning-making and inform more effective pedagogical practices in ESP/EMP classrooms.

1.2 Identification of the Problem

Over the past three decades, Multimodal Discourse Analysis (MDA) has developed into a significant paradigm within applied linguistics, expanding the scope of discourse studies beyond language to encompass image, layout, gesture, sound, and other semiotic resources (Jewitt, 2014). MDA foregrounds the idea that meaning-making emerges through the patterned interplay of modes, each with distinctive affordances and constraints (Kress, 2010). However, in educational settings and particularly within ESP – research and pedagogical practices often remain linguistically oriented, treating visuals as ancillary rather than constitutive components of disciplinary knowledge (Bateman, 2014). This discrepancy has resulted in a

critical gap between theoretical advancements in multimodality and their pedagogical application.

Within EMP, textbooks function as primary multimodal environments through which learners encounter specialised knowledge. These materials integrate verbal, visual, and typographical resources to represent anatomical processes, clinical procedures, and biomedical mechanisms. Yet, despite the centrality of visuals in constructing disciplinary understanding, few studies have systematically examined how linguistic and visual modes interact to convey medical meaning. EMP research internationally has often focused on terminology development (Bosher & Smalkoski, 2002), discourse structures (Salager-Meyer, 2014), or pragmatic communication in clinical settings (Le et al., 2021), with limited attention to multimodal design.

In Vietnam, this research gap is even more pronounced. Although English has become indispensable in medical education due to the globalisation of scientific knowledge (Bui, 2024), local studies on English medical materials primarily examine linguistic skills or translation practices (Bui, 2016). Consequently, the semiotic and multimodal dimensions of EMP textbooks such as how diagrams, photographs, and infographics articulate with linguistic explanations – remain under-investigated.

Moreover, while international scholarship offers robust frameworks for analysing multimodal texts such as Kress and van Leeuwen's GVD and Unsworth's multimodal metafunctions – few studies have integrated these approaches or adapted them to local ESP/EMP contexts. The absence of an integrated analytical framework limits our understanding of how multimodal meanings are orchestrated pedagogically across modes.

Additionally, textbook analysis in Vietnam has traditionally relied on content analysis or linguistic description, providing little insight into visual-

verbal coherence or semiotic complementarity. In the medical domain, where visual representation is epistemic rather than merely illustrative (Goodwin, 1994), the lack of multimodal evaluation criteria poses risks: poorly aligned images may hinder comprehension and impede the development of professional vision.

Thus, the central problem addressed in this study is the lack of empirical and theoretical understanding of how visual images contribute to meaning-making in English medical textbooks used in Vietnam. This problem encompasses three interrelated gaps:

(i) *A theoretical gap*: current MDA research lacks an integrated framework uniting the descriptive strengths of GVD with the interpretive depth of multimodal metafunctions.

(ii) *A pedagogical gap*: ESP/EMP teaching practices and textbook evaluation rarely consider multimodal literacy or the semiotic interplay between text and image.

(iii) *A contextual gap*: there appears to be no comprehensive study examining how EMP textbooks used in Vietnam represent medical knowledge multimodally.

Addressing these gaps, this study proposes an Integrated Multimodal Analytical Framework (IMAF) to analyse how linguistic and visual modes co-construct medical meaning in selected EMP textbooks. Through qualitative multimodal discourse analysis, the study aims to describe and interpret intersemiosis between text and image, thereby contributing to theoretical development in MDA and pedagogical advancement in ESP/EMP education.

1.3 Research Aims and Questions

The preceding discussion situates this research at the intersection of MDA and EMP. EMP textbooks constitute highly multimodal sites of

disciplinary meaning-making, yet little empirical work has examined how their visual and linguistic resources function together to construct medical knowledge, particularly within Vietnam.

1.3.1 Overall Aim of the Study

The overarching aim of this doctoral research is to analyse how visual and linguistic modes interact to construct medical knowledge and professional meaning in English medical textbooks used in Vietnam. This aim is pursued through the development and application of an IMAF, grounded in Systemic Functional Linguistics and synthesising Kress and van Leeuwen's (2021) GVD with Unsworth's (2008) multimodal metafunctional model.

1.3.2 Specific Research Objectives

To achieve this aim, the study pursues the following specific objectives:

1. To identify and categorise the types of visual images employed in selected English medical textbooks, based on their representational, interactional, and compositional structures.
2. To analyse how visual and linguistic modes interact to realise ideational, interpersonal, and textual meanings, drawing on the integration of visual grammar and multimodal metafunctions.
3. To interpret how multimodal configurations construct medical concepts, roles, and relationships, contributing to learners' disciplinary literacy and professional reasoning.
4. To evaluate the pedagogical implications of these multimodal practices for ESP/EMP material design and instructional approaches in Vietnamese medical universities.

These objectives are framed by the recognition that multimodal literacy constitutes an essential competence for contemporary medical professionals,

where the interpretation of visual-verbal synthesis forms the foundation of biomedical communication (Jewitt, 2014; Kress, 2010).

1.3.3 Research Questions

In line with the overarching aim and specific objectives outlined above, this study addresses one overarching research question and two subordinate questions. Together, they are designed to capture complementary analytical dimensions of multimodal meaning-making in English medical textbooks, as interpreted through the IMAF.

Overarching Research Question (RQ)

How do visual and linguistic modes interact to construct medical meaning and disciplinary knowledge in English medical textbooks used in Vietnam, as interpreted through the Integrated Multimodal Analytical Framework?

Subordinate Research Questions

RQ1. In what ways are visual images employed and organised representationally, interactively, and compositionally across the selected English medical textbooks?

RQ2. How do visual and linguistic modes jointly realise ideational, interpersonal, and textual meanings within the IMAF-based analysis?

These questions move beyond descriptive cataloguing of visual or linguistic features. Instead, they adopt an interpretive and evaluative stance, seeking to elucidate how multimodal configurations embody disciplinary ways of seeing and knowing in medical education (Goodwin, 1994). In doing so, the research questions establish the analytical pathways that will be developed through the IMAF in Chapter II and operationalised in the methodological procedures of Chapter III.

1.3.4 Justification of Research Focus

The formulation of the above research questions is grounded in both conceptual considerations and empirical needs identified in the literature and in the Vietnamese EMP context.

First, Kress and van Leeuwen's (2021) GVD explicates the semiotic organisation of images by detailing representational, interactive and compositional structures, alongside analytical categories such as participants, vectors, gaze, salience and information value. While this framework is powerful in describing how images mean, it does not fully account for how they mean with language in integrated multimodal texts. Secondly, Unsworth (2006) theorises intersemiosis – the dynamic interplay between language and image – through metafunctional alignment, mapping ideational, interpersonal and textual meanings across modes. However, his work offers less elaboration on the internal structural organisation of images themselves.

Taken separately, these frameworks therefore provide partial but incomplete accounts of multimodal meaning-making. Integrating them within the IMAF allows for a more comprehensive analysis of how intra-modal structures (within image and text) and inter-modal relations (between image and text) jointly contribute to the construction, negotiation and pedagogical realisation of disciplinary knowledge in educational discourse (Royce, 2007; Unsworth, 2008).

Moreover, in the specific context of Vietnamese higher education, English medical textbooks are predominantly imported or adapted from Western publishers such as Pearson and Express Publishing. The visual culture embedded in these materials may reflect Western epistemological orientations in medicine that do not always align with local pedagogical expectations or cultural assumptions (Le, 2017). A multimodal analysis of

such textbooks can therefore reveal not only textual meanings but also potential ideological orientations embedded in visual representation, resonating with Fairclough's (2010) notion of the semiotic framing of knowledge.

Finally, the research responds to calls in both global and Vietnamese educational policy for the development of multimodal literacy, particularly within foreign language education (Ministry of Education and Training [MOET], 2018). As Kress (2010) argues, the future of literacy is fundamentally multimodal, requiring learners to interpret and produce texts in which images, layout and other semiotic resources are central rather than peripheral. Examining the semiotic architecture of English medical textbooks thus becomes both an academic necessity and a pedagogical imperative.

In sum, the research questions are designed to extend multimodal discourse theory through the synthesis of visual grammar and metafunctional models; provide empirical insight into multimodal design in ESP/EMP materials in an under-researched Asian context; and inform pedagogical practices that integrate multimodal literacy into medical English education.

In this way, the focus of the study builds directly on the background and problematisation presented in Sections 1.1 and 1.2, and operationalises the aims and objectives specified in Section 1.3.

1.4 Scope of the Study

The scope of this research is delimited by its focus, data set, and analytical orientation. Building on the background, problematisation and aims outlined in the preceding sections, the study investigates the interaction between linguistic and visual modes in two English medical textbooks currently used in Vietnamese tertiary institutions. Its primary objective is to

describe and interpret how multimodal resources work together to construct disciplinary meaning and pedagogical function within these materials.

The analysis concentrates on two representative textbooks:

Medical Terminology: A Living Language (7th ed.) by Fremgen and Frucht (2019), published by Pearson Education; and

Career Paths: Medical by Evans, Dooley, and Tran (2012), published by Express Publishing.

These textbooks were selected because they are widely adopted in Vietnamese medical universities (e.g., Hai Phong University of Medicine and Pharmacy, Da Nang University of Medical Technology and Pharmacy, Can Tho University of Medicine and Pharmacy) and exemplify two pedagogical orientations in EMP: a terminology-based model and a more skills-based communication-oriented model. This purposive selection allows for multimodal analysis across distinct instructional genres (Miles, Huberman, & Saldaña, 2014).

The study focuses on units and lessons that include both visual and verbal content – specifically illustrations, diagrams, photographs, captions, and the accompanying texts. Non-pictorial elements such as exercises, glossaries, or pronunciation guides are examined only when they contribute directly to multimodal meaning-making (for instance, through typographical emphasis or iconographic symbols). The study is limited to printed, officially published materials, excluding supplementary digital components (audio, video, online exercises). This delimitation ensures analytical consistency, as dynamic digital modes would require additional theoretical and methodological tools beyond the scope of the present framework..

The analytical focus is qualitative, descriptive, and interpretive, employing the IMAF derived from Kress and van Leeuwen's (2021) GVD

and Unsworth's (2008) multimodal metafunctions. The study therefore seeks to illuminate patterns of multimodal meaning-making rather than to quantify frequencies or test causal relationships. In line with Creswell and Poth's (2018) view of qualitative inquiry, priority is given to depth of understanding over breadth of coverage.

Geographically and institutionally, the study is situated within the Vietnamese higher education context, focusing on the adaptation and pedagogical use of imported English medical textbooks. Although the materials originate from global publishers, their analysis is interpreted through the lens of local educational needs, policy directions and cultural considerations (Le, 2017). Consequently, the scope of the research is bounded within the intersection of multimodal discourse, ESP/EMP pedagogy and Vietnamese tertiary education, providing a focused yet contextually grounded basis for the subsequent chapters of the thesis.

1.5 Significance of the Study

The significance of this study lies in its contribution to the theoretical development, methodological refinement and pedagogical enhancement of multimodal discourse research within ESP, and more specifically within English for Medical Purposes. By examining the interaction between linguistic and visual modes in English medical textbooks used in Vietnam, the study addresses a crucial yet underexplored dimension of disciplinary meaning-making in educational discourse and responds to the gaps outlined in the previous sections.

1.5.1 Theoretical Significance

From a theoretical perspective, the study advances Multimodal Discourse Analysis by proposing and applying an Integrated Multimodal Analytical Framework that synthesises two influential yet distinct models:

Kress and van Leeuwen's (2021) GVD and Unsworth's (2008) multimodal metafunctional framework. Kress and van Leeuwen conceptualise the visual mode as a semiotic system capable of realising representational, interactive and compositional meanings; however, their account remains primarily descriptive and intra-modal in orientation. Conversely, Unsworth provides a metafunctional model that interprets the intersemiosis between language and image but does not elaborate detailed visual structural categories.

The integration of these frameworks addresses an analytical gap in multimodal discourse studies (Jewitt, 2014; Royce, 2007) by offering a more comprehensive lens for examining how multimodal resources function together to construct knowledge in disciplinary contexts. The resulting IMAF enables a dual-level analysis: it simultaneously attends to intra-modal structures within each mode and to inter-modal relationships across modes. In this way, the study contributes to ongoing theoretical conversations about how multimodal meaning systems are realised, co-ordinated and interpreted in educational discourse – a concern central to both systemic functional linguistics (Halliday, 1978) and multimodal social semiotics (van Leeuwen, 2005).

Furthermore, by applying this integrated framework to medical English textbooks, the research extends the scope of multimodal theory beyond more commonly studied domains such as picture books (Painter, Martin, & Unsworth, 2013) and media discourse (Machin & Mayr, 2023). It demonstrates the applicability of MDA to disciplinary pedagogic texts, thereby recontextualising multimodal theory within ESP/EMP material analysis and providing a conceptual and analytical template for future research in other specialised domains such as nursing, engineering or business.

1.5.2 Methodological Significance

Methodologically, the study exemplifies how qualitative multimodal analysis can be operationalised in an educational context through a systematic and transparent procedure. By integrating visual grammar, metafunctional mapping and pedagogical discourse interpretation, it offers a layered approach to meaning analysis that transcends the limitations of single-mode textual analysis.

The use of a purposively selected corpus of two widely used medical English textbooks – *Medical Terminology: A Living Language* (7th ed.; Fremgen & Frucht, 2019) and *Career Paths: Medical* (Evans, Dooley, & Tran, 2012) – demonstrates how multimodal methods can yield fine-grained insights into meaning-making practices across textual, visual and layout dimensions. In doing so, the study responds to Kress's (2010) call for methodologically coherent accounts of multimodal design in communication.

In addition, the development of a codebook and data inventory matrix, together with procedures such as peer debriefing to enhance dependability, contributes to the methodological transparency and trustworthiness of multimodal analysis – an area sometimes criticised for being overly interpretive or subjective (Cresswell, 2009). The study thus shows that qualitative research in MDA can maintain interpretive depth while meeting expectations of analytic rigour and accountability.

1.5.3 Pedagogical and Practical Significance

The pedagogical relevance of this research lies in its potential to enhance multimodal literacy and ESP/EMP pedagogy in Vietnamese higher education. Medical knowledge is frequently conceptualised and communicated through diagrams, illustrations, symbolic representations and visual metaphors (Lemke, 1998). Understanding how these multimodal

resources are deployed in language learning materials is therefore crucial for designing effective curricula and instructional practices.

The findings of this study are expected to:

- inform ESP/EMP material designers about how to balance textual and visual content in order to optimise learners' comprehension of complex medical concepts;
- provide teachers with analytical awareness of how multimodal affordances can be exploited to scaffold discipline-specific literacy; and
- support policy initiatives by the Vietnamese Ministry of Education and Training (MOET, 2008) concerning the development of learners' multimodal competence within the National Foreign Language Project.

Furthermore, the study offers a critical basis for reconsidering the pedagogical use of imported textbooks. By revealing how visual and verbal modes are configured, it can assist curriculum developers and teacher educators in localising instructional materials so that they better align with Vietnamese learners' cognitive, cultural and disciplinary needs (Le, 2017). In this way, the research bridges a practical gap between global multimodal design and local pedagogical implementation.

1.5.4 Contextual and Societal Significance

In a broader educational and societal context, the study contributes to Vietnam's ongoing transition towards knowledge-based and communicatively competent higher education. By foregrounding multimodal meaning-making, it underscores the importance of visual and semiotic awareness as integral components of professional communication in healthcare. As Goodwin (1994) observes, professional expertise often depends on the ability to interpret and produce complex multimodal artefacts, such as medical charts, diagnostic images and clinical documentation. Enhancing understanding of

the multimodal construction of English medical textbooks can therefore support not only language education but also interdisciplinary communication competence within the medical profession.

1.5.5 Overall Contribution

In summary, the significance of this study is threefold:

Theoretical contribution: It extends multimodal discourse theory by integrating visual grammar and metafunctional models into a unified analytical framework applicable to disciplinary discourse.

Methodological contribution: It demonstrates the feasibility of systematic, replicable qualitative multimodal analysis in ESP/EMP research, establishing procedures that enhance transparency and validity.

Pedagogical and contextual contribution: It provides the feasibility of systematic, replicable qualitative multimodal analysis in ESP/EMP research, establishing procedures that enhance transparency and validity.

Collectively, these contributions position multimodal discourse analysis not merely as a descriptive tool but as a transformative approach to understanding and improving the ways in which specialised knowledge is communicated, learned and taught, thereby reinforcing the rationale and aims articulated in Sections 1.1 – 1.3.

1.6 Structure of the Thesis

This doctoral thesis is organized into seven chapters that collectively develop the theoretical basis, methodological approach, analytical findings, and pedagogical implications of the study. The structure progresses logically from contextualization to interpretation, ensuring coherence between research questions, analysis, and educational relevance.

Chapter 1: Introduction establishes the foundation of the research. It outlines the background and rationale, identifies the problem addressed, and

formulates the research aims and questions. The chapter clarifies the significance of examining visual meaning in English medical textbooks within the context of ESP and medical education in Vietnam. The methodological orientation, scope, and limitations are also introduced, positioning the study within the intersecting fields of MDA and medical English pedagogy.

Chapter 2: Literature Review surveys theoretical and empirical work on multimodal meaning-making in English medical textbooks. It positions the study within ESP, multimodal discourse analysis, and social semiotics, foregrounding visual and visual-verbal resources. Synthesising global and Vietnamese research, it identifies conceptual and methodological gaps, motivating conceptual frameworks underpinning the thesis's empirical analysis.

Chapter 3: Research Methodology outlines the methodology used to investigate how visual and linguistic modes interact to construct medical meaning in English medical textbooks used in Vietnam. Anchored in a theory-driven, interpretive-explanatory social semiotic paradigm, it conceptualises meaning as multimodal design rather than transparent representation. The chapter details the research design, corpus construction, and data procedures, establishes the page ensemble as the unit of analysis, and presents the IMAF, coding phases, codebook, and trustworthiness measures guiding the analysis.

Chapter 4: Visual Meaning in Medical Textbook Imagery investigates how visual images construct medical meaning in *Medical Terminology: A Living Language* and *Career Paths: Medical*. Drawing on the Integrated Multimodal Analytical Framework and the Grammar of Visual Design, it addresses RQ1.1 through representational, interactive, and

compositional analysis. Using 40 page ensembles and systematically coded evidence, the chapter identifies corpus-wide visual patterns that ground the subsequent intersemiotic analysis.

Chapter 5: Intersemiotic Meaning in Medical Textbooks extends the analysis from visual meaning to intersemiotic meaning, examining how images and language are systematically coordinated within page ensembles to construct medical knowledge. Building on Chapter IV, it explains how image–text relations organise conceptual understanding, procedural reasoning, and professional role enactment, showing how multimodal design renders medical knowledge pedagogically learnable.

Chapter 6: Conclusions synthesises the thesis’ core insights by integrating empirical findings, theoretical contributions, and Vietnam-relevant implications from the comparative multimodal analysis of two EMP textbook series. It argues that medical meaning and pedagogy are semiotically designed through visual-intersemiotic orchestration, then restates the research purpose, answers the research questions, clarifies contributions and limitations, and outlines directions for future research.

References and Appendices follow the main chapters.

CHAPTER II: LITERATURE REVIEW

This chapter establishes the theoretical and empirical grounding for analysing how English medical textbooks construct disciplinary meaning through multimodal resources. As EMP materials increasingly integrate specialised language with diagrams, labels, and layout design, meaning-making must be examined beyond language alone. Accordingly, this chapter reviews key ESP/EMP developments, synthesises social-semiotic and SFL-informed multimodal perspectives, identifies persistent gaps - especially in Vietnam - and thereby justifies the need for an integrated framework to analyse both visual structure and image–text relations in textbook page ensembles.

2.1 Theoretical Background

2.1.1 English for Specific Purposes

English for Specific Purposes has been widely recognised as a major strand within applied linguistics, developed in response to the post-war expansion of scientific, technological, and professional communication in English. This historical context created a strong demand for functional language teaching aligned with discipline-specific and workplace needs, rather than general proficiency goals. Early ESP scholarship was therefore shaped by register-oriented approaches, which aimed to describe the linguistic features that distinguish specialised discourse from everyday language. This tradition is well exemplified by foundational work on scientific and technical English, where researchers attempted to identify recurring lexico-grammatical patterns and communicative routines that could support the teaching of specialised varieties (Ewer & Latorre, 1969). Such work provided an important starting point for understanding how discipline-specific language differs in terms of terminology, syntax, and textual density.

However, as ESP matured, register analysis was increasingly seen as insufficient for explaining how specialised communication operates as a social practice. Focusing mainly on linguistic surface features risks reducing disciplinary discourse to a set of technical forms, while overlooking the pragmatic intentions, rhetorical strategies, and epistemic work through which professional communities construct and legitimise knowledge. This critique paved the way for a major conceptual shift in the 1980s and 1990s, when ESP research increasingly incorporated discourse analysis and genre-based perspectives. A pivotal turning point was Swales' genre theory, which framed genres as socially recognised communicative events shaped by shared purposes and conventions within discourse communities (Swales, 1990). This reframing transformed ESP from a largely descriptive approach to specialised language into an analytical lens concerned with how texts function to achieve professional goals.

Building on Swales' insights, Bhatia (1993) further established genre analysis as a core foundation for ESP by demonstrating how professional genres are strategically organised to enact institutional actions. Rather than treating specialised discourse as a static linguistic system, genre-based ESP conceptualises disciplinary writing and speaking as purposeful, socially situated practices through which professional identities and community membership are negotiated. Within this view, ESP learning is not primarily the accumulation of technical vocabulary, but the development of competence to interpret and produce texts that conform to disciplinary expectations and rhetorical norms. This is closely aligned with the notion of disciplinary literacy, where learners are progressively inducted into the communicative conventions and epistemological assumptions of their target fields (Dudley-Evans & St John, 1998; Hyland, 2005). Importantly, this induction involves

mastering not only what is said, but also how knowledge claims are structured, evaluated, and positioned in relation to readers and professional standards.

More recently, ESP scholarship has continued to expand in response to transformations in professional communication practices. Contemporary academic and workplace texts increasingly operate as multimodal ensembles, integrating language with images, diagrams, tables, symbols, and spatial design. Consequently, meaning-making cannot be fully captured through linguistic analysis alone. Multimodality research has highlighted that texts are structured through multiple semiotic resources, and these resources work together to realise disciplinary meanings in systematic ways (Jewitt, 2014; Bezemer & Kress, 2016). This has significant implications for ESP, particularly in domains where visual representations play a central epistemic role. In scientific and medical contexts, for example, diagrams and visualisations often function not merely as decorative illustrations, but as semiotic tools that enable conceptual reasoning and communicate complex processes and relations. As Lemke (1998) argues, scientific meaning is frequently distributed across verbal and visual modes, requiring learners to interpret these modes in combination rather than in isolation. These developments point to a pedagogical need for ESP curricula to incorporate multimodal literacy, supporting learners in interpreting and producing meaning across interconnected semiotic systems.

Within this evolving theoretical landscape, English for Medical Purposes (EMP) has emerged as a specialised and pedagogically demanding sub-field of ESP. Medical discourse encompasses a wide spectrum of genres, including clinical notes, diagnostic reports, patient histories, procedural guidelines, and research articles, each shaped by distinctive pragmatic and

ethical constraints as well as discipline-specific terminologies and conventions (Candlin & Candlin, 2002; Salager-Meyer, 2014). Notably, medical knowledge is profoundly visual: anatomical diagrams, radiological images, procedural illustrations, and chart-based representations are central to how medical understanding is constructed and communicated. Therefore, EMP competence requires learners to coordinate verbal and visual resources as part of professional meaning-making. In this sense, EMP textbooks play a crucial mediating role in disciplinary enculturation, and their design warrants careful analysis not only in linguistic terms but also in relation to multimodal meaning potential - an issue developed further in the next section.

2.1.2 Medical ESP Textbooks and Their Pedagogical Roles

Within the broader landscape of EMP, textbooks occupy a particularly consequential position as semiotic, epistemological, and pedagogical mediators. Unlike general English teaching materials, EMP textbooks do not merely introduce specialist vocabulary or reading skills; they construct access to biomedical knowledge through interdependent linguistic and visual systems. As Troncoso (2010) notes, teaching materials are always “*cultural and ideological artefacts*” encoding specific assumptions about knowledge, legitimacy, and learning. Accordingly, medical ESP textbooks function not only as linguistic scaffolds but also as semiotic architectures through which learners encounter, interpret, and internalise disciplinary ways of seeing and knowing – a process especially salient in visually dense fields such as medicine.

2.1.2.1 Textbooks as Pedagogical Artefacts

Textbooks, particularly within ESP, have long been recognised as highly structured pedagogical artefacts that embody an underlying pedagogical logic and epistemological framing (Cunningsworth, 1995;

McGrath, 2013). Their design reflects deliberate choices concerning what counts as essential disciplinary knowledge, how such knowledge should be sequenced, and what literacy practices learners are expected to acquire. Following Hyland (2006), ESP textbooks can be understood as mediating devices that recontextualise complex professional genres into pedagogically manageable forms, transforming authentic discourse practices into simulations suited to instructional purposes.

In EMP contexts, this recontextualisation is particularly intricate. Medical discourse is cognitively demanding and structurally specialised, involving a dense network of technical terminology, nominalisations, schematic reasoning, and pragmatic conventions tied to professional roles (Salager-Meyer, 2014). Textbooks must therefore balance linguistic simplification with conceptual accuracy, offering learners a structured pathway into biomedical epistemology. This process resonates with what Levin and Mayer (1993) term pedagogical recontextualisation, wherein textbooks actively shape how learners conceptualise medicine – not only through linguistic representation but also through the selection, modification, and sequencing of visual resources.

Yet, despite their pedagogical importance, research indicates that many EMP materials continue to privilege lexico-grammatical coverage over genre-based and knowledge-based representation (Tran, 2024). In Vietnam, students often perceive EMP textbooks as inadequately aligned with clinical communicative demands, especially regarding authentic discourse patterns or reasoning structures. These observations foreground the need to examine not only the linguistic dimension of textbooks but also their multimodal and epistemic orchestration.

2.1.2.2 The Multimodal Dimension

A defining characteristic of contemporary EMP textbooks is their inherently multimodal semiotic design. Grounded in social semiotics, Kress and van Leeuwen (2021) argue that meaning in educational texts emerges not from language alone but from the orchestrated interplay of linguistic, visual, spatial and typographic resources. In medical education, this orchestration is indispensable rather than optional. Much of biomedical knowledge – structural, spatial, functional or procedural – cannot be adequately represented through verbal language alone. Anatomical diagrams, cross-sectional illustrations, radiographic images, procedural schematics and clinical infographics therefore carry substantial ideational load, enabling learners to perceive relations that written description tends to abstract (Lemke, 1998; O'Halloran, 2008).

Insights from cognitive psychology reinforce this multimodal necessity. Mayer's (2014) multimedia learning theory shows that comprehension improves when text and images are semantically aligned and spatially integrated, allowing learners to construct coherent mental models. Conversely, fragmented or poorly coordinated multimodal design may overload cognitive processing and hinder the assimilation of complex scientific concepts. This interplay between verbal and visual modes thus becomes central to pedagogical effectiveness in EMP textbooks.

Empirical research in science and medical education supports these claims. Analyses of educational materials in biology and health sciences demonstrate that images routinely perform representational functions (depicting entities and processes), interactive functions (positioning viewers through gaze, angle or social distance) and compositional functions (structuring information flow and salience) (Dimopoulos et al., 2003; Slough

et al., 2010; Makhuble et al., 2022).). These findings underscore that visuals are not ancillary ornaments but semiotic resources contributing directly to the construction of disciplinary meaning.

Although research in Vietnam remains comparatively limited, emerging studies point in the same direction. Hoang (2015) highlights the conceptual scaffolding role of visuals in Vietnamese ELT materials, while Le (2017) show that multimodal layouts shape learners' interpretive engagement and inferential strategies. Most pertinently, Lam and Pham (2024) demonstrate that visuals in Vietnamese medical English textbooks guide attention to clinically salient features, model procedural steps and visually convey institutional or professional authority. Yet, even these studies tend to privilege the intra-modal dimension, analysing images in isolation from their accompanying texts.

Taken together, this body of research reveals a persistent gap: while visuals in EMP textbooks have been recognised as pedagogically consequential, few studies systematically examine the intersemiotic complementarity (Royce, 2007) through which language and image jointly construct medical meaning. Addressing this gap requires an integrated multimodal analytical framework capable of capturing both the internal structure of visual resources and the dynamic relationships between visual and verbal modes – an endeavour undertaken in this thesis.

2.1.2.3 Pedagogical and Ideological Dimensions

Textbooks are not merely pedagogical tools; they are also ideological artefacts shaping learners' perceptions of professionalism, authority, and identity. Gray (2010) and Canagarajah (2013) argue that ELT materials frequently embody Western-centric epistemologies, privileging specific discourses and representational norms. In EMP, this ideological layer may be

even more pronounced due to the dominance of Anglophone biomedical traditions in global medical communication. Similar concerns are echoed in medical humanities scholarship, where scholars note that the visual culture of medicine often reproduces implicit hierarchies of race, gender, and clinical authority (Bleakley, 2015; Lupton, 2013).

Visual representations often naturalise implicit assumptions: recurrent depictions of particular skin tones, gendered professional hierarchies, technologically advanced clinical settings, or culturally specific patient – doctor interactions may subtly shape learners’ expectations of what constitutes “*authentic*” or “*modern*” medicine (Jewitt & Oyama, 2004). Such ideological encoding operates through both linguistic and visual modes, influencing learners’ emergent professional vision (Goodwin, 1994). This issue has also been raised in Vietnamese educational research, where imported textbooks in medicine and healthcare are observed to reflect Western institutional norms that do not always correspond to local medical realities (Hoang, 2015).

Thus, understanding EMP textbooks necessitates attention not only to pedagogical design but also to the ideological work they perform through multimodal representation.

2.1.2.4 Limitations in Existing EMP Textbook Research

Despite the centrality of textbooks in EMP instruction, existing research reveals notable gaps. International evaluations often prioritise vocabulary load, skills balance, or user perceptions (Hsu & Yang, 2007), while multimodal dimensions – particularly visual-verbal integration – remain understudied. Studies in multimodal science and health-education textbooks also suggest that visual design is frequently analysed descriptively rather than through a theoretically integrated lens (Dimopoulos et al., 2003; Slough et al., 2010).

In Vietnam, textbook analyses (Nguyen & Tran, 2023) similarly highlight issues of content relevance, authenticity, and alignment with clinical communication needs, yet they seldom address the semiotic organisation of pedagogic materials in a systematic manner. Complementary findings are reported by Tran and Lam (2025), whose evaluation of ESP materials for medical students reveals that most textbooks prioritise lexical coverage and reading comprehension while overlooking the multimodal configuration through which medical knowledge is visually constructed. Likewise, Mukundan (2011) identifies persistent challenges in using EMP textbooks in Vietnamese medical universities, including insufficient visual scaffolding and limited alignment between textbook representations and real-world clinical practices. More broadly, research in Vietnamese educational linguistics has shown that visuals in ELT and ESP materials tend to be under-analysed despite their significant pedagogical role (Hoang, 2015, Febrianti & Vu, 2023). These patterns collectively indicate that although textbook evaluation is not new in Vietnam, multimodal analysis – particularly visual-verbal integration – remains largely absent, reinforcing the need for a theoretically grounded multimodal framework.

Emerging multimodal studies (Lam & Pham, 2024) contribute valuable insights yet remain limited either in analytical scope (e.g., examining only representational meaning) or in multimodal integration (e.g., analysing images independently of text). These limitations underscore the need for a theoretically robust, multimodally integrated approach capable of examining how intra-modal structures (within images and language) and inter-modal relations (between image and language) collectively construct medical meaning – thereby justifying the development of the Integrated Multimodal Analytical Framework applied in this thesis.

2.1.3 Social Semiotics and Systemic Functional Linguistics

Social semiotics provides the overarching epistemological foundation for this study by conceptualising meaning not as a neutral transfer of information but as a socially organised and culturally motivated process. From a social semiotic perspective, sign-making involves patterned choices among semiotic resources that are shaped by institutional norms, disciplinary practices, and historical conditions. Meaning is therefore understood as contingent, interest-driven, and ideologically inflected: semiotic forms are selected and legitimised by communities to enact particular versions of knowledge, identity, and authority (van Leeuwen, 2005). This orientation is particularly relevant to professional domains such as medicine, where specialised knowledge is produced and circulated through highly regulated communicative practices, and where textual representations may carry consequential implications for decision-making, professional accountability, and risk management.

The intellectual origins of social semiotics are closely aligned with Halliday's conceptualisation of language as a "social semiotic," a view that constitutes the core premise of SFL (Halliday, 1978). In contrast to formalist traditions that treat language as an autonomous system, SFL theorises language as a meaning-making resource shaped by its functions in social life. Linguistic patterns are not arbitrary structural regularities; rather, they are motivated by the communicative work language performs within specific contexts. This shared commitment establishes a principled bridge between social semiotics and SFL: social semiotics foregrounds the social, cultural, and ideological embeddedness of meaning-making, while SFL offers a systematic model for analysing how such meanings are realised through structured choices in language - and, by extension, other semiotic modes.

A central concept in SFL is the metafunctional principle, which proposes that all meaning-making simultaneously realises three metafunctions: ideational, interpersonal, and textual (Halliday & Matthiessen, 2014). The ideational metafunction accounts for how experience is construed, including the representation of entities, processes, and logical relations. The interpersonal metafunction concerns the enactment of social relations, including stance, evaluation, authority, alignment, and degrees of certainty or obligation. The textual metafunction organises meanings into coherent messages, supporting information flow and cohesion through patterns such as thematic organisation and textual linking. Importantly, these metafunctions operate concurrently, enabling texts to function as multidimensional semiotic artefacts rather than as sequences of grammatical forms.

This metafunctional architecture is especially consequential in specialised discourse. In medical communication, ideational choices underpin biomedical categorisation and diagnostic reasoning, shaping how symptoms, conditions, and procedures are represented within established knowledge frameworks. Interpersonal resources regulate the management of expertise, sensitivity, and risk, influencing how certainty, responsibility, and ethical positioning are communicated. Textual resources, in turn, support the packaging and sequencing of complex information across genres such as clinical histories, case reports, diagnostic interpretations, and procedural protocols. Through this lens, SFL provides an explanatory framework for understanding how medical knowledge is linguistically constructed and socially enacted, and how professional authority is embedded within patterned textual practices.

Closely related to metafunctions is the systemic principle of meaning-as-choice-in-context. SFL models language as a network of interrelated

options, where texts are formed through selections from meaning potentials shaped by contextual demands. This “choice-based” view enables SFL to interpret recurrent linguistic tendencies as semiotic decisions carrying epistemic and institutional significance. For example, the prevalence of nominalisation in clinical and academic medical texts can be understood not simply as a stylistic feature but as a resource for construing processes and qualities as stable entities, aligning with medicine’s classificatory and objectifying epistemology. Likewise, passive voice and agent suppression frequently foreground procedures, observations, or outcomes over individual actors, reflecting professional orientations toward standardisation and procedural accountability. Such patterned choices illustrate how grammatical resources participate in the construction and legitimisation of disciplinary knowledge.

Crucially, SFL’s theoretical scope extends beyond language because metafunctions are formulated as functional principles rather than as exclusively linguistic properties. This abstraction provides a theoretical basis for multimodal expansions in which meaning across different modes can be analysed within a shared functional logic. The grammar of visual design, developed within multimodal social semiotics, systematically adapts Hallidayan metafunctions to the analysis of images, proposing that representational, interactive, and compositional meanings correspond broadly to ideational, interpersonal, and textual meanings in language (Kress & van Leeuwen, 2021). This alignment is essential for multimodal research on educational materials, including ESP/EMP textbooks, where meaning is distributed across written language, visual representations, layout, and typographic orchestration. Within such ensembles, diagrams, visual schemata, and image-text alignments may carry substantial epistemic weight,

functioning as tools for conceptual reasoning and professional enculturation rather than as decorative additions.

In addition to metafunctions, SFL offers a stratified architecture that links contextual configurations to linguistic realisation. A standard SFL account distinguishes the strata of context (culture and situation), discourse semantics, lexicogrammar, and expression (phonology/graphology). Stratification is analytically valuable because it provides a principled pathway from macro-level contextual variables to micro-level textual patterning without collapsing one into the other. Context of situation is commonly operationalised through register variables - field, tenor, and mode - which shape the probabilities of meaning selection and textual realisation. In medical education contexts, for instance, the field may involve diagnostic explanation or procedural instruction; tenor may reflect expert–novice roles (teacher-student, clinician-trainee); and mode may be multimodal, integrating technical language with images such as radiographs, anatomical diagrams, and workflow charts. These contextual configurations constrain and motivate semiotic choices across strata, allowing researchers to explain not only what meaning patterns occur but why they emerge as institutionally relevant and disciplinarily conventional.

In sum, social semiotics and SFL jointly provide a coherent theoretical foundation for this study. Social semiotics frames meaning-making as socially organised and ideologically inflected, highlighting how disciplinary communities legitimate knowledge through semiotic practices. SFL supplies a stratified and operationalisable model capable of analysing how contextual demands are realised through semantic organisation and lexicogrammatical choices. Because SFL's metafunctional architecture can be extended across modes, it also offers a principled bridge to multimodal social semiotics,

enabling systematic investigation of how verbal and visual resources in EMP textbooks co-function to construct medical knowledge and professional discourse. This integrated theoretical grounding provides the springboard for the multimodal expansion developed in the following section.

To clarify how SFL conceptualises meaning-making as a stratified and recursive process, Figure 2.1 presents a cyclic model showing how context activates meaning potential, which is realised through lexicogrammatical form and instantiated through expression.

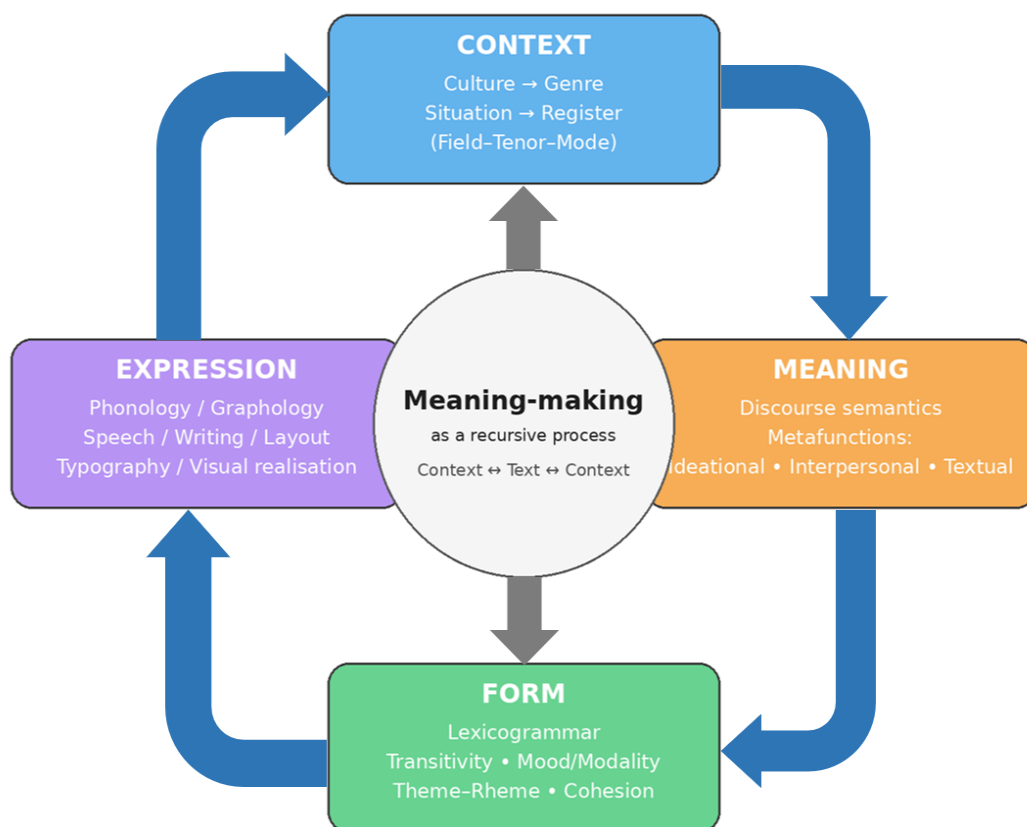


Figure 2.1 Cyclic Stratified Model of Meaning in SFL

This figure represents SFL meaning-making as a stratified and recursive process. Context (culture and situation) shapes discourse through genre and register variables (field, tenor, and mode), which activate the meaning potential of a text. Meanings are organised at the level of discourse

semantics through the ideational, interpersonal, and textual metafunctions, and are then realised through lexicogrammatical choices. These meanings are instantiated materially through expression in phonology or graphology, and the resulting texts circulate within communities, feeding back into and potentially reshaping contextual norms and communicative practices.

The foregoing discussion has shown that Social Semiotics conceptualises meaning as a socially constructed process realised through various semiotic resources. Since semiotic practices are embedded in particular social and cultural environments, meanings cannot be regarded as fixed or universally shared. Rather, they are shaped by the contexts in which semiotic resources are produced, circulated, and interpreted (Halliday, 1978; Hodge & Kress, 1988). This perspective is important for understanding visual communication because images, like language, function as socially organised systems of meaning whose interpretation is inherently context-dependent.

This view is consistent with the visual grammar developed by Kress and van Leeuwen (2021). Building on Halliday's social-semiotic theory of meaning, Kress and van Leeuwen conceptualise images as semiotic resources capable of realising representational, interactive, and compositional meanings. Importantly, they argue that semiotic resources are socially shaped and continuously transformed through changing social and cultural practices. Therefore, visual categories such as gaze, social distance, perspective, and viewer involvement should not be interpreted as universal codes with fixed meanings. Instead, their meanings need to be understood in relation to the social and cultural contexts in which visual communication takes place.

From a social-semiotic perspective, the meaning potential of semiotic modes is shaped by social practices and the contexts in which these modes are used (Hodge & Kress, 1988; van Leeuwen, 2005). Consequently, the

interpretation of visual meanings requires attention not only to the structural features of images but also to the broader contexts that influence how these semiotic resources are produced and interpreted. This contextual orientation is particularly important in educational discourse, where visual representations may be shaped by disciplinary conventions, pedagogical purposes, and institutional settings.

Accordingly, although the present study adopts the social-semiotic framework and visual grammar proposed by Kress and van Leeuwen (2021) as its primary analytical tools, it does not assume that visual meanings are universally interpreted. Instead, the analysis recognises the role of contextual and cultural factors in shaping the interpretation of gaze, social distance, perspective, and viewer positioning in English medical textbooks. Such a perspective provides a more context-sensitive understanding of how visual and verbal resources interact to construct meaning in multimodal educational discourse.

2.1.4 Multimodal Discourse Analysis

MDA builds upon the principles of SFL to investigate how meaning is realised through the orchestration of multiple semiotic modes. Developed over the past three decades in response to increasingly visual and digitally mediated communication, MDA extends the analytic focus beyond language to include image, typography, colour, gesture, gaze, sound, and spatial arrangements (Jewitt, 2014).

A key premise of MDA is that many contemporary discourses – especially in science, engineering, and healthcare – are inherently multimodal. Kress (2010) argues that multimodality is not a peripheral phenomenon but the normative condition of communication in modern societies. Meaning emerges from the interplay of modes, each offering distinct representational

affordances and constraints.

The most influential multimodal model is Kress and van Leeuwen's (2021) GVD, which systematically extends Halliday's metafunctions into the visual domain. They propose three metafunctionally aligned visual structures: Representational meaning (depicting participants, processes and circumstances); Interactive meaning (positioning viewers in relation to represented participants through gaze, angle and distance); Compositional meaning (integrating elements through information value, salience, framing and layout). This visual grammar has become the dominant framework for analysing images in educational and professional discourses due to its systematic, replicable categories.

Complementing GVD, O'Halloran (2008, 2011) advances Systemic Functional-Multimodal Discourse Analysis (SF-MDA), which models intersemiotic relations across modes. Her work demonstrates how linguistic, symbolic and visual resources operate synergistically in mathematical and scientific discourse. This is highly applicable to medical discourse, which similarly relies on specialised visualisation systems – anatomical diagrams, radiological images, biochemical pathways and procedural schematics – to construct and validate knowledge.

Jewitt (2014) further emphasises that multimodal meaning arises from modal configuration, the patterned orchestration of semiotic resources within a communicative event. This perspective aligns with Royce's (2007) construct of intersemiotic complementarity, which highlights how images and language collaboratively build coherent meaning.

These multimodal theories are especially relevant for medicine, a field fundamentally shaped by the integration of visual and verbal modes. Medical thinking often depends on image-based reasoning – reading X-rays,

interpreting MRI scans, analysing pathological slides – and these visual processes are inseparable from linguistic explanation (Unsworth, 2014, van Leeuwen, 2015). Thus, multimodality is not supplemental but constitutive of biomedical knowledge production.

In ESP and particularly in EMP, MDA provides an essential framework for analysing how textbooks mediate disciplinary literacy. Studies show that multimodal resources help learners understand spatial, causal and functional relations central to clinical reasoning. However, existing EMP research often focuses on text or image in isolation, overlooking the intersemiotic mechanisms through which meaning is jointly constructed. This methodological limitation underscores the need for integrated multimodal frameworks – precisely the purpose of combining GVD and Unsworth’s multimodal metafunctions in the present study.

In conclusion, MDA bridges SFL’s metafunctional theory with multimodal visual grammar and intersemiotic analysis, offering a comprehensive framework for examining how medical knowledge is constructed across modes. This theoretical synergy provides the conceptual foundation for the analytical frameworks discussed in Sections 2.3 and 2.4.

2.2 Grammar of Visual Design (Kress & van Leeuwen, 2021)

2.2.1 Representational Meaning

Kress and van Leeuwen’s Grammar of Visual Design (2021) constitutes one of the most authoritative frameworks in multimodal discourse analysis, offering a systematic account of how images realise meaning through socially and culturally patterned semiotic choices. Drawing on Halliday’s (1978) conceptualisation of language as a “*social semiotic*” they argue that visual communication, like verbal communication, is governed by a set of regularities that users draw upon to represent experience, enact

relations, and organise information. Importantly, these regularities are not prescriptive rules but socially sedimented conventions shaped by institutional practices such as those of medicine, science, and education.

Representational meaning, corresponding to Halliday's ideational metafunction, concerns how images depict participants, processes, and circumstances to construct experiential meaning (Kress & van Leeuwen, 2021). The framework distinguishes between narrative and conceptual structures. Narrative representations denote unfolding actions through vectors – visual lines of motion or gaze that connect participants and establish relations such as Actor and Goal. For instance, in an illustration of venipuncture, the needle forms a vector linking healthcare professional and patient, visually narrating the procedural sequence and rendering clinical action perceptually accessible.

By contrast, conceptual structures represent generalised, atemporal relations. Classificational structures map biomedical taxonomies (e.g., organ systems), analytical structures show part-whole relations (e.g., labelled cardiac chambers), and symbolic structures convey abstract attributes such as expertise or authority (van Leeuwen, 2008). Medical textbooks rely heavily on such conceptual diagrams – anatomical schematics, physiological maps, biochemical pathways – to organise epistemic knowledge in stable, recognisable formats.

Moreover, representational meaning carries ideological weight. Standardised anatomical images frequently depict an abstract, “*universal*” body that foregrounds biomedical neutrality while effacing demographic diversity (van Leeuwen, 2005; Kress, 2010). Critically, such choices normalise particular epistemic assumptions about whose body counts as the

medical default. Thus, analysing representational meaning requires attention not only to what is visually present but also to what is strategically omitted.

Pedagogically, representational structures play a central role in disciplinary learning. Through semantic expansion (Lemke, 1998), visuals complement linguistic explanation by offering perceptual instantiation of complex phenomena such as neuronal firing or blood circulation. This produces intersemiotic complementarity (Unsworth, 2008), wherein text and image collaboratively build conceptual understanding. Over time, these recurrent visual conventions cultivate learners' professional vision – the capacity to notice clinically salient features and interpret patterns according to disciplinary norms (Goodwin, 1994, Azkiyah et al., 2021). Consequently, representational meaning is not merely illustrative but fundamentally epistemological.

In this study, representational meaning is analysed to uncover how Vietnamese EMP textbooks visualise medical knowledge and position learners as emerging participants in biomedical epistemic practices.

2.2.2 Interactive Meaning

While representational meaning focuses on how images depict experiential content, interactive meaning addresses how images position viewers in relation to the represented participants and the image producer. This dimension aligns with Halliday's interpersonal metafunction, which governs how communicative acts negotiate roles, stance, and evaluative orientation (Halliday & Matthiessen, 2014). In the visual mode, interactive meaning is realised through four interrelated systems: gaze, social distance, angle, and modality (Kress & van Leeuwen, 2021).

Gaze distinguishes demand images – those in which represented participants look directly at the viewer – from offer images, which present

information for detached examination. Notably, medical textbooks overwhelmingly employ offer images, reinforcing the normative biomedical stance of clinical objectivity (Lemke, 1998). The rarity of demand images reveals an implicit disciplinary ideology privileging distance over engagement.

Social distance, conveyed through framings and shot scales, encodes degrees of intimacy. Close-ups of organs or tissues create cognitive proximity essential for diagnostic reasoning, whereas long shots of clinical environments maintain formality and institutional hierarchy (van Leeuwen, 2005). These spatial choices subtly shape how learners relate to clinical content.

Angle contributes to the construction of epistemic and social relations. High angles position the viewer as dominant, reflecting the clinical gaze that surveys and evaluates; low angles elevate participants, conveying authority; eye-level angles imply equality and empathy (Kress & van Leeuwen, 2021). Importantly, medical imagery frequently oscillates between the authoritative gaze of clinical expertise and the ethical gaze emphasised in patient care (Foucault, 1973).

Modality concerns truth-value. Paradoxically, in medicine, schematic diagrams – precise, simplified, decontextualised – often carry higher epistemic authority than photographic realism because abstraction enhances clarity and generalisability (van Leeuwen, 2008). Thus, modality choices reflect institutional preferences for idealised representations over literal depiction.

Taken together, these systems show that interactive meaning encodes a pedagogical and ideological stance: learners are positioned as observers being gradually socialised into biomedical ways of seeing. This is critical in

contexts such as Vietnam, where students engage with imported visual conventions that may implicitly shape professional dispositions.

In the present study, analysing interactive meaning illuminates how EMP textbooks position learners within disciplinary hierarchies of knowledge, authority, and clinical perception.

2.2.3 Compositional Meaning

Compositional meaning corresponds to Halliday's textual metafunction and governs how representational and interactive meanings are organised into coherent visual messages. According to Kress and van Leeuwen (2021), composition is structured along three core systems: information value, salience, and framing.

Information Value assigns meaning to spatial zones such as Given-New (left/right), Ideal-Real (top/bottom), and Centre-Margin. These culturally sedimented arrangements guide interpretation. For example, placing conceptual overviews at the top (Ideal) and detailed labels below (Real) mirrors the pedagogical movement from abstract principle to concrete detail. Similarly, left-right contrasts organise diagnostic comparisons – for instance, contrasting normal and pathological conditions.

Salience orchestrates attention and cognitive load through contrast, colour, size, sharpness, and foregrounding. In medical textbooks, salient features often highlight clinically crucial structures, directing learners to the most relevant anatomical or pathological information (Sweller, 1994). Critically, salience also reflects disciplinary priorities: what is made visually prominent often mirrors what the discipline deems epistemically important (van Leeuwen, 2008).

Framing regulates connectedness or separation among elements through borders, spacing, and connectors. Strong framing promotes analytical

decomposition – useful for studying anatomical segments – whereas integrated framing fosters holistic understanding, essential for grasping systems-level physiology (Unsworth, 2008).

Together, these compositional resources create the visual architecture through which learners interpret and hierarchise information. Importantly, composition does not merely organise content; it encodes epistemological assumptions about how medical knowledge should be structured, sequenced, and cognitively navigated.

Integrated Perspective

Across its representational, interactive, and compositional dimensions, GVD reveals that visual design is inherently epistemic, ideological, and pedagogic. Images in medical textbooks are not decorative supplements; they are semiotic engines that organise perception, legitimise disciplinary hierarchies, and cultivate professional habits of seeing.

For this thesis, GVD forms a key analytical layer within the Integrated Multimodal Analytical Framework, enabling a rigorous analysis of intra-modal visual structures before examining how these interface with linguistic modes to construct medical meaning in Vietnamese EMP textbooks.

2.3 Multimodal Metafunctions: Extending Halliday through Unsworth

2.3.1 Multimodal Metafunctions

The multimodal metafunctional framework developed by Unsworth (2008) represents one of the most significant expansions of SFL into the domain of multimodal analysis. While Halliday (1978, 1994) originally theorised ideational, interpersonal and textual metafunctions as inherent to language, Unsworth reconceptualises them as *transmodal organising principles* that operate across linguistic and visual systems. This shift marks

an epistemologically consequential development: meaning-making in contemporary texts is no longer understood as language-led, but as an inherently multimodal process in which distinct semiotic resources collaborate to construct disciplinary knowledge.

Importantly, Unsworth's work fills a theoretical gap between linguistic functionalism and visual social semiotics. Kress and van Leeuwen (2021) had previously demonstrated how images realise representational, interactive and compositional meaning. However, the relationship between these visual meanings and Hallidayan metafunctions was only partially articulated. Unsworth (2008) addresses this gap by arguing that *multimodal metafunctions* provide the conceptual foundation through which we can systematically analyse how linguistic and visual modes jointly construct experience, negotiate stance and build textual coherence.

This reconceptualisation is particularly relevant for education and specialised discourse. In domains such as medicine – where knowledge is inherently visual-verbal metafunctions are not parallel systems operating independently; rather, they are *integrated semiotic principles* shaping how learners interpret complex phenomena. Accordingly, the present study adopts Unsworth's multimodal metafunctional model as a core lens for examining EMP textbooks in Vietnam.

Ideational Meaning: Construing Experience Across Modalities

The ideational metafunction concerns how experience is represented and how logical relations among entities, processes and circumstances are construed. Extending Halliday's transitivity framework, Unsworth (2006) demonstrates that images – like clauses – encode processes (action, reaction, classification), participants (actors, goals, carriers) and circumstances (location, manner, cause). This multimodal transitivity is not merely a visual

analogue to language; it constitutes a complementary and, at times, epistemically generative system.

In EMP textbooks, ideational meaning typically arises through *intersemiotic redundancy and expansion* (Royce, 2002). Linguistic text may describe a cardiac conduction process sequentially, while the accompanying diagram spatialises this process, making temporal and functional relations perceptually accessible. Such visualisations are not decorative; they restructure biomedical knowledge into cognitively tractable forms – what Lemke (1998) describes as *semantic amplification*, where images add meaning potentials not easily captured lexically.

Unsworth (2008) identifies three main forms of ideational interaction:

- **Elaboration**, where visuals restate or clarify textual meaning.
- **Extension**, where images add new experiential details – e.g., anatomical orientation.
- **Enhancement**, where visuals depict causal or functional relations more explicitly than text.

These processes are essential in medical education, where complex, often invisible physiological mechanisms must be conceptualised visually to be understood. Thus, ideational meaning in multimodal medical discourse performs both *representational* and *epistemological* work, shaping how learners internalise clinical knowledge.

Interpersonal Meaning: Negotiating Stance and Professional Vision

The interpersonal metafunction concerns how social relations, evaluation and stance are enacted. In multimodal contexts, Unsworth (2008) argues that interpersonal meaning emerges through a dialogic interplay between linguistic modality and visual perspective.

Linguistically, medical textbooks generally adopt an impersonal, authoritative tone – nominalisation, passive constructions and technical lexis serve to foreground biomedical processes rather than agents (Hyland, 2005). Visually, however, interpersonal meaning is realised through gaze, distance, angle and modality (Kress & van Leeuwen, 2021).

Medical images tend to employ what Kress and van Leeuwen call offer structures, presenting information for detached observation rather than interpersonal engagement. This visual detachment aligns with the epistemic stance of medicine, wherein objective observation and clinical neutrality are privileged (Lemke, 1998). In this sense, visual design reinforces the disciplinary habitus of “*seeing clinically*” – privileging anatomical clarity over affective engagement.

Nevertheless, interpersonal meaning occasionally shifts. Images depicting patient-physician interactions may use eye-level angles or moderate distance to foster empathy, aligning with the broader pedagogical goal of cultivating patient-centred professionalism. Hence, interpersonal meaning in EMP textbooks is multifaceted: it trains learners to adopt the *clinical gaze* (Foucault, 1973) while also modelling the interpersonal norms of bedside communication.

Textual Meaning: Achieving Coherence Across Semiotic Modes

The textual metafunction concerns how meanings are organised into coherent messages. Unsworth (2006) reframes this as *intersemiotic texture*, arguing that coherence in multimodal texts arises from alignment between linguistic cohesion (reference, thematic structure, lexical relations) and visual compositional systems (information value, salience, framing).

Textual meaning emerges through several interactional mechanisms:

- **Lexical-visual alignment**, where terminological choices echo visual labels.
- **Sequential coherence**, where stepwise diagrams align with procedural clauses.
- **Thematic coupling**, where visual placement reinforces linguistic thematic progression.

This multimodal texture is particularly pronounced in medical materials. Procedural sequences such as interpreting ECG waves or describing steps of a clinical manoeuvre – require precise textual-visual harmonisation so that learners can integrate temporal and spatial information. When alignment is strong, multimodal coherence increases comprehension; when misaligned, it produces cognitive dissonance.

Therefore, textual metafunctions are central to how EMP textbooks scaffold disciplinary reasoning. They provide the structural backbone through which multimodal ensembles form interpretable, pedagogically meaningful wholes.

2.3.2 Intersemiosis and Image-Text Relations

Intersemiosis as the Core of Multimodal Meaning

Intersemiosis refers to the process by which different semiotic modes interact to create integrated meaning. Whereas metafunctions operate within and across modes, intersemiosis examines the *relationships* between modes – how meanings are projected, expanded, reinforced or transformed (Unsworth, 2008; Royce, 2007).

In multimodal medical discourse, intersemiosis is not optional; it is the *mechanism* through which biomedical knowledge is constructed. Diagrams, captions, tables, formulae and explanatory paragraphs work together to map

complex physiological systems. Thus, intersemiosis constitutes the epistemological foundation of medical literacy.

Typologies of Image-Text Relations

Unsworth (2008) identifies three primary relations:

1. **Complementarity** – modes combine to form a unified meaning (e.g., cardiac cycle diagram + functional description).
2. **Elaboration** – one mode restates or specifies meaning in the other.
3. **Extension** – one mode adds new experiential meaning not present in the other.

Martinec and Salway (2005) further refine these relations using two dimensions:

- **status** (equal vs. unequal prominence),
- **logico-semantic relation** (elaboration, extension, enhancement).

This nuanced taxonomy allows precise analysis of semiotic hierarchy within medical textbooks. For instance, a labelled anatomical diagram typically holds primary status, while the caption elaborates or enhances its meaning.

Mechanisms of Intersemiotic Complementarity

Three metafunctional alignments underpin intersemiosis:

- **Ideational alignment:** text explains, image spatialises; together they represent biomedical processes.
- **Interpersonal alignment:** linguistic modality and visual stance jointly negotiate authority and engagement.
- **Textual alignment:** cohesive ties and visual composition work together to produce interpretive pathways.

Mayer & Fiorella's (2022) multimedia learning theory empirically supports this, showing that spatial and temporal contiguity between modes

enhances comprehension. In medical education, this principle is indispensable: learners must synchronise visual and verbal information to build coherent mental models of anatomical and physiological relationships.

Relevance to English for Medical Purposes

The multimodal metafunctional and intersemiotic framework is essential for analysing EMP textbooks for three reasons.

First, ideational intersemiosis reveals how biomedical knowledge is recontextualised for learners. In the Vietnamese context, where students often rely heavily on visuals to comprehend complex concepts, the alignment between diagrams and explanations determines whether learning is conceptually accessible.

Second, interpersonal meanings shape professional enculturation. Visual stance and linguistic authority jointly construct the epistemic norms of clinical reasoning. Understanding this helps clarify how textbooks implicitly train learners to adopt professional ways of seeing and acting.

Third, textual meanings expose how multimodal coherence is – or is not achieved. Given that many locally used EMP materials integrate imported visuals with locally written texts, compositional and cohesive inconsistencies can undermine disciplinary literacy development.

Taken together, Unsworth's multimodal metafunctions and theories of intersemiosis provide a robust analytical foundation for the present study. They enable a detailed, principled examination of how linguistic and visual modes co-construct medical meaning and how this multimodal ensemble shapes learners' emergent professional identities in the Vietnamese EMP context.

2.4 Previous Research

2.4.1 *Global Developments in Research on Visual and Multimodal Texts*

Research on educational discourse initially privileged linguistic forms, treating written text as the primary vehicle for conveying disciplinary knowledge. Early work in applied linguistics and genre studies, such as Swales (1990) and Hyland (2006), predominantly examined rhetorical organisation, lexical bundles, and disciplinary conventions, while visual elements were seen as marginal or merely decorative. However, a major intellectual shift occurred with the rise of social semiotics (Hodge & Kress, 1988) and Systemic Functional Linguistics (Halliday, 2004), which reconceptualised communication as inherently multimodal. This shift disrupted linguistic monomodality by positioning images, layout, typography, and spatial design as equally meaningful semiotic resources.

Kress and van Leeuwen's *Reading Images* (2021) solidified this transformation by proposing a “*visual grammar*” parallel to linguistic systems. Their representational, interactive, and compositional metafunctions offered scholars a systematic lens for decoding visual meaning, highlighting that images participate in knowledge construction rather than merely supporting text. Lemke (1998) extended this, demonstrating that scientific discourse inherently distributes meaning across linguistic, mathematical, and graphical resources. In his formulation, disciplinary knowledge is co-constructed across modes.

As multimodality gained prominence, scholars emphasized that educational materials – especially those in Science – Technology – Engineering – Mathematics and health sciences – cannot be understood without analysing visual structures, spatial arrangements, and intersemiotic

relations (Bezemer & Kress, 2016). This reconceptualization marks a turning point: visuals are now recognized not as passive reflections of textual meaning but as autonomous, epistemically significant contributors that shape how learners conceptualize scientific phenomena, interpret evidence, and build schematic reasoning.

Thus, global research has moved from linguistic centrality toward a multimodal paradigm, establishing the theoretical foundation for contemporary analyses of medical and ESP textbooks.

2.4.2 Empirical Findings from Multimodal Educational Research

Empirical studies in science and technical education have repeatedly confirmed the pedagogical and epistemic significance of visual resources. Mayer's (2014) Cognitive Theory of Multimedia Learning demonstrates that learning improves when visual and verbal information are presented in coordinated, contiguous forms. His dual-channel and coherence principles explain why diagram-text integration supports conceptual retention and reduces cognitive load – insights directly relevant to medical pedagogy.

From a linguistic-semiotic standpoint, Royce (2007) formalised intersemiotic complementarity, proposing that verbal and visual elements collaborate to create cohesive multimodal meanings. Unsworth (2008) extended Hallidayan metafunctions to multimodal settings, showing how ideational, interpersonal, and textual meanings materialize through the interplay of images and language.

Empirical work further demonstrates that scientific diagrams argue rather than merely illustrate. Lemke (1998) showed that graphs and anatomical visuals encode causal logic, classification hierarchies, and spatial reasoning central to disciplinary thinking. Studies in mathematics and science education (O'Halloran, 2008; Lim et al., 2012) reveal that meaning emerges through

intersemiotic synergy, whereby symbolic, graphical, and linguistic resources interact to articulate abstract relationships.

Textbook-focused research echoes these findings. Bezemer and Kress (2016) treat textbooks as designed multimodal ensembles shaped by technological, institutional, and curricular constraints. Their work demonstrates that students rely on visuals to interpret complex processes, and that poorly designed visual-verbal integration leads to misconceptions, cognitive overload, or loss of coherence.

Recent multimodal analyses in health sciences (e.g., Rowley-Jolivet, 2002; Bateman, 2014) underscore that diagrams, clinical photographs, and radiological images encode epistemic claims about medical phenomena and diagnostic reasoning. Importantly, such visuals guide perception by directing gaze, structuring attention, and highlighting clinically salient features.

Together, these empirical findings establish that multimodal design is not an aesthetic accessory but a central determinant of how learners conceptualise, internalise, and operationalise disciplinary knowledge – an insight critical for analysing ESP and EMP textbooks.

2.4.3 ESP/EMP-related Multimodal Studies

Within ESP, research has increasingly acknowledged the multimodal nature of disciplinary communication. Hyland (2006) emphasizes that ESP genres – laboratory reports, case presentations, research articles – exhibit multimodal density, integrating visuals, numerical data, specialised terminology, and disciplinary reasoning. This density renders visual meaning-making particularly crucial in ESP pedagogy.

In EMP, visuals play an especially central role due to medicine's reliance on anatomical diagrams, disease progression charts, flow diagrams, imaging modalities, and procedural schematics. Studies by Rowley-Jolivet (2002) and

Carney & Levin (2002) revealed that visuals in scientific and medical discourse serve argumentative and evidential functions, modelling causal, temporal, and classificatory relations. Unsworth & Cléirigh (2009) demonstrated how multimodal configurations in medical texts create epistemic coherence, enabling learners to conceptualize dynamic biological systems.

Textbook-oriented EMP research further reinforces this point. Studies in science and medical education consistently show that effective diagram-text integration enhances comprehension and supports the construction of disciplinary schemas. For example, Tang, Delgado, and Moje (2014) demonstrated that coordinated verbal-visual representations in science textbooks significantly improve learners' conceptual understanding, particularly when visuals elaborate processes that language alone cannot fully convey. Likewise, Slough et al. (2010) found that well-designed multimodal configurations help students form coherent mental models of scientific phenomena, confirming that meaning emerges through intersemiotic coordination rather than simple juxtaposition. In the context of textbook design, Bateman (2014) further argues that analysing image-text relations reveals how instructional materials encode disciplinary reasoning, procedural logic, and hierarchical structures of knowledge. Collectively, these studies underscore that in EMP – as in other scientific domains – conceptual clarity is contingent on the cohesive integration of linguistic and visual semiotic resources.

Critically, scholars such as Jewitt and Oyama (2004) underscore that visuals in medical discourse also carry ideological weight, representing bodies, expertise, and clinical authority in culturally situated ways. This raises important questions about how medical textbooks shape learners' perceptions

of professionalism, patienthood, and medical norms – particularly when materials are imported into non-Western contexts.

Overall, research in ESP/EMP highlights the necessity of integrating multimodal analysis into textbook evaluation and curriculum design.

2.4.4 Vietnam-based Research on Multimodality and ESP/EMP Textbooks

In Vietnam, interest in multimodal discourse has increased, but scholarship remains fragmented and underdeveloped, particularly regarding ESP and EMP materials. Most studies address linguistic challenges, needs analysis, or curriculum alignment rather than multimodal meaning-making. For example, Ton (2017) evaluated ESP materials for medical students, highlighting issues of authenticity and insufficient alignment with clinical communication. Similarly, Le (2017) identified challenges in using EMP textbooks, including limited contextual relevance and a lack of pedagogical scaffolding.

Although these studies offer valuable insights, they tend to approach visuals only implicitly or at a surface level. Research in general English education reveals a similar pattern. Hoang (2015), in his analysis of the development of Vietnam's national English textbook series, notes that while images are systematically included, their pedagogical functions are rarely theorised beyond illustration. Extending this critique, Pham (2022) demonstrates that visuals in Vietnamese EFL textbooks frequently operate within limited rhetorical roles, with insufficient attention to their semiotic potential in constructing meaning. Collectively, these findings indicate that although visuals are pervasive in Vietnamese educational materials, their analytical treatment remains underdeveloped, leaving the deeper multimodal, epistemic, and pedagogical dimensions largely unexplored.

Multimodal analyses specifically focusing on medical English textbooks are still rare but emergent. Lam and Pham (2024), for instance, conducted a preliminary multimodal analysis of a Vietnamese medical textbook, examining representational structures and noting the dominance of biomedical visual conventions. Their findings underscore that visuals in EMP materials guide attention to clinically salient features but often inadequately integrate with textual explanations.

Collectively, Vietnamese scholarship indicates a gap: while textbook quality is frequently discussed, systematic multimodal analyses – particularly those grounded in Kress and van Leeuwen’s visual grammar or Unsworth’s multimodal metafunctions – are lacking. This gap justifies the present study’s focus on the multimodal construction of medical knowledge in Vietnamese EMP textbooks.

2.4.5 Synthesis and Implications

Across global, ESP/EMP, and Vietnam-based research, three conclusions emerge.

First, the maturation of multimodal theory demonstrates that textbook visuals constitute a core epistemic resource rather than supplemental illustration. Global studies provide strong evidence that scientific and medical understanding depends on multimodal orchestration.

Second, empirical findings reveal that meaning in educational texts emerges from intersemiotic integration – how images, text, symbols, and layout work together. Yet many analyses, especially in ESP/EMP, remain monomodal or descriptive, without examining multimodal coherence.

Third, Vietnamese research, while growing, has not yet adopted systematic multimodal approaches. Existing EMP textbook evaluations identify surface-level pedagogical issues but rarely analyse representational, interactive,

compositional, or intersemiotic structures. This creates a methodological and contextual gap.

Accordingly, the present study responds to all three gaps by integrating (i) Kress and van Leeuwen's Grammar of Visual Design and (ii) Unsworth's multimodal metafunctional model to analyse how Vietnamese EMP textbooks construct ideational, interpersonal, and textual meanings across modes. By doing so, the study contributes to global scholarship, advances ESP/EMP research, and fills a critical knowledge gap in Vietnam's medical English education.

2.5 Constructing a Social-Semiotic Multimodal Framework: A Systemic Integration of Visual Grammar and Multimodal Metafunctions

The present study adopts a social-semiotic view of meaning in which all semiotic modes are understood as culturally and institutionally shaped resources for making meaning rather than neutral channels of transmission (Halliday, 1978; Kress, 2010). Within this paradigm, any analytical framework used to interrogate English medical textbooks must be both systemic – capable of accounting for patterned choice in a regular, rule-governed way and multimodal – able to trace how linguistic and visual resources work together to construct disciplinary knowledge. This section therefore argues for a critical, systemic integration of Kress and van Leeuwen's GVD with Unsworth's multimodal metafunctional model as the basis for the IMAF employed in this thesis.

2.5.1 Shared social-semiotic foundations

Both GVD and Unsworth's model are Hallidayan-inspired social-semiotic extensions, drawing on SFL metafunctions to theorise meaning across modes. Halliday's (1978) conception of language as a “*social semiotic*” and his metafunctional model – ideational, interpersonal and textual

– provide a general theory of how semiotic systems simultaneously construe experience, enact social relations and organise messages (Halliday & Matthiessen, 2014). Kress and van Leeuwen (2021) explicitly extend these metafunctions into the visual domain, arguing that images realise representational, interactive and compositional meanings which are functionally parallel to Halliday’s metafunctions.

Unsworth (2008), working from the same theoretical tradition, proposes that metafunctions should be understood as transmodal principles rather than properties of language alone. He reconceptualises them as multimodal metafunctions, showing how ideational, interpersonal and textual meanings can be mapped across language and image in school science and literacy texts. Crucially, both frameworks treat multimodality not as an add-on but as a necessary consequence of social-semiotic theory: if meaning is functional and contextual, semiotic modes participating in social communication can be systematically analysed in metafunctional terms. (Jewitt, 2014; van Leeuwen, 2005).

This common philosophical ground legitimises their combination. It ensures that integrating visual grammar and multimodal metafunctions is not an eclectic gesture but a theoretically coherent move within systemic functional social semiotics.

2.5.2 Systemic complementarity: from intra-modal to inter-modal analysis

Despite this shared foundation, the two frameworks operate at different but complementary levels of abstraction. GVD provides a detailed intra-modal description of the visual mode. It offers fine-grained categories such as narrative versus conceptual structures, vectors, gaze, social distance, information value, salience and framing – that make it possible to describe

how particular images are organised and how they foreground specific aspects of experience (Kress & van Leeuwen, 2021). In the context of EMP textbooks, these tools are indispensable for describing how anatomical diagrams, procedural illustrations or clinical photographs are constructed as visual artefacts.

However, GVD is less explicit about how visual structures combine with language to form multimodal wholes. It can show, for instance, that an illustration represents a dynamic physiological process through narrative structure, but it says relatively little about how that depiction interacts with the accompanying paragraph, caption or activity instructions. As O'Halloran (2008) notes, a major challenge in multimodal analysis lies in articulating relationships between semiotic strata, not merely describing each mode separately.

Unsworth's multimodal metafunctional framework addresses precisely this issue. Rather than providing a fine-grained "*grammar of image*", it models inter-modal relations through constructs such as intersemiosis, intersemiotic complementarity and intersemiotic texture (Unsworth, 2006, 2008). Ideationally, it accounts for how language and image jointly construe processes, participants and circumstances; interpersonally, it tracks alignment between evaluative stance and visual positioning; textually, it explains how cohesion and coherence are distributed across modes (Royce, 2007).

Critically, then, the two frameworks are not alternatives but systemic complements. GVD excels at describing how images mean on their own terms; Unsworth's model explains how images and language mean together in context. Taken separately, each leaves an analytical gap: GVD risks remaining intra-modal and descriptive, whereas Unsworth's metafunctional mapping can become too abstract if not anchored in a concrete visual

grammar. A social-semiotic multimodal framework for EMP textbooks therefore needs to mobilise both: GVD to analyse visual structure, and Unsworth to explain multimodal orchestration.

2.5.3 Critical integration for analysing English medical textbooks

The need for such an integrated framework is particularly acute in the analysis of English medical textbooks used in Vietnam. Empirical work on ESP/EMP materials has shown that these textbooks are central sites of disciplinary enculturation, yet existing evaluations tend to foreground linguistic coverage and skills balance, giving only limited attention to multimodal design (Ton & Nguyen, 2020). Emerging multimodal research on Vietnamese medical textbooks suggests that visuals carry substantial pedagogic and epistemic load, but these studies often focus either on isolated visual categories or on broad textual analysis without a systematic account of image-text relations (Lam & Pham, 2024).

Against this backdrop, a critical integration of GVD and Unsworth's framework offers several advantages. First, at the visual stratum, GVD allows the present study to classify the representational, interactive and compositional patterns through which medical images are deployed (RQ1.1). It makes it possible, for example, to distinguish between diagrams that function as conceptual taxonomies of body systems and images that narrate diagnostic or procedural sequences. It also enables scrutiny of interactive choices such as gaze, angle and social distance – that position learners as detached observers or empathetic participants, thus shaping professional vision and identity (Goodwin, 1994; Kress, 2010).

Second, at the intersemiotic stratum, Unsworth's metafunctional model provides the conceptual machinery for addressing overarching RQ and Sub-RQs: how visual and linguistic modes interact to construct medical meaning

and disciplinary knowledge. By analysing ideational alignment (e.g., whether diagrams elaborate, extend or exemplify textual explanation), interpersonal coordination (e.g., how textual modality and visual perspective jointly encode certainty, authority or care) and textual coherence (e.g., how captions, labels and layout contribute to intersemiotic texture), the framework explains not only what images and texts do separately but how their combination becomes pedagogically meaningful (Unsworth, 2008; Mayer, 2014).

Third, and importantly, this integration supports a critical reading of how imported EMP textbooks mediate global biomedical epistemologies in local Vietnamese classrooms. Visual and verbal design jointly encode assumptions about what counts as legitimate medical knowledge, who is authorised to speak, and how patients and professionals should appear (Bezemer & Kress, 2016). The combined framework makes it possible to interrogate, for instance, whether diagrams privilege a universalised, decontextualised body, or whether image–text pairings leave room for contextualised, patient-centred understandings relevant to Vietnamese healthcare settings.

2.5.4 A stratified social-semiotic model: from modes to multimodal ensembles

To translate this critical integration into an operational analytical system, the IMAF developed in this thesis is conceptualised as a stratified social-semiotic model comprising three mutually dependent yet analytically distinguishable layers. This stratification reflects the systemic organisation of meaning in Hallidayan social semiotics, where semiosis is viewed as distributed, layered, and contextually patterned rather than confined within isolated modes.

First, the visual stratum is examined through the categories of Kress and van Leeuwen's (2021) GVD – representational, interactive, and compositional meanings. This stratum captures how individual images function as structured social-semiotic artefacts, revealing how medical phenomena are depicted (e.g., through narrative or analytical structures), how visual relations with viewers are enacted, and how visual elements are arranged to construct hierarchy, coherence, and emphasis. In medical textbooks, this enables fine-grained analysis of how diagrams, anatomical illustrations, and clinical photographs materialise biomedical knowledge and encode disciplinary expectations.

Second, the linguistic stratum is analysed through selected SFL tools – specifically transitivity, mood/modality (evaluation), and thematic structure. This layer uncovers how verbal language construes processes, participants, evaluative stance, and information flow, thereby illuminating how medical discourse enacts authoritative voice, clinical objectivity, or instructional guidance. The linguistic stratum thus reveals the functional semantics through which specialist knowledge is verbalised and pedagogically sequenced.

Third, and crucially, While Unsworth (2008) provides the overarching intersemiotic logic (multimodal metafunctions and intersemiotic texture), Royce's (2007) intersemiotic complementarity offers operational descriptors for mapping image-text relations at metafunctional levels, strengthening the analytical tractability of the intersemiotic stratum. This stratum enables the analysis of how diagrams enrich textual explanations, how captions mediate viewer interpretation, and how layout orchestrates verbal-visual synergy on the textbook page. It provides the analytical means to understand multimodal cohesion and the epistemic labour performed by the interaction between words and images.

Taken together, these strata represent a systemic model of multimodal semiosis, aligning with Bateman (2014) and O'Halloran's (2011) assertion that multimodal meaning emerges across interlocking semiotic layers rather than within discrete modes. This system allows analytical movement from micro-level observations (e.g., how salience foregrounds a cardiac structure), through meso-level multimodal configurations within units or lessons (e.g., how anatomical diagrams and textual explanations cohere), to macro-level pedagogical interpretations (e.g., how these multimodal ensembles support or limit the development of disciplinary literacy in medical students).

Importantly, the integration is conducted with a critical stance. Kress and van Leeuwen's GVD has been critiqued for its Eurocentric assumptions and for implying greater stability in visual conventions than may actually exist in diverse cultural contexts (Machin, 2013). Similarly, Unsworth's metafunctional framework – originally formulated within Anglophone primary and secondary schooling – may not fully capture the epistemic density and abstraction characteristic of higher-education medical discourse. These limitations matter significantly in the Vietnamese EMP context, where imported textbooks mediate Western biomedical knowledge within local pedagogical traditions.

Therefore, in this thesis, both frameworks are employed not as prescriptive or universal systems but as heuristic, adaptable analytical tools. Where empirical data reveal frictions such as hybrid diagrammatic genres, culturally inflected representational choices, or multimodal sequences that exceed existing taxonomies – these tensions are treated as theoretical opportunities rather than methodological anomalies. Such divergences are foregrounded explicitly, enabling the study to contribute to ongoing refinements of multimodal analysis in specialised educational discourse.

Collectively, this stratified approach enables the thesis to address its overarching research question: how linguistic and visual modes interact to construct medical meaning in English medical textbooks used in Vietnam. By systematically articulating the transitions from intra-modal structure to inter-modal orchestration, the IMAF aligns analytical granularity with epistemic purpose, offering a robust basis for examining how multimodal resources construct disciplinary knowledge, professional reasoning, and pedagogical intent in the textbooks under study.

This systemic, socially grounded multimodal framework consolidates the analytical foundations for examining how English medical textbooks mobilise visual and linguistic resources to construct disciplinary meaning. By synthesising the descriptive power of GVD with the functional explanatory capacity of multimodal metafunctions, the framework provides a coherent pathway for addressing RQ1 and its sub-questions. Importantly, the review of theoretical debates also highlights that such an integrated model has not yet been applied to the Vietnamese EMP context. Section 2.7 therefore identifies the specific conceptual, methodological and contextual gaps that remain unaddressed and presents the IMAF underpinning the empirical analysis of this thesis.

2.6 Research Synthesis

The preceding review has shown that research on EMP, multimodal discourse, and visual design has developed into a rich and diverse field, spanning linguistic, cognitive, social-semiotic, and educational perspectives. However, the juxtaposition of these literatures also reveals critical inconsistencies, under-examined intersections, and methodological limitations – particularly in the Vietnamese context – thereby justifying the necessity of the present study. This section synthesises these bodies of work, identifies the

research gaps that emerge from their intersections, and establishes a conceptually integrated framework that guides the empirical analysis.

2.6.1 Research Gap

A comprehensive comparison across international, ESP/EMP-specific, multimodal, and Vietnam-based research reveals four major gaps that remain unaddressed despite the abundance of literature.

2.6.1.1 Theoretical Gap: Absence of a Fully Integrated Multimodal Analytical Model

Although the global literature on multimodal discourse has produced several influential frameworks, these approaches remain theoretically fragmented, each addressing only one dimension of how multimodal meaning is constructed. The GVD (Kress & van Leeuwen, 2021) offers a highly elaborated account of visual representation, interaction, and composition, yet its analytical power is largely confined to intra-modal structures within the visual mode. Conversely, Unsworth's multimodal metafunctions (2008), together with frameworks of intersemiosis developed by Royce (2007) and Martinec and Salway (2005), illuminate the inter-modal relations through which images and language collaborate to produce cohesive meaning, but they provide limited tools for describing the internal organisation of visual resources themselves.

This disjunction creates a significant theoretical gap: despite the richness of the existing landscape, no study has yet integrated these complementary traditions into a unified, stratified analytical model that can simultaneously account for (i) the internal visual grammar of images, (ii) the linguistic structures of accompanying text, and (iii) the orchestration of meaning across modes. Such an integrated framework is essential for capturing the full complexity of multimodal semiosis.

The absence of a consolidated model is particularly consequential in the field of medicine, where knowledge is inherently and irreducibly multimodal (Lemke, 1998; Kress, 2000, Unsworth, 2001, Jewitt, 2009a). Medical understanding routinely emerges through the interplay of verbal explanation, anatomical schematics, procedural diagrams, spatially organised labels, and representational conventions that cannot be reduced to language alone. Without a framework capable of analysing both intra-modal structure and inter-modal synergy, research cannot adequately account for how medical concepts are visualised, sequenced, and pedagogically communicated. It is this precise theoretical gap – at the intersection of visual grammar, linguistic metafunctions, and intersemiotic relations – that the present study seeks to address.

2.6.1.2 Methodological Gap: Fragmented Analyses of EMP Textbooks Internationally

Despite substantial international scholarship on multimodal educational materials, methodological fragmentation remains a persistent issue in the analysis of EMP textbooks. Studies by Bezemer and Kress (2016), O'Halloran (2008), and Tang et al. (2014) have convincingly demonstrated that educational discourse is inherently multimodal, yet most EMP-focused analyses continue to approach textbooks through single-mode or loosely multimodal lenses. Linguistically oriented studies typically foreground vocabulary load, genre structures, or skill distribution, while visually oriented research often classifies images by type or representational function. Other strands adopt cognitive perspectives, examining diagram comprehension or knowledge retention. However, these distinct strands rarely converge into a systemic, social-semiotic methodology capable of analysing how linguistic and visual modes co-realise meaning.

Moreover, research explicitly targeting medical imagery such as Dimopoulos et al. (2003) and Slough et al. (2010) – tends to describe visual forms rather than interrogating how images interact with accompanying text to construct biomedical reasoning, clinical logic, or epistemic authority. Critically, these studies seldom trace the ideational, interpersonal, and textual intersemiosis through which multimodal resources jointly produce disciplinary knowledge. This absence is consequential: as medical discourse fundamentally depends on the coordinated mobilisation of diagrams, labels, explanatory paragraphs, spatial layout, and textbook page architecture, descriptive accounts of isolated images cannot reveal how learners are socialised into professional ways of seeing and knowing.

Importantly, this methodological fragmentation in the global literature parallels gaps observed in the Vietnamese context. As reviewed earlier, domestic studies on EMP materials (e.g., Bui, 2016; Le, 2017; Lam & Pham, 2024) often prioritise linguistic alignment or content relevance, while multimodal investigations – though emerging – remain limited in scope and analytical depth. Together, these patterns underscore a broader methodological issue: international and Vietnamese research alike lack a unified multimodal procedure that integrates visual grammar, linguistic metafunctions, and intersemiotic analysis into a coherent analytic design.

Taken collectively, these methodological shortcomings reveal the absence of a robust analytical apparatus capable of addressing the core concerns of this thesis. Because RQ1 requires an examination of how visual and linguistic modes interact to construct medical meaning, and RQ1.1 – RQ1.2 necessitate systematic analysis of both intra-modal structures and inter-modal orchestration, existing approaches are insufficient. What is needed – and what this study proposes – is a critically integrated, socially-

semiotic, and stratified multimodal methodology that moves beyond fragmented accounts to capture the full complexity of multimodal semiosis in EMP textbooks.

2.6.1.3 Contextual Gap: Lack of Multimodal Studies on EMP Textbooks in Vietnam

Although research on ESP and EMP in Vietnam has expanded considerably in recent decades (Ton, 2017; Pham, 2022, Tran & Lam, 2025), the majority of studies remain anchored in linguistic, curricular, or perceptual concerns. Much of this scholarship prioritises needs analyses, evaluates language proficiency requirements, or examines issues related to content relevance and learner perceptions. While these contributions illuminate structural and pedagogical challenges facing EMP instruction, they provide only partial insight into how disciplinary knowledge is actually constructed through the multimodal design of teaching materials.

Moreover, recent attempts to introduce multimodal perspectives into Vietnamese ELT and ESP research such as Hoang (2015), Le (2017), Pham (2022), and Lam and Pham (2024) – mark a valuable shift towards recognising the role of visual resources in shaping learning. However, these studies remain methodologically limited. Analyses often focus on isolated images or on representational dimensions alone, without systematically examining how visual, linguistic, and spatial resources operate as interdependent semiotic strata. Consequently, intersemiotic modelling remains underdeveloped, and the pedagogical implications of multimodal design for disciplinary literacy are seldom explored. This stands in contrast with international scholarship, where multimodal integration is increasingly viewed as central to understanding how scientific and medical knowledge is mediated (Bezemer & Kress, 2016; O’Halloran, 2008, Jewitt, 2009b).

Critically, no study in Vietnam has yet applied a systemic, social-semiotic multimodal framework capable of analysing how visual grammar (e.g., representational, interactive, compositional structures) interfaces with linguistic metafunctions to construct coherent medical meaning. The absence of a theoretically integrated approach – one that combines Kress and van Leeuwen’s visual grammar with multimodal metafunctions and intersemiosis – means that Vietnamese research has not addressed how EMP textbooks induct learners into biomedical ways of seeing, reasoning, and professional knowing. This omission is significant, as EMP materials used in Vietnamese medical universities often blend imported visual conventions with locally produced textual explanations, creating multimodal configurations whose coherence cannot be evaluated through monomodal or descriptive methods alone.

Importantly, this contextual gap directly intersects with the research questions driving the present study. RQ1 requires an examination of how linguistic and visual modes jointly construct medical meaning, while RQ1.1 and RQ1.2 demand a systematic account of intra-modal structures and inter-modal orchestration. The current Vietnamese literature is not methodologically equipped to answer these questions. As a result, a critical need emerges for an integrated, stratified, and contextually sensitive multimodal framework, one capable of capturing the complex semiotic processes through which EMP textbooks mediate biomedical knowledge within Vietnamese higher education.

2.6.1.4 Pedagogical Gap: Lack of Understanding of How Multimodal Design Shapes Disciplinary Literacy

Although a substantial body of international research demonstrates that multimodal design fundamentally shapes comprehension, reasoning, and

disciplinary knowledge construction (Royce, 2007; Bateman, 2014), pedagogical practice in EMP in Vietnam continues to treat visual resources in a largely instrumental manner. Images are frequently positioned as supplementary embellishments, vocabulary cues, or static illustrations rather than as integral components of how biomedical concepts are construed and communicated. This pedagogical orientation contrasts sharply with social-semiotic perspectives, which hold that meaning is distributed across linguistic, visual, and spatial modes, and that disciplinary learning depends on the orchestrated interplay of these semiotic resources (Kress, 2010; Bezemer & Kress, 2016).

Moreover, existing EMP evaluations – both globally and within Vietnam – have seldom examined how multimodal ensembles cultivate what Goodwin (1994) terms professional vision: the socially and epistemically patterned ways of seeing that underpin clinical reasoning. While scholars such as Jewitt and Oyama (2014) have shown how visual representations encode professional authority, and Foucault (1973) has theorised the clinical gaze as a historically situated mode of perception, these insights have not been systematically applied to the analysis of EMP textbooks. As a result, the pedagogical functions of multimodal design – how diagrams guide attention to clinically salient features, how layout structures cognitive pathways, or how intersemiotic coherence scaffolds biomedical explanation – remain largely untheorised in the Vietnamese context.

Critically, this lack of pedagogically informed multimodal research means that current evaluations fail to capture the semiotic mechanisms through which disciplinary literacy is cultivated. Biomedical literacy is not acquired solely through lexical or grammatical exposure but through engagement with multimodal constellations of text, image, label, vector, and

spatial arrangement. International studies in science and health education (e.g., O'Halloran, 2008; Tang et al., 2014) have demonstrated that students learn complex phenomena by integrating visual and verbal cues into coherent mental models. However, Vietnamese EMP materials have not been analysed through frameworks that can explain how such multimodal reasoning is facilitated – or hindered – by their design.

Importantly, this pedagogical gap directly intersects with the objectives of the present study. RQ1 requires an account of how multimodal design constructs medical meaning, while RQ1.1 and RQ1.2 demand an understanding of how intra-modal structures and intersemiotic relations jointly support disciplinary learning. Without a systematic, social-semiotic investigation of multimodal design, current pedagogical evaluations risk overlooking the very mechanisms through which learners develop the disciplinary reasoning, epistemic stance, and professional perception required in medical contexts.

Taken together, these limitations highlight an urgent need for a critically integrated, multimodal, and pedagogically grounded analytical framework that can illuminate how visual and linguistic resources collaboratively shape medical disciplinary literacy in Vietnamese EMP textbooks.

Overall, across international and Vietnamese scholarship, a single overarching gap becomes evident: no existing study combines visual grammar with multimodal metafunctions in a systematic framework to examine how linguistic and visual modes jointly construct medical meaning in EMP textbooks. Addressing this critical absence constitutes the central contribution of the present study.

2.6.2 Conceptual Framework for the Study

This thesis is conceptually anchored in social semiotics, which treats meaning as socially situated semiotic choice rather than as the neutral

transmission of information. From this epistemological standpoint, English medical textbooks are understood not merely as pedagogic containers of biomedical facts, but as designed semiotic artefacts through which disciplinary knowledge and professional relations are selected, organised, and stabilised. Accordingly, what counts as “*medical meaning*” in a textbook emerges from patterned choices in representation, stance, and organisation that are shaped by institutional norms and disciplinary histories; in turn, these choices orient learners towards sanctioned ways of seeing, knowing, and speaking within the medical domain (Halliday, 1978; Hodge & Kress, 1988; van Leeuwen, 2005; Kress, 2010).

Because social semiotics assumes that semiotic resources are culturally organised and institutionally regulated, it follows that interpretation must be systematic and principled rather than impressionistic. The analytical aim, therefore, is not to describe isolated textual features or visual elements as “effective” or “attractive”, but to explain how meaning potentials are activated through conventionalised selections and how these selections become pedagogically functional in context (van Leeuwen, 2005). This requirement for accountable explanation is particularly salient in medical education, where knowledge is characteristically technical, classificatory, and normative, and where textbooks routinely encode epistemic authority and disciplinary legitimacy through semiotic design.

To provide a coherent organising logic for this social-semiotic stance, the thesis draws on Systemic Functional Linguistics (SFL) as a theory of meaning-making. SFL is especially appropriate because it conceptualises language as a “*social semiotic*” and models meaning as functional choice in context rather than as autonomous form (Halliday, 1978; Halliday & Matthiessen, 2014). Importantly, at this stage SFL functions primarily as a conceptual backbone that constrains interpretation and clarifies what counts as analytic evidence, rather than as an independent analytical tier involving

exhaustive linguistic coding. This design choice preserves the theoretical clarity of the conceptual framework while preparing the ground for later multimodal operationalisation.

Central to SFL is the metafunctional hypothesis, which holds that meaning-making simultaneously realises ideational, interpersonal, and textual meanings (Halliday & Matthiessen, 2014). This principle is crucial because it prevents analysis from reducing textbook meaning to “*content*” alone. The ideational metafunction motivates attention to how biomedical experience is construed through processes, participants, relations, and taxonomies; the interpersonal metafunction foregrounds how stance, authority, evaluation, and professional role relations are enacted; and the textual metafunction accounts for how information is staged, sequenced, and made navigable within dense instructional discourse.

Equally significant is SFL’s principle of stratification, which models semiosis as layered from context → semantics → lexicogrammar → expression (Halliday & Matthiessen, 2014). Stratification enables the study to connect observable semiotic forms - wording, organisation, and layout - to higher-level meaning potentials, while avoiding “*level errors*” in interpretation. In this view, surface patterns are not treated as self-contained features but as social realisations of meaning shaped by contextual conditions such as curricular aims, institutional assessment regimes, and disciplinary epistemologies.

In addition, SFL explains context through register variables - field, tenor, and mode - which specify how situational configurations shape meaning potential (Halliday & Matthiessen, 2014; Eggins, 2004). This is conceptually productive for EMP textbook discourse because it highlights that biomedical “*content*” (field) is inseparable from pedagogic authority and role relations (tenor), as well as from the communicative channel and organisation of discourse (mode). Register therefore provides a principled bridge between

the social-semiotic claim that meaning is socially patterned and the textual evidence through which such patterning can be demonstrated.

Finally, SFL's notion of instantiation clarifies the relationship between particular textbook texts and the wider semiotic resources of a culture. Textbooks can be understood as instantiations of broader disciplinary meaning potentials, stabilised into conventionalised pedagogic forms (Halliday & Matthiessen, 2014). This perspective supports analytic caution: rather than assuming that textbooks transparently reflect biomedical reality, the thesis examines how they semiotically construct what counts as legitimate knowledge and how they position learners in relation to that knowledge.

Taken together, social semiotics provides the epistemological grounding for meaning as socially situated choice, while SFL supplies the core systemic logic - metafunctions, stratification, register, and instantiation - through which this grounding becomes analytically workable. Consequently, this conceptual framework establishes what meaning is assumed to be, how meaning is organised in relation to context and realisation, and why textbook discourse can be examined as patterned semiotic design. Building on this theoretical rationale, subsequent chapters operationalise these commitments into an explicit analytical procedure for examining multimodal textbook page ensembles in a transparent and auditable manner.

Figure 2.2 summarises the conceptual foundation of the study by showing how social semiotics provides the epistemological base, while SFL supplies the core organising logic for describing meaning potential in textbook discourse.

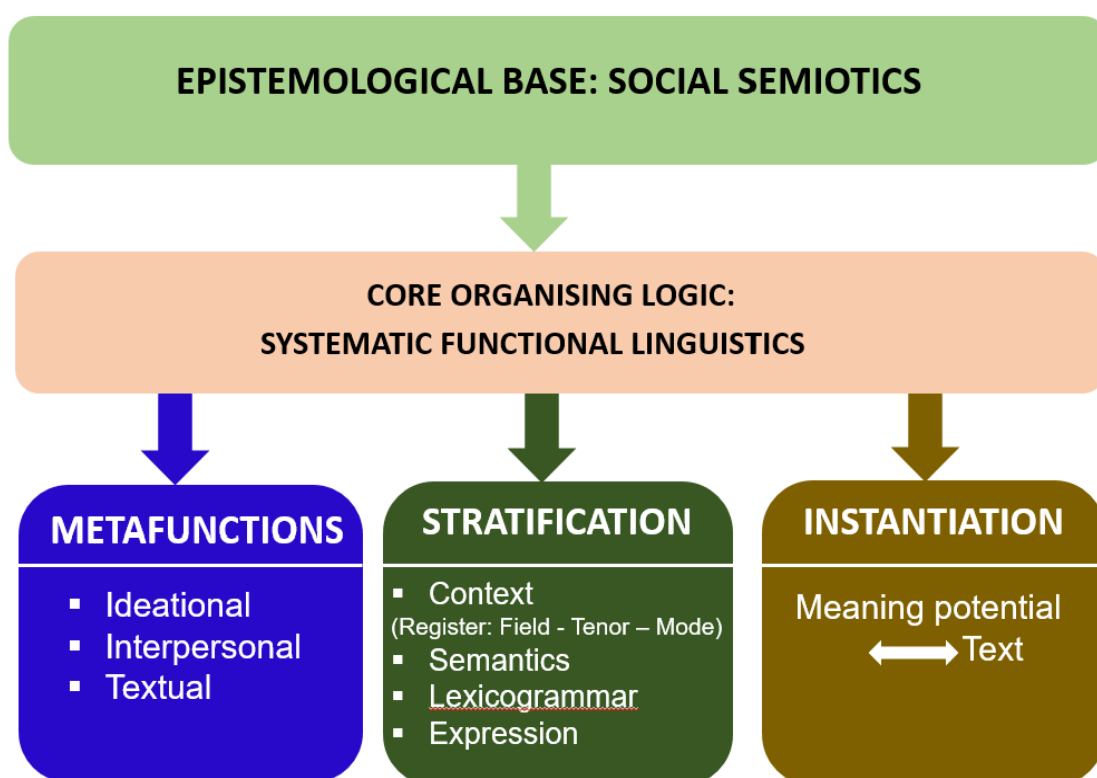


Figure 2.2 Conceptual Framework for the Study

As illustrated in Figure 2.2, the study is grounded in social semiotics, which treats meaning as socially situated semiotic choice. On this epistemological basis, SFL provides the systemic logic for explaining how meaning is organised through metafunctions (ideational, interpersonal, textual), stratification (context-semantics-exicogrammar-expression), and register (field-tenor-mode). Instantiation further clarifies how particular textbook texts realise selections from broader cultural meaning potentials, allowing textbook discourse to be examined as patterned semiotic design rather than neutral representation.

2.7 Concluding Remarks

This chapter has established the theoretical and empirical foundation for examining EMP textbooks as multimodal pedagogic artefacts. It first traced

the evolution of ESP from early register-oriented descriptions of specialised language (Ewer & Latorre, 1969) to discourse- and genre-based perspectives that reconceptualise disciplinary communication as socially situated practice (Swales, 1990; Bhatia, 1993; Dudley-Evans & St John, 1998; Hyland, 2005). Within this trajectory, EMP emerges as a particularly complex domain because medical discourse is simultaneously linguistically specialised and epistemically visual, requiring learners to integrate technical language with diagrammatic and procedural representations (Candlin & Candlin, 2002; Salager-Meyer, 2014).

The chapter then argued that EMP textbooks cannot be treated as neutral containers of biomedical information. Rather, they function as designed mediating artefacts that organise disciplinary knowledge, encode pedagogic priorities, and position learners within professional relations (Cunningsworth, 1995; Tomlinson, 2011; McGrath, 2013). In medicine, this mediating role is intensified because key concepts depend on visual and spatial representation, meaning that textbook meaning-making must be evaluated not only through linguistic content but also through multimodal orchestration and image-text coherence.

To ground such analysis, the chapter positioned social semiotics as the epistemological base of the study, where meaning is understood as socially organised semiotic choice shaped by institutional norms and disciplinary histories (Halliday, 1978; Hodge & Kress, 1988; van Leeuwen, 2005; Kress, 2010). Within this paradigm, Systemic Functional Linguistics (SFL) provides the organising logic that renders interpretation principled and accountable, through its metafunctional model and stratified architecture linking context, meaning, and realisation (Halliday & Matthiessen, 2014; Eggins, 2004).

Crucially, SFL also offers a transmodal rationale that supports systematic extension into multimodal analysis.

Building on this grounding, the chapter reviewed major contributions in multimodal discourse analysis, demonstrating that educational meaning - especially in scientific and medical discourse - is distributed across semiotic modes rather than conveyed by language alone (Lemke, 1998; O'Halloran, 2008; Jewitt, 2014; Bezemer & Kress, 2016). The Grammar of Visual Design provides a robust metalanguage for analysing how medical images realise representational, interactive, and compositional meanings (Kress & van Leeuwen, 2021). However, multimodal explanation also requires modelling how images and language operate together. Here, Unsworth's multimodal metafunctions and intersemiotic relations enable principled analysis of cross-modal coupling (Unsworth, 2008), while Royce's intersemiotic complementarity strengthens operational mapping between visual and verbal meaning potentials (Royce, 2007).

Despite these advances, the chapter identified a persistent limitation in EMP textbook research: multimodal meaning is rarely analysed as an integrated system. International studies often prioritise linguistic coverage, skills balance, or user perceptions, while visual-verbal integration remains marginal or descriptive (Dimopoulos et al., 2003; Slough et al., 2010; Rahman, 2015). In Vietnam, EMP textbook research has similarly focused on language-based criteria and content relevance, with limited systematic attention to how biomedical knowledge is visually constructed and intersemiotically coordinated (Le, 2017; Ton, 2020; Tran & Lam, 2025). Although emerging multimodal work signals progress, it tends to remain intra-modal, leaving image-text relations under-theorised (Lam & Pham, 2024). Consequently, there remains a clear need for an analytical framework

that can capture both intra-modal organisation and intersemiotic meaning-making in EMP textbook page ensembles.

CHAPTER III: RESEARCH METHODOLOGY

Chapter III presents the methodological architecture for examining multimodal meaning-making in English medical textbooks used in Vietnam. Grounded in a theory-driven, interpretive-explanatory social semiotic paradigm, it shows how the paradigm is operationalised into a qualitative design, a contrastive corpus, and transparent procedures for data selection and collection. It then details the staged analytical procedure and the Integrated Multimodal Analytical Framework guiding the move from systematic visual description to theory-constrained intersemiotic explanation, and concludes with the study's approach to trustworthiness, ethics, and reflexive rigour, providing the foundation for the analyses in Chapters IV and V.

3.1 Research Paradigm

This study adopts a theory-driven, interpretive-explanatory social-semiotic research paradigm, drawing on social semiotics and multimodal discourse analysis as its principal epistemic and analytical traditions. Within this paradigm, meaning is not treated as an objective property embedded in texts or images, but as an outcome of semiotic design and situated interpretation. Disciplinary knowledge, particularly as represented in educational materials, is therefore understood as something that is made available through patterned configurations of visual, linguistic, and spatial resources, and realised through interpretive engagement rather than accessed as pre-existing truth.

Ontologically, the paradigm assumes that medical knowledge as represented in textbooks does not constitute a transparent reflection of biomedical reality. Instead, it is a discursively and semiotically organised representation, shaped by institutional norms, pedagogical intentions, and disciplinary conventions. Anatomical diagrams, procedural illustrations, and

schematic representations do not simply depict medical facts; rather, they select, foreground, and hierarchise aspects of reality in ways that are meaningful within particular traditions of medical education. In this sense, textbook visuals function as epistemic artefacts – designed representations that structure how learners are invited to see, conceptualise, and reason about the medical domain (Lemke, 1998; Davies & Elder, 2004, Creswell & Poth, 2018, Kress & van Leeuwen, 2021).

Epistemologically, the study advances an interpretive-explanatory stance in which explanation is achieved through theoretically disciplined interpretation rather than through causal prediction or statistical generalisation. Meaning is approached as meaning potential (Halliday & Matthiessen, 2014): a range of semiotic possibilities activated through representational, interactional, and compositional choices and realised through intersemiotic relations across modes. Importantly, images and texts are not treated as self-sufficient carriers of meaning. Instead, meaning emerges through the interaction between semiotic resources and the analytic frameworks brought to bear on them. This conception of explanation is well established in multimodal scholarship, where explanation involves accounting for how semiotic configurations work, why particular design choices recur, and what kinds of understanding they make available for intended audiences (Kress, 2010; Jewitt, 2014; Bezemer & Kress, 2016).

Crucially, this paradigm supports an explicitly top-down analytic orientation. The study does not proceed inductively from data to theory; instead, it mobilises established theoretical frameworks from the outset to guide observation, coding, and interpretation. The GVD (Kress & van Leeuwen, 2021) and Unsworth's multimodal metafunctional model (2006, 2008) provide principled analytical lenses through which multimodal patterns

are identified, compared, and explained. However, a theory-driven orientation does not entail theoretical imposition. These frameworks operate as heuristic constraints that discipline interpretation, ensuring analytical rigour while allowing meanings to be explicated in relation to context, pedagogic purpose, and semiotic configuration.

Moreover, foregrounding visual semiosis is indispensable in the medical domain, where disciplinary knowledge is characteristically multimodal. Processes are spatialised through diagrams, relationships are rendered through layout and salience, and authority is frequently constructed through visual abstraction and standardisation. Visual images in medical textbooks therefore function not as ancillary supports to language but as semiotic resources that actively participate in constructing what counts as relevant, normal, and authoritative medical knowledge (Lemke, 1998; O'Halloran, 2008). Recognising this epistemic role of visuals is central to the study's overarching aim of analysing how visual and linguistic modes interact to construct medical knowledge and professional meaning.

However, when English medical textbooks circulate beyond their original contexts of production, multimodal designs are recontextualised within local pedagogical traditions, learner expectations, and institutional constraints (van Leeuwen, 2005). While research in ESP and EMP has increasingly acknowledged tensions between imported materials and local teaching realities, systematic analyses of visual-verbal meaning-making in medical textbooks remain limited. This study addresses this gap by offering a context-sensitive explanation of how multimodal resources are designed to function pedagogically, without assuming uniform interpretation or learning effects.

This paradigm also entails a clear and explicit positioning of the researcher. The researcher is not a medical practitioner and does not claim authority to evaluate biomedical accuracy. Rather, the researcher assumes the role of an EMP researcher and educator, analysing how medical knowledge is semiotically constructed and pedagogically mediated for intended non-expert learners. Critically, this positioning is not epistemically deficient but analytically appropriate, given that textbooks are pedagogic artefacts designed to bridge expert knowledge and novice understanding. Interpreting them from the standpoint of informed pedagogic literacy allows the study to examine how meaning is made accessible, rather than how it is verified.

Nevertheless, the paradigm recognises that interpretation is unavoidably shaped by the researcher's disciplinary background and theoretical commitments. For this reason, reflexivity is treated as an epistemic requirement, not an optional methodological add-on. Interpretive claims are anchored in explicit analytical categories derived from established frameworks, documented through systematic coding procedures, and continuously checked against alternative readings. Such practices align with recognised qualitative standards for analytical rigour and transparency (Lincoln & Guba, 1985; Creswell, 2009, Miles, Huberman, & Saldaña, 2019), and respond directly to long-standing critiques of subjectivity in multimodal analysis.

Finally, the paradigm is tightly aligned with the study's aims, objectives, and research questions. The overarching research question – how visual and linguistic modes interact to construct medical meaning – presupposes an explanatory account of semiotic interaction rather than a descriptive inventory of features. RQ1 requires a principled analysis of visual organisation, enabled through the top-down application of visual grammar,

while RQ2 requires an interpretive explanation of how multimodal metafunctions are jointly realised across modes. Although RQ1 foregrounds intra-modal visual organisation, it is examined at the level of the page ensemble so that visual meanings are interpreted in relation to co-present linguistic and layout cues. Together, these questions position explanation as the systematic elucidation of semiotic mechanisms through which disciplinary meaning is constructed in English medical textbooks used in Vietnam.

In sum, the research paradigm adopted in this study provides a coherent philosophical and epistemic foundation for a theory-driven, interpretive – explanatory investigation of multimodal meaning-making in EMP textbooks. By foregrounding semiotic design, researcher positionality, and contextual interpretation, it binds the study’s conceptual frameworks, analytical procedures, and research questions into a unified inquiry into how medical knowledge is visually and linguistically constructed for Vietnamese learners of English.

3.2 Research Design

Building upon the research paradigm outlined in Section 3.1, this study adopts a theory-driven, interpretive-explanatory qualitative research design to operationalise the analysis of multimodal meaning-making in English medical textbooks used in Vietnam. The design translates the study’s epistemic commitments into a coherent analytical architecture through which established theoretical frameworks are systematically applied to explain how visual and linguistic resources interact to construct disciplinary and pedagogical meaning.

The design is explicitly top-down in orientation. Rather than generating analytical categories inductively from the data, the study mobilises established frameworks from social semiotics and multimodal discourse

analysis to guide data selection, coding, and interpretation from the outset. This approach reflects a central principle of multimodal research: explanation depends on theoretically specified systems for describing how meaning is designed and organised across modes (Kress, 2010; Jewitt, 2014, Schwandt, 1998, Schwandt, 2007). At the same time, theoretical frameworks function as heuristic resources that discipline interpretation while remaining sensitive to contextual variation and pedagogical purpose.

Methodologically, the study adopts a qualitative multimodal design situated within the social semiotic tradition. Visual and linguistic modes are treated as interdependent semiotic resources whose meanings emerge through patterned interaction within multimodal ensembles (Bezemer & Kress, 2016, Sandelowski, 1995, Royce, 1998). Accordingly, the analysis integrates visual grammar, linguistic metafunctions, and intersemiotic relations as complementary dimensions of meaning-making.

Within this design, a comparative multiple-case approach is employed, focusing on two English for Medical Purposes textbooks: *Medical Terminology: A Living Language* (Fremgen & Frucht, 2019) and *Career Paths: Medical* (Evans, Dooley, & Tran, 2012). These cases are theoretically selected to represent contrasting pedagogical and semiotic orientations within the same disciplinary domain. While *Medical Terminology* foregrounds classificatory and taxonomic representations through labelled diagrams and schemata, *Career Paths: Medical* emphasises procedural and interactional meanings through contextualised visuals and simulated professional discourse. Comparing these cases enables explanation of how different multimodal configurations realise disciplinary knowledge in distinct ways.

The unit of analysis is the textbook page ensemble, conceptualised as a multimodal constellation in which visual, linguistic, and spatial resources are

intentionally arranged to guide interpretation (van Leeuwen, 2005; Bezemer & Kress, 2016). Analysing the page ensemble, rather than isolated images or texts, allows the study to capture how representational structure, intersemiotic alignment, and layout jointly contribute to pedagogical meaning. Page ensembles are selected across instructional domains such as anatomical description, diagnostic explanation, and patient-practitioner interaction to ensure analytical breadth.

Analytically, the design follows a layered and recursive procedure aligned with the IMAF and its analytical realisation. Visual analysis addresses representational, interactional, and compositional structures (RQ1), while intersemiotic analysis explains how visual and linguistic modes jointly realise ideational, interpersonal, and textual meanings (RQ2). These analytical layers are treated as interdependent perspectives rather than linear steps, enabling iterative refinement of interpretation.

To ensure analytical rigour, the study employs theory-linked codebooks, an audit trail of analytic decisions, and reflexive memos documenting interpretive choices and alternative readings (Lincoln & Guba, 1985; Dornyei, 2007; Miles, Huberman, & Saldaña, 2019). These procedures render interpretation transparent and theoretically accountable, addressing critiques of subjectivity in multimodal analysis.

Importantly, the research design is deliberately bounded. It does not examine learner reception or learning outcomes, but focuses on explaining how meaning is designed to function pedagogically within multimodal textbook artefacts. This focus aligns directly with the study's aims and research questions, ensuring coherence between paradigm, design, and scope.

3.3 Data Corpus

Situated within the theory-driven, interpretive-explanatory paradigm articulated in Section 3.1 and operationalised through the research design in Section 3.2, the data corpus of this study is conceptualised as a pedagogically designed semiotic environment rather than a neutral collection of textbook pages. From a social semiotic perspective, meaning is not located in isolated images or linguistic units, but emerges from patterned configurations of visual, verbal, and spatial resources designed to support disciplinary understanding (Kress & van Leeuwen, 2021; Unsworth, 2008). Accordingly, the corpus is constructed to enable a principled examination of how multimodal resources are deployed to construct medical meaning and disciplinary knowledge in English medical textbooks used in Vietnam.

Consistent with the Conceptual Framework (Figure 2.2), the corpus is designed to support two closely related analytical foci. First, it enables systematic analysis of how visual images are employed and organised representationally, interactively, and compositionally, thereby directly addressing RQ1. Second, it provides a robust empirical basis for examining how visual and linguistic modes interact to realise ideational, interpersonal, and textual meanings across modes, as required by RQ2. At a higher level, the corpus also supports an integrative explanatory interpretation within the scope of RQs, allowing the study to account for how multimodal configurations are pedagogically oriented toward learners in Vietnamese EMP contexts.

3.3.1 Rationale for Book-level Selection

The corpus comprises two internationally published English medical textbooks that are widely used in Vietnamese medical and health science universities, yet embody contrasting multimodal and pedagogic orientations within the shared domain of EMP: *Medical Terminology: A Living Language*

(7th ed.) by Fremgen and Frucht (2019) and *Career Paths: Medical* by Evans, Dooley, and Tran (2012).

The selection is theory-driven rather than representative. The two textbooks were chosen because they instantiate different semiotic strategies for constructing medical meaning. *Medical Terminology* foregrounds a conceptual-classificatory orientation, in which disciplinary knowledge is construed through systematic terminology building and conceptual visualisation. In contrast, *Career Paths: Medical* adopts a contextualised, task-oriented orientation, in which medical lexis and meanings are embedded within staged communicative activities and simulated professional contexts. Analysed together, the two textbooks form a contrastive corpus that enables systematic comparison of multimodal meaning-making across distinct pedagogic designs while remaining within a common disciplinary frame.

3.3.2 Textbook 1: Medical Terminology: A Living Language (Fremgen & Frucht, 2019)

Medical Terminology: A Living Language is a foundational EMP textbook designed specifically for introductory medical terminology courses. Its primary pedagogic aim is to support novice learners in acquiring, decoding, and applying specialised medical lexis, foregrounding terminology mastery as the principal mode of access to disciplinary knowledge (Fremgen & Frucht, 2019).

Structurally, the textbook follows a highly regularised and pedagogically scaffolded organisation. After introductory chapters on word-building principles and body organisation, subsequent chapters are organised by body systems. Each system chapter unfolds through a recurrent multimodal sequence: System at a Glance pages listing combining forms, prefixes, and suffixes; System Illustrated sections presenting labelled

anatomical diagrams; concise anatomy and physiology explanations; and extended terminology sections covering pathology, diagnostic procedures, therapeutic procedures, pharmacology, and abbreviations. This stable organisation reinforces a view of medical knowledge as systematically classifiable and terminologically decomposable.

From a visual semiotic perspective, the textbook is characterised by a strong dominance of conceptual representations as described in the GVD (Kress & van Leeuwen, 2021). Images are predominantly static, non-narrative, and decontextualised from social action, depicting anatomical entities through analytical part-whole relations, taxonomic structures, and schematic cross-sections. Human figures, when present, function primarily as anatomical carriers rather than as social actors. Compositional resources such as centre-margin and top-bottom information value, tight framing, and high salience of labelled elements are systematically deployed to stabilise canonical biomedical perspectives and direct learners' attention toward terminological precision.

These visual designs are tightly integrated with dense definitional discourse. Linguistically, the textbook relies heavily on relational processes, extensive nominalisation, and complex technical noun groups, resulting in high lexical density characteristic of scientific discourse (Halliday & Matthiessen, 2014). Theme-Rheme progression frequently mirrors the spatial logic of diagrams, whereby visually foregrounded anatomical elements become thematic anchors in accompanying definitions. This alignment exemplifies ideational and textual intersemiosis, in which visual classification and linguistic definition converge to construe disciplinary meaning (Unsworth, 2008).

Within the present study, *Medical Terminology* constitutes a particularly productive corpus for addressing RQ1, given the clarity and consistency with which visual representational and compositional structures are realised. It also provides a robust baseline for RQ2, enabling close examination of how visual and linguistic modes jointly realise ideational and textual meanings through systematic visual-verbal alignment.

3.3.3 Textbook 2: *Career Paths: Medical* (Evans, Dooley, & Tran, 2012)

In contrast, *Career Paths: Medical* constructs medical knowledge through a skills-oriented, task-based multimodal design that foregrounds lexicalised content-in-context rather than abstract conceptual systematisation. Organised into short, thematically coherent units (e.g. Heart Disease, Traumatic Injuries, Infections), the textbook adopts an ESP-oriented instructional sequence in which disciplinary meanings are progressively introduced, practised, and recontextualised across reading, listening, speaking, and writing activities.

Visually, the textbook relies on a heterogeneous mix of photographs, illustrative scenes, posters, and simple infographics embedded within instructional sequences. These visuals rarely function as stand-alone representations of biomedical knowledge. Instead, they operate as contextualising prompts that anchor technical lexis and textual genres within recognisable clinical situations. Representational meanings are predominantly narrative or contextualised conceptual, foregrounding participants, settings, and artefacts rather than detailed internal anatomy (Kress & van Leeuwen, 2021, Unsworth, 2006).

Interactive meanings are salient across the textbook. Images frequently employ oblique or frontal angles and medium social distance, positioning

learners as legitimate peripheral participants rather than detached observers. Gaze is typically indirect, aligning with the pedagogic function of orienting learners to professional environments rather than demanding interpersonal engagement. Linguistically, the textbook privileges controlled technical lexis embedded in simplified expository and dialogic texts, with frequent use of declaratives, interrogatives, and selective modality (should, must, need to) to enact basic professional stance.

Meaning-making in *Career Paths: Medical* is achieved primarily through intersemiotic complementarity rather than intersemiotic density. Visuals extend and contextualise linguistic meaning by situating terminology within familiar clinical scenarios, while ideational content is largely realised through language (Unsworth, 2008). This configuration makes the textbook analytically central to RQ2, particularly for examining how visual and linguistic modes jointly realise interpersonal and textual meanings in pedagogically staged contexts.

3.3.4 Unit of Analysis and Corpus Delimitation

Consistent with a social semiotic view of meaning as orchestrated across modes, the unit of analysis in this study is the page ensemble – a designed constellation of images, captions, adjacent text, typography, and layout that together realise pedagogic meaning (Bezemer & Kress, 2016). Analysing meaning at the level of the page ensemble preserves multimodal coherence and avoids fragmenting intersemiotic relations.

The corpus was delimited through maximum variation sampling (Miles, Huberman, & Saldaña, 2019) to capture diversity in visual type, pedagogic function, and disciplinary focus. Approximately forty page ensembles were selected: twenty from *Medical Terminology* and twenty from *Career Paths: Medical*. Inclusion criteria included (i) semiotic salience, (ii) pedagogic

centrality, and (iii) representational diversity. A complete inventory of selected page ensembles, including textbook source and page references, is provided in Appendix A (Page-Ensemble Inventory), ensuring analytical transparency and traceability from interpretive claims to empirical evidence.

3.4 Data Selection Criteria

Data selection in this study is treated as an explicitly theory-calibrated stage of research design rather than a preliminary technical step. Consistent with the theory-driven, interpretive-explanatory social-semiotic paradigm (Section 3.1) and the top-down analytic design (Section 3.2), the corpus is constructed to make analytically visible how multimodal meaning is designed and why particular visual-verbal configurations are pedagogically consequential in EMP textbooks used in Vietnam. Importantly, the Conceptual Framework (Figure 2.2) determines what semiotic work counts as analytically focal (visual grammar + multimodal metafunctions) and specifies how that work is traced through a staged reading of visual, linguistic, and intersemiotic organisation. Therefore, the criteria below function as operational filters that ensure the selected data can sustain a rigorous explanation of RQs rather than producing only descriptive commentary.

3.4.1 Construct-to-Criterion Alignment with the Research Questions

The first principle of selection was that every inclusion rule must be traceable to a construct embedded in the research questions.

For the overarching RQ (mode interaction in constructing medical meaning): multimodal integrativeness.

Units were included only when visual and linguistic resources were demonstrably co-designed for a shared local pedagogic purpose (e.g., defining a concept, stabilising terminology, staging a clinical exchange), rather than merely co-present on the same page. In other words, interaction was treated as a design relation, not a proximity effect. This criterion is consistent with

multimodality research that locates meaning in the orchestration of modes within designed ensembles rather than in isolated semiotic components (Bezemer & Kress, 2016).

For RQ1 (visual organisation): diagnostic richness in GVD terms.

Eligible units had to enable principled coding of representational, interactive, and compositional meaning. Accordingly, selection prioritised ensembles where (i) representational structures were legible (conceptual part-whole, taxonomy, or narrative action), (ii) interactive cues were interpretable (gaze, social distance, angle, modality cues), and (iii) composition was purposeful (information value, salience, framing) (Kress & van Leeuwen, 2021). Critically, this does not mean “*choosing the most illustrated pages*”; it means selecting pages where image design does semiotic work that can be analysed systematically.

For RQ2 (metafunctional co-realisation across modes): explicit intersemiotic anchoring.

Units were included only when there was identifiable cross-modal coupling – for example, labelled visual elements taken up by definitional clauses, procedure visuals aligned with material/imperative clause patterns, or layout sequencing that supported textual organisation and reading paths. Such anchoring is essential for analysing ideational/interpersonal/textual meanings as jointly realised across modes, instead of describing image and language in parallel tracks (Unsworth, 2008; Martinec & Salway, 2005; Royce, 2007).

3.4.2 Unit of Analysis: Page Ensemble

The unit of analysis is the page ensemble, defined as a designed constellation of image(s), caption(s), co-present verbal text, typography, and layout cues that together realise an instructional message. This choice is methodologically decisive. A diagram’s meaning potential, for instance, is rarely exhausted by the image alone; it is typically completed by definitional

wording, labels, typographic hierarchy, and layout positioning. Conversely, textual explanations frequently presuppose visual salience and spatial organisation. Therefore, analysing isolated images or isolated paragraphs would fragment the very intersemiotic relations that RQs aim to explain. The page ensemble preserves designed cohesion and supports accountable interpretation of multimodal interaction (Bezemer & Kress, 2016; Kress & van Leeuwen, 2021).

3.4.3 Inclusion and Exclusion Rules

To achieve breadth while maintaining interpretive depth, page ensembles were selected through maximum variation sampling, ensuring diversity in multimodal forms and pedagogic functions (Patton, 2015).

Inclusion rules. A page ensemble was included if it met three conditions:

- (i) multimodal co-functionality, where image and language jointly served a coherent pedagogic action (e.g., defining, classifying, explaining, or staging interaction);
- (ii) analytical tractability, enabling principled coding under GVD categories (RQ1) and metafunctional co-realisation across modes (RQ2); and
- (iii) pedagogic centrality, meaning the ensemble contributed substantively to the instructional sequence rather than functioning as peripheral decoration.

Exclusion rules. Pages were excluded if they contained:

- (i) purely decorative visuals without meaningful anchoring in the co-text;
- (ii) text-dominant content with negligible visual contribution to meaning-making; or
- (iii) analytically discontinuous ensembles where captions, co-text, or layout relations were absent or insufficient for intersemiotic interpretation.

These exclusions do not indicate pedagogic insignificance; they simply fall beyond the study's design-focused scope of analysable visual-verbal interaction.

3.4.4 Textbook-Specific Selection Trajectories

Although the same core criteria applied across both cases, selection was also calibrated to each textbook's semiotic ecology in order to maximise analytic leverage.

Medical Terminology: A Living Language (Fremgen & Frucht, 2019): conceptual-terminological ensembles.

Priority was given to ensembles that instantiate the book's stabilised instructional architecture – particularly System At a Glance and System Illustrated sequences where terminology-building resources (combining forms, prefixes/suffixes) are tightly coordinated with labelled conceptual images and definitional prose. These ensembles are optimal for examining (i) conceptual representation (part-whole, classification), (ii) compositional hierarchy (salience, information value), and (iii) ideational/textual convergence between diagrammatic organisation and definitional discourse (Kress & van Leeuwen, 2021; Unsworth, 2008).

Career Paths: Medical (Evans, Dooley, & Tran, 2012): contextualised task-based ensembles.

Selection prioritised ensembles where visuals function as contextualising prompts within staged tasks (reading, listening, dialogue, controlled speaking/writing), enabling examination of how meaning is distributed across modes in pedagogically sequenced activity structures. Such ensembles are particularly productive for analysing interpersonal and textual meanings realised through the coordination of image positioning (e.g., distance/angle/gaze patterns) with dialogic and modalised language typical of introductory clinical communication tasks (Kress & van Leeuwen, 2021;

Royce, 2007). Moreover, because this textbook often favours intersemiotic complementarity rather than dense visual explication, selection targeted pages where the division of semiotic labour between image and language could be interpreted and explained with clarity (Unsworth, 2008).

3.4.5 Corpus Size, Saturation Logic, and Documentation

The final dataset comprises 40 page ensembles (20 from each textbook). This scale is justified by interpretive manageability and category saturation: pilot coding indicated that additional pages tended to reproduce already-observed configurations rather than extending the analytical system. Moreover, given the study's top-down explanatory aim, the goal is not statistical representativeness but theoretical coverage of the major multimodal configurations needed to explain the research questions within the defined scope.

To support transparency, each selected ensemble is logged in an audit matrix (source, unit, page reference, image type, key linguistic function, compositional cues, and preliminary intersemiotic relations). The complete inventory is provided in Appendix A, enabling traceability from interpretive claims to page-level evidence, in line with established qualitative standards for credibility and auditability (Lincoln & Guba, 1985).

3.4.6 Analytical Payoff

Taken together, these criteria produce a corpus that is fit for purpose in three ways. First, it ensures sufficient visual diversity and structural clarity to address RQ1 through systematic GVD-based analysis. Second, it secures explicit intersemiotic anchoring so that RQ2 can be examined as metafunctional co-realisation rather than parallel description. Third, by adopting the page ensemble as the unit of analysis, it preserves the designed integrity of multimodal meaning-making and supports traceable interpretation

of how EMP textbooks make medical meaning available to Vietnamese learners.

3.5 Data Collection Procedure

Data collection in this study was conducted as a theoretically guided and systematically staged process rather than as a mechanical extraction of textbook pages. Consistent with the theory-driven, interpretive-explanatory paradigm articulated in Section 3.1 and operationalised through the research design in Section 3.2, the procedure was structured to ensure that the collected data were analytically fit for examining visual organisation (RQ1) and visual-linguistic interaction (RQ2) within multimodal textbook design. To make the workflow explicit and replicable, the overall procedure is summarised in Figure 3.1 as a five-step sequence moving from corpus orientation to verification and traceability.

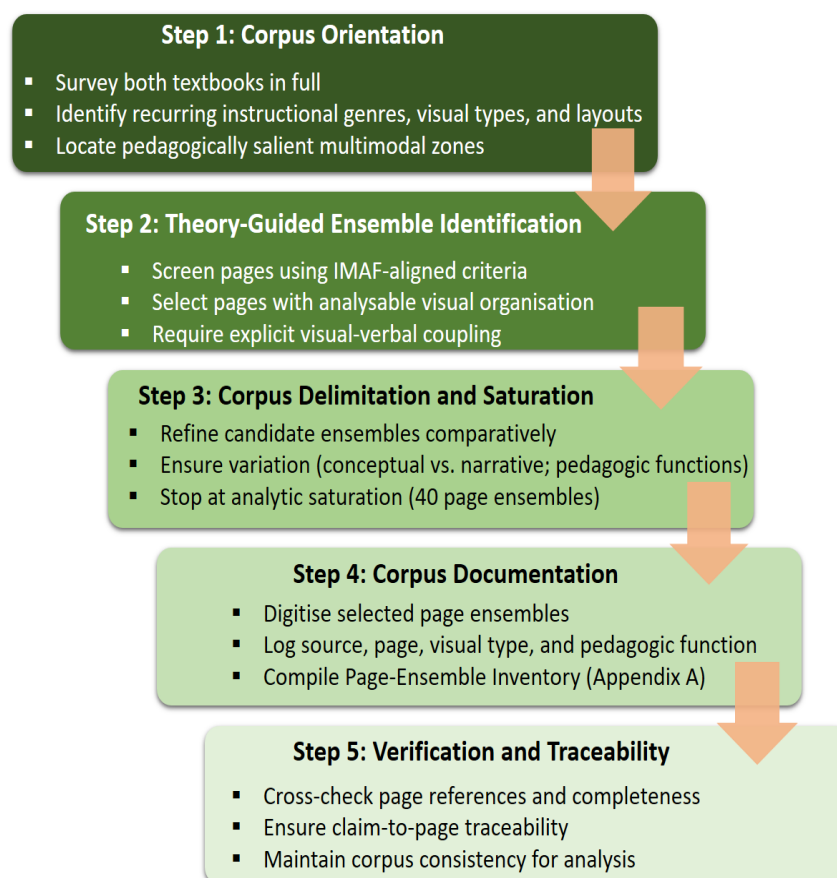


Figure 3.1 Data Collection Procedure

As shown in Figure 3.1, the procedure began with *Step 1: Corpus Orientation*, in which the full print versions of the two selected textbooks – *Medical Terminology: A Living Language* (Fremgen & Frucht, 2019) and *Career Paths: Medical* (Evans, Dooley, & Tran, 2012) – were surveyed in their entirety to establish an overview of their multimodal architecture. This familiarisation stage focused on mapping recurring instructional genres (e.g. system overviews, terminology pages, reading-visual sequences, dialogue-based tasks), dominant visual types, and typical page layouts. Rather than functioning as preliminary coding, this step provided an informed baseline for identifying where multimodal design is pedagogically mobilised across each textbook.

Building on this orientation, *Step 2: Theory-Guided Ensemble Identification* narrowed the material through a theory-led screening process aligned with the Integrated Multimodal Analytical Framework (Figure 2.1). At this stage, pages were provisionally marked when they displayed (i) salient visual organisation analysable under the GVD, and (ii) explicit visual-verbal coupling enabling metafunctional interpretation. Importantly, this screening was deliberately conservative: pages that were text-dominant, visually decorative, or analytically discontinuous were set aside so that the eventual corpus would foreground analysable intersemiotic design rather than incidental co-presence.

Following initial identification, *Step 3: Corpus Delimitation and Saturation* refined the candidate set through iterative comparison within and across the two textbooks. Here, ensembles were reviewed to ensure variation in representational type (conceptual versus narrative), pedagogic function (expository, classificatory, procedural, interactional), and layout configuration. This refinement continued until analytic saturation was reached – namely, when additional ensembles yielded recurring multimodal patterns without extending the analytical categories required for RQ1 and RQ2. The

outcome of this step was a balanced corpus of forty page ensembles, comprising twenty ensembles from each textbook.

Once the final corpus was delimited, *Step 4: Corpus Documentation* consolidated the dataset into an auditable research record. Each selected ensemble was digitised and logged in a corpus inventory matrix documenting textbook source, unit, page reference, visual type, pedagogic function, and brief analytical notes. This step served a dual purpose: it prepared the corpus for systematic coding and, at the same time, established transparent traceability from interpretive claims to page-level evidence. The completed inventory is reported in Appendix A (Page-Ensemble Inventory).

Finally, *Step 5: Verification and Traceability* ensured corpus integrity before analysis proceeded. Page references and inventory entries were cross-checked for accuracy and completeness, and digital files were stored in organised folders with consistent naming conventions to maintain stable links between (i) the archived page ensembles, (ii) the corpus log, and (iii) subsequent analytic outputs. In this way, verification functioned not as an administrative afterthought but as a methodological safeguard, reinforcing transparency and ensuring that the corpus could sustain rigorous, theory-driven explanation of multimodal meaning-making.

Taken together, these five steps demonstrate that data collection in this study was conducted independently of analysis but not independently of theory. By aligning corpus construction with the study's analytical frameworks and research questions from the outset, the procedure established a corpus that is methodologically coherent, analytically tractable, and fully traceable – thereby providing a robust empirical foundation for the multimodal analyses reported in the following sections.

3.6 Data Analysis

3.6.1 Data Analysis Procedure

Following Burns' (1999) view that data analysis involves both describing and explaining, the present study operationalised analysis as a staged movement from systematic multimodal description to theory-constrained interpretation. Accordingly, the workflow was organised into two phases: Phase A focused on descriptive coding using Kress and van Leeuwen's (2021) Grammar of Visual Design (GVD), while Phase B developed explanatory interpretations of image-text relations using Unsworth's (2008) multimodal metafunctional framework.

Stage 1: Dataset finalisation (raw page ensembles).

The dataset comprised forty multimodal page ensembles (twenty from each textbook), selected according to the criteria specified in Section 3.4. Each ensemble was treated as a coherent unit in which images, written language, layout, and typographic resources jointly contributed to pedagogic meaning.

Stage 2: Data preparation and analytical organisation.

All ensembles were digitised and organised in a corpus inventory matrix documenting textbook source, unit, page reference, dominant visual type, and pedagogic function. This stage established traceability between page-level evidence and subsequent coding outputs, forming an auditable analytic record.

Stage 3: Pilot coding and codebook calibration.

A pilot analysis of five ensembles sampled across both textbooks was conducted to test the clarity and applicability of the coding scheme. Pilot

coding refined category definitions, resolved boundary cases, and stabilised annotation conventions. Based on the pilot, a rubricised codebook was finalised to guide full-scale coding.

Stage 4: Descriptive coding of the full dataset (visual grammar).

The full dataset was coded using GVD categories to describe visual organisation in terms of: (i) representational meanings (narrative/conceptual structures; participants, processes, circumstances), (ii) interactive meanings (gaze, distance, angle, modality cues), and (iii) compositional meanings (information value, salience, framing). This stage addressed RQ1 by establishing a systematic account of how visual design functions within each page ensemble.

Stage 5: Explanatory interpretation (multimodal meaning-making).

Building on descriptive visual coding, the analysis proceeded to explanatory interpretation by modelling how visual and linguistic modes interacted across the ensemble. Drawing on Unsworth's (2008) multimodal metafunctional account, image-text relations were interpreted in terms of: (i) ideational relations (elaboration/extension/enhancement), (ii) interpersonal alignment (visual positioning with linguistic mood/modality), and (iii) textual organisation (layout, captioning, salience, and lexical cohesion shaping reading paths). Interpretations were formulated as theory-led explanations rather than impressionistic commentary and were grounded in explicit categories and page-referenced evidence.

In line with Burns'(1999) view that data analysis involves both describing and explaining, Figure 3.2 presents the staged analytical pathway adopted in this study, moving from systematic description to theory-constrained interpretation.

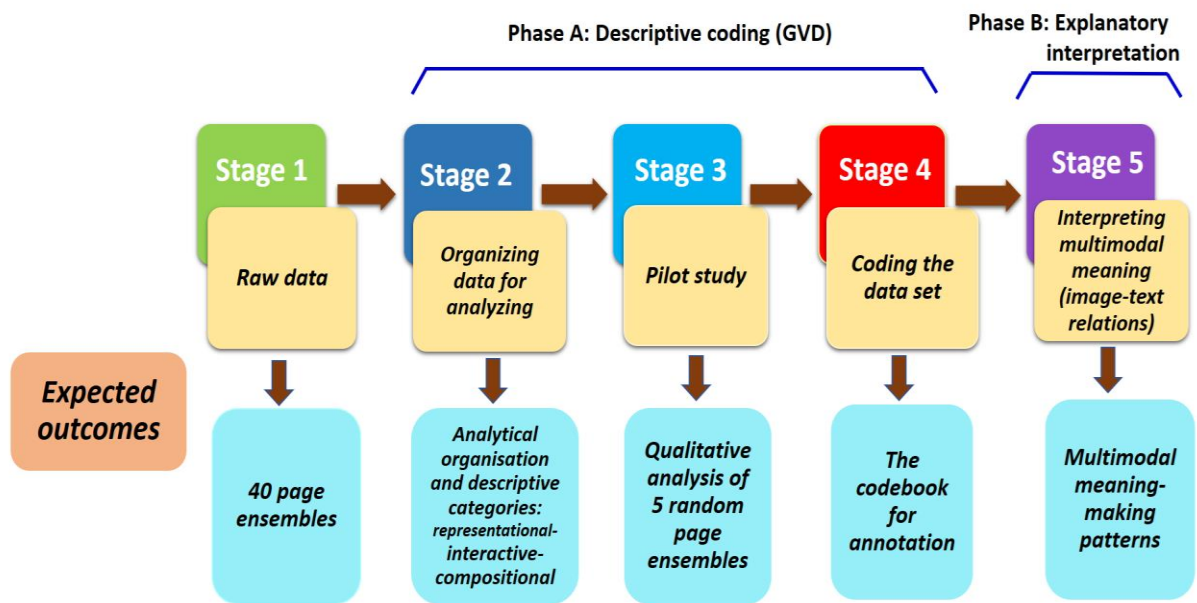


Figure 3.2 Data Analysis Procedure

As illustrated in Figure 3.2, the analysis was conducted in five sequential stages. Stages 1-4 constitute Phase A, where page ensembles were finalised, organised, piloted, and descriptively coded using the GVD systems of representational, interactive, and compositional meaning. Stage 5 forms Phase B, where the coded results were interpreted through multimodal metafunctions and intersemiotic relations to explain how visual-verbal orchestration constructs medical meaning. This staged design ensured that interpretation remained grounded in traceable descriptive evidence rather than impressionistic commentary. Overall, this staged procedure enabled the study to progress from structured description to explanatory interpretation, thereby aligning the analytical workflow with Burns’(1999) distinction between describing and explaining while maintaining theoretical coherence with social semiotic multimodal analysis.

3.6.2 Pilot Study: Trial Qualitative Analysis of Five Random Page Ensembles

Before coding the full dataset, a pilot study was conducted to test the feasibility and operationalisability of the IMAF at the level of the page ensemble, and to calibrate the rubricised codebook for consistent annotation. In multimodal discourse analysis, methodological robustness depends not only on adopting established theories but also on demonstrating that analytical categories can be applied systematically to authentic artefacts without collapsing into impressionistic description (Jewitt, 2014; Bezemer & Kress, 2016). This requirement is particularly salient in EMP textbook research because medical knowledge is frequently distributed across images, labels, layout, and specialised language, making it analytically risky to treat modes as parallel yet separate channels (Lemke, 1998; O'Halloran, 2008). In the Vietnamese context, EMP materials research has largely prioritised content relevance, linguistic coverage, and user perceptions, while multimodal organisation and visual-verbal coupling remain less systematically modelled. Consequently, a pilot trial was necessary to ensure that the present study's multimodal claims would be grounded in explicit categories and traceable page-based evidence rather than intuitive commentary (e.g., Nguyen, 2019; Le et al., 2021; Luu, 2021; Lam & Pham, 2024).

Sampling rationale and selection

Five page ensembles were randomly sampled across the two textbooks (three from *Medical Terminology: A Living Language* and two from *Career Paths: Medical*) as a pragmatic trial set for category testing and calibration. While the selection was random, the sampled pages nevertheless exhibited high variation in visual type and pedagogic function (e.g., labelled anatomical diagrams, medical photographs, tabulated terminology, procedural instruction

sequences, and brochure-like layouts). This combination is methodologically appropriate because a pilot in qualitative research is not intended to generate substantive findings, but to test whether the proposed analytical system can handle variation, ambiguity, and boundary cases that are typical of the corpus (Patton, 2015; Miles, Huberman, & Saldaña, 2019). The pilot therefore served as a controlled “*stress test*” of the framework’s descriptive and explanatory components across contrasting semiotic ecologies.

Analytical procedure and framework alignment

Consistent with Burns’(1999) distinction between describing and explaining, the pilot proceeded in two aligned phases, corresponding to the IMAF’s descriptive and interpretive strata.

Phase A: Descriptive trial coding (GVD)

Each sampled ensemble was provisionally coded using Kress and van Leeuwen’s (2021) GVD to establish a disciplined description of visual organisation. Coding focused on the three metafunctional systems: (i) representational meanings (narrative vs. conceptual structures; part-whole and taxonomic relations; participants and circumstances), (ii) interactive meanings (gaze, social distance, angle, and modality cues), and (iii) compositional meanings (information value, salience, framing, and reading paths). This phase operationalised the study’s commitment to “*category-led*” description, ensuring that visual analysis was conducted with a transparent metalanguage rather than with informal aesthetic judgement (Kress & van Leeuwen, 2021; Bezemer & Kress, 2016).

Phase B: Explanatory trial interpretation (intersemiotic relations).

Building on Phase A, each ensemble was then examined for demonstrable image-text coupling through Unsworth’s (2008) multimodal metafunctional account, supported where relevant by established intersemiotic

models (e.g., Martinec & Salway, 2005; Royce, 2007). Intersemiotic relations were interpreted at three aligned levels: (i) ideational relations (elaboration, extension, enhancement), (ii) interpersonal alignment (correspondences between visual positioning - distance/angle/gaze - and linguistic stance such as mood and modality), and (iii) textual organisation (layout and salience in relation to sequencing, cohesion, and instructional flow). Importantly, linguistic descriptors (e.g., transitivity, mood/modality, Theme–Rheme) were used selectively and only when they could be explicitly mapped onto visual evidence within the same ensemble. This design choice prevented analytical inflation and maintained coherence with the unit of analysis (the page ensemble), as recommended in multimodal scholarship that treats meaning as orchestrated across modes rather than distributed into independent modal tracks (van Leeuwen, 2005; Jewitt, 2014, Patton, 2015).

Outcomes and codebook calibration

The pilot did not aim to produce reportable results about the textbooks. Instead, it generated methodological outputs that strengthened the subsequent full-scale analysis in four ways.

First, it verified the analytical adequacy of the page ensemble as a unit of analysis by showing that multimodal meanings were not exhausted by isolated images or paragraphs but were routinely completed through captioning, labelling, typographic hierarchy, and layout sequencing. This finding supports the study's ensemble-based design and aligns with established multimodal approaches to educational texts (Bezemer & Kress, 2016).

Second, it exposed category boundary cases such as visuals that combine conceptual classification with procedural implication, or captions that operate simultaneously as elaboration and enhancement - thereby

enabling refinement of category definitions and decision rules. In line with qualitative best practice, these refinements were documented as part of an audit trail to ensure that later interpretations remained consistent and theoretically constrained (Lincoln & Guba, 1985; Miles, Huberman, & Saldaña, 2019, Kress & van Leeuwen, 2001).

Third, it stabilised annotation conventions (e.g., what counts as the focal image unit within a multi-image page; how to record salience and framing in pages with layered design; how to log evidence for intersemiotic coupling). Stabilising such conventions is crucial for multimodal research, where analytic transparency depends on replicable coding routines rather than on researcher intuition (Jewitt, 2014; Bezemer & Kress, 2016).

Finally, the pilot confirmed that the explanatory phase (intersemiotic interpretation) could be conducted as theory-led modelling rather than as parallel description of image and language. This point is methodologically consequential for EMP research in Vietnam, where materials evaluations have tended to privilege linguistic criteria and under-specify how visuals and language jointly construct disciplinary meaning. The pilot therefore provided an empirically grounded justification for the study's integrated framework and the subsequent full-corpus coding strategy (e.g., Hoang, 2015; Le, 2017; Le et al., 2021; Lam & Pham, 2024).

In sum, the pilot study functioned as a methodological calibration stage that ensured the IMAF was operational, the codebook was internally coherent, and the subsequent analytic claims would be auditable and evidence-based. By establishing that the framework can handle variation and resolve boundary cases across contrasting textbook designs, the pilot strengthened the credibility and transparency of the full dataset analysis reported in the

following sections (Lincoln & Guba, 1985; Miles, Huberman, & Saldaña, 2019).

2.6.3 Integrated Multimodal Analytical Framework

To explain how visual and linguistic modes interact to construct medical meaning in English medical textbooks used in Vietnam, this study adopts an IMAF. Importantly, the IMAF is not proposed as a new theory, but as a theory-led analytical apparatus that synthesises established frameworks in social semiotics into a coherent, operational system for empirical analysis at the level of the page ensemble.

Specifically, the IMAF integrates two complementary strands of multimodal scholarship:

- (i) the GVD as a systematic metalanguage for analysing visual organisation (Kress & van Leeuwen, 2021);
- (ii) models of intersemiotic relations and multimodal metafunctions, which explain how visual and linguistic meanings are coupled within the same pedagogic artefact (Unsworth, 2008).

In this sense, the IMAF functions as the methodological bridge between the study's interpretive-explanatory paradigm (Section 3.1), its research design (Section 3.2), and the empirical analysis of textbook pages. It ensures that interpretation is not impressionistic or intuitive, but theoretically constrained, systematically traceable, and analytically auditable.

Figure 3.3 below visualises the Integrated Multimodal Analytical Framework as a three-stratum architecture, progressing from visual semiotic description, through intersemiotic explanation, to pedagogic-disciplinary interpretation.

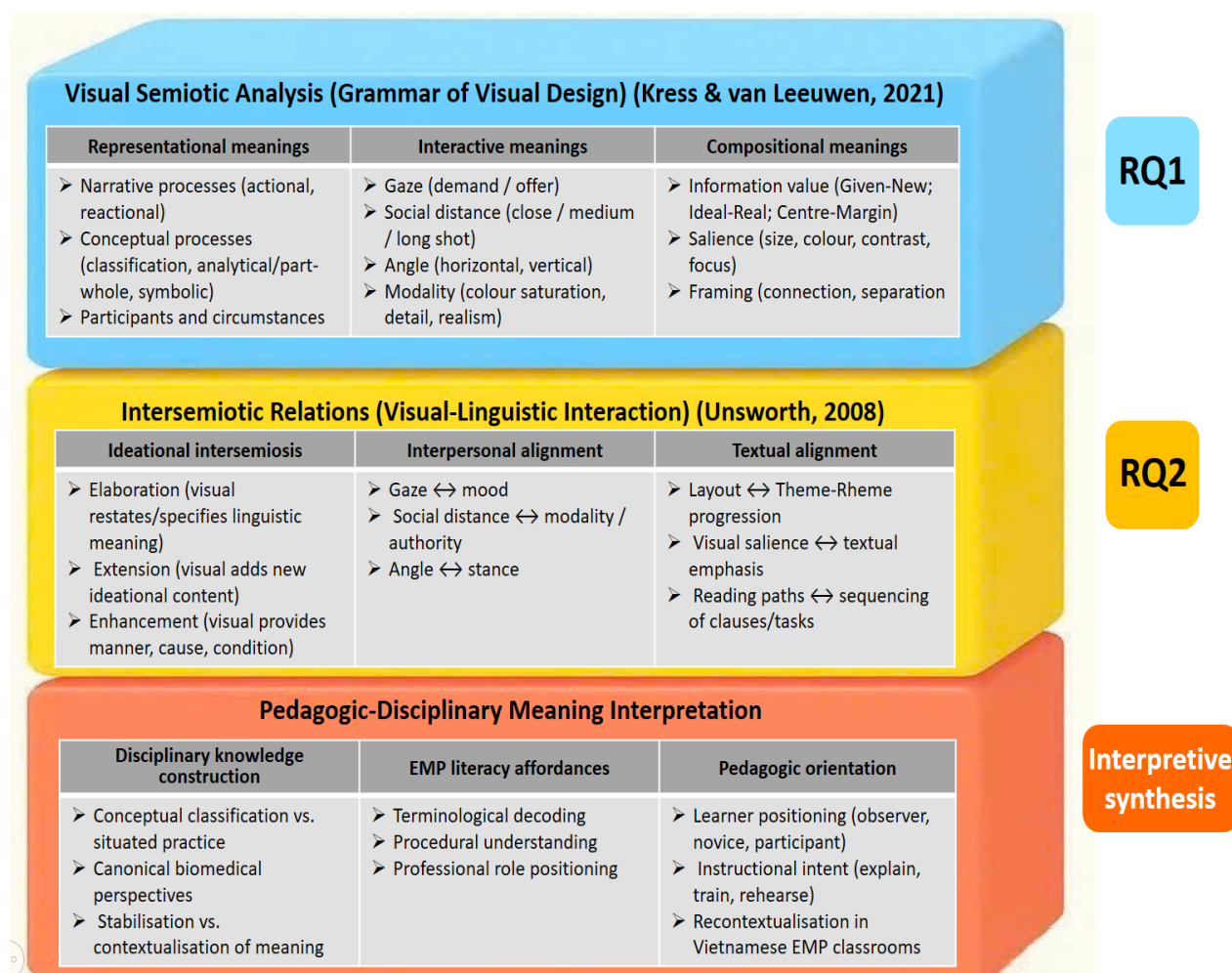


Figure 3.3 Integrated Multimodal Analytical Framework

As shown in Figure 3.3, the IMAF is organised into three analytically distinct but conceptually interdependent strata, each of which performs a specific explanatory function within the overall framework.

At the first stratum, the framework conducts Visual Semiotic Analysis, drawing explicitly on the GVD (Kress & van Leeuwen, 2021). This stratum addresses RQ1 and provides a disciplined account of how images are organised internally. Visual meaning is analysed through three interrelated systems.

Representational meanings examine what kinds of medical entities, processes, and relations are depicted, distinguishing between narrative processes (such as actional and reactional structures) and conceptual processes (including classification, analytical part-whole relations, and symbolic representations), together with the participants and circumstances involved.

Interactive meanings address how viewer-image relations are staged, focusing on gaze (demand versus offer), social distance (close, medium, long shot), angle (horizontal and vertical), and modality cues such as colour saturation, detail, and degree of realism.

Compositional meanings examine how visual elements are arranged to guide interpretation, including information value (Given-New, Ideal-Real, Centre-Margin), salience (size, colour, contrast, focus), and framing (connection versus separation).

However, the IMAF does not treat visual organisation as self-sufficient. While this stratum establishes what the visual design is doing, explanation requires moving beyond description to account for how visual meanings are integrated with language.

This is the focus of the second stratum, which examines Intersemiotic Relations between visual and linguistic modes, drawing on Unsworth's (2008) multimodal metafunctional model and established accounts of intersemiotic relations (e.g., Martinec & Salway, 2005; Royce, 2007). Here, interaction is not assumed but demonstrated through identifiable relations.

At the ideational level, visual-linguistic coupling is analysed in terms of elaboration (where one mode restates or specifies the other), extension (where one mode adds new ideational content), and enhancement (where one mode

provides circumstantial information such as manner, cause, sequence, or condition).

At the interpersonal level, alignment is examined through correspondences between visual choices and linguistic stance – for example, gaze in relation to mood, social distance in relation to modality and authority, and angle in relation to evaluative positioning.

At the textual level, intersemiotic alignment is analysed through correspondences between layout and thematic organisation (Theme-Rheme patterns), visual salience and textual emphasis, and reading paths in relation to the sequencing of clauses or tasks.

Critically, Linguistic features (e.g., transitivity, mood/modality, Theme-Rheme) were only coded insofar as they enabled explicit mapping with visual evidence within the same page ensemble, thereby maintaining analytic alignment with the study's unit of analysis and avoiding parallel monomodal coding.

The third stratum of the IMAF moves from semiotic explanation to Pedagogic-Disciplinary Meaning Interpretation. At this level, the framework interprets how multimodal configurations function as epistemic resources for constructing medical knowledge and professional meaning. Analysis focuses on three interrelated dimensions:

First, how disciplinary knowledge is constructed, for instance through conceptual classification versus situated clinical practice, canonical biomedical perspectives, and the stabilisation or contextualisation of meaning;

Second, what kinds of EMP literacy affordances are made available, including terminological decoding, procedural understanding, and professional role positioning; and

Third, how pedagogic orientation is realised, in terms of learner positioning (observer, novice, participant), instructional intent (explaining, training, rehearsing), and recontextualisation within Vietnamese EMP classrooms (van Leeuwen, 2005).

Importantly, this pedagogic-disciplinary stratum does not claim to measure learning outcomes. Rather, it provides an interpretive account of how learning is designed to be made possible through multimodal orchestration. This orientation aligns with both international multimodal research on textbook discourse (e.g. Lemke, 1998; O'Halloran, 2008; Bezemer & Kress, 2016) and Vietnamese EMP scholarship, which has increasingly emphasised the need to examine not only linguistic content but also visual and pedagogic mediation in imported instructional materials.

Taken together, the IMAF enables the study to move systematically from visual description, through intersemiotic explanation, to pedagogic interpretation, without collapsing these levels or treating them as independent. By specifying what counts as analysable evidence at each stratum, the IMAF provides a coherent analytical language for answering RQs and ensures that claims about multimodal interaction are grounded in explicit categories, cross-modal mapping, and page-referenced data. In doing so, the framework operationalises the study's theoretical commitments and renders the analysis of multimodal medical textbooks both rigorous and epistemically accountable.

3.6.4 Coding and Analytical Phases

Situated within the interpretive-explanatory research paradigm articulated in Section 3.1 and operationalised through the IMAF presented in Section 3.7 (Figure 3.2), the present study conceptualises coding not as a neutral technical procedure but as a theoretically disciplined interpretive practice. In social semiotic and multimodal discourse traditions, coding is understood as an

epistemic act through which semiotic patterns are rendered analytically visible and explanatorily meaningful, rather than mechanically extracted from texts (Jewitt, 2014; Fairclough, 1995). Consequently, the analytical process adopted in this study is recursive, stratified, and theory-driven, rather than linear or sequential.

Importantly, the coding and analytical phases are organised in direct alignment with the stratal logic of the IMAF. This alignment ensures that analytical decisions are not made ad hoc, but are systematically constrained by established multimodal theory. In particular, the study deliberately avoids treating language as an autonomous object of analysis. Linguistic resources are examined only insofar as they participate in intersemiotic relations with visual design. This methodological decision reflects the study's central concern with multimodal interaction rather than parallel modal description, and prevents the fragmentation of meaning across modes (Bezemer & Kress, 2016).

The analytical process unfolds through three interrelated and recursively connected phases, each corresponding to one stratum of the IMAF and collectively addressing the overarching research question and its two subordinate questions (RQ1 and RQ2).

Phase 1: Visual Semiotic Coding (RQ1)

The first analytical phase focuses on visual semiotic coding and corresponds to Stratum 1 of the IMAF. Anchored in the GVD developed by Kress and van Leeuwen (2021), this phase establishes a principled account of how visual images are organised to construe medical meaning. Each page ensemble is analysed as a designed visual artefact rather than a transparent representation of biomedical reality. From this perspective, images are treated as epistemic resources that select, hierarchise, and stabilise particular ways of knowing the medical domain (Lemke, 1998).

Visual coding proceeds through the three metafunctional domains specified in GVD. Representational analysis identifies whether images construe narrative processes – such as actional or reactional structures involving vectors, actors, and goals – or conceptual processes, including classificatory, analytical (part-whole), and symbolic relations. Particular attention is paid to how anatomical entities, procedures, and instruments are configured as participants within these structures, and to how circumstantial elements contribute to disciplinary explanation.

Interactive analysis examines how viewer-image relations are staged through gaze (demand versus offer), social distance (close, medium, or long shot), angle (horizontal or vertical), and modality cues such as colour saturation, detail, and degrees of realism. These features are interpreted as semiotic mechanisms that position learners epistemically—whether as detached observers, guided novices, or legitimate peripheral participants in medical practice (Kress & van Leeuwen, 2021).

Compositional analysis investigates how information is spatially organised through information value (Give-New, Ideal-Real, Centre-Margin), salience, and framing. This level of analysis reveals how visual layout directs reading paths, foregrounds canonical knowledge, and stabilises interpretive priorities within the page ensemble. Importantly, Phase 1 does not yet explain interaction between image and language; rather, it establishes a rigorous baseline of visual meaning potential upon which subsequent intersemiotic interpretation depends.

Phase 2: Intersemiotic Coding and Metafunctional Alignment (RQ2)

Building on the visual foundation established in Phase 1, the second analytical phase corresponds to Stratum 2 of the IMAF and operationalises RQ2 through intersemiotic coding. Drawing on Unsworth's (2008) model of multimodal metafunctions and the intersemiotic frameworks developed by

Martinec and Salway (2005) and Royce (2007), this phase examines how visual and linguistic resources are coupled to jointly realise meaning across modes.

Crucially, language enters the analysis at this stage only as a relational resource. Clauses, captions, labels, and instructional text are examined insofar as they interact with visual design, not as independent linguistic systems. This approach reflects a core social semiotic assumption: meaning in multimodal texts emerges through orchestration and alignment across modes rather than through additive accumulation (Jewitt, 2014; Bezemer & Kress, 2016).

Intersemiotic relations are identified through three principal mechanisms. Elaboration occurs when one mode restates, specifies, or identifies the meaning realised by another, as in labelled anatomical diagrams where linguistic naming anchors visual classification. Extension is identified when one mode contributes new but related information that the other does not provide, such as when explanatory text extends the functional implications of a diagram. Enhancement captures cases where one mode supplies circumstantial meaning – temporal, causal, procedural, or conditional – that frames the interpretation of the other mode.

These relations are interpreted through metafunctional alignment. Ideational alignment is evidenced when visually represented entities are realised linguistically as participants in definitional or classificatory clauses. Interpersonal alignment is identified when viewer positioning encoded visually coheres with directive or advisory language. Textual alignment is examined through correspondences between layout organisation and information flow, such as when visual salience converges with reading paths to guide interpretive emphasis. Through this layered analysis, interaction between modes is demonstrated analytically rather than asserted impressionistically.

Phase 3: Interpretive Synthesis of Pedagogic-Disciplinary Meaning (The overarching research question)

The third analytical phase corresponds to Stratum 3 of the IMAF and functions as an interpretive synthesis rather than a coding operation. At this stage, findings from visual and intersemiotic analyses are recontextualised to address the overarching research question: how visual and linguistic modes interact to construct medical meaning and disciplinary knowledge in English medical textbooks used in Vietnam.

Drawing on van Leeuwen's (2005) concept of recontextualisation, this phase interprets how multimodal configurations stabilise canonical biomedical perspectives while mediating knowledge for novice learners. The analysis focuses on how disciplinary knowledge is constructed – whether through abstract conceptual classification or situated representations of practice – and how such constructions afford particular forms of EMP literacy, including terminological decoding, procedural understanding, and professional role positioning. Importantly, this phase does not evaluate learning outcomes or biomedical accuracy; rather, it explains how pedagogic meaning is semiotically made available through patterned multimodal design.

Interpretive Rigour and Calibration across the Corpus

Across all phases, the analytical process is explicitly recursive, reflecting a hermeneutic spiral of understanding in which insights generated at one level prompt refinement at others (Jewitt et al., 2001). To support interpretive rigour and transparency, reflexive analytic memos and an audit trail were maintained, documenting coding decisions, theoretical tensions, and alternative readings (Lincoln & Guba, 1985; Miles et al., 2014).

To maintain interpretive consistency across the forty page ensembles while avoiding reductive quantification, the study employs a calibrated interpretive salience scale (0-3). This scale functions as a methodological

device for stabilising judgments of semiotic prominence across visual and intersemiotic analyses.

Table 3.1 Interpretive Saliency Calibration Scale (0-3)

Score	Descriptor	Interpretive Meaning	Illustrative Interpretation at Page-Ensemble Level
0	Not Realised	No evidence of the targeted visual or intersemiotic feature	An anatomical diagram appears without labels, captions, or textual reference, preventing visual-verbal integration.
1	Minimally Realised	Feature present but weak or fragmented	Image and text co-occur on the same page, but referential links are implicit or ambiguous, leaving interpretation largely inferential.
2	Functionally Realised	Feature supports meaning construction clearly	Visual structures and accompanying text align sufficiently to explain a medical concept or procedure in a stable manner.
3	Strongly and systematically realised	Feature integrated to stabilise disciplinary meaning	Visual organisation, captions, terminology, and layout jointly construct a stepwise conceptualisation of an anatomical system or clinical process.

Importantly, this scale does not transform qualitative interpretation into numerical measurement; rather, it provides a shared analytical reference point through which degrees of multimodal realisation can be compared across page ensembles.

To make explicit how the analytical phases operationalise the research questions, Table 3.2 maps each research question onto its corresponding analytical focus and coding phase.

Table 3.2 Mapping Research Questions to Analytical Focus and Coding Phases

Research Question	Analytical Focus	IMAF Component Activated	Coding / Analytical Phase	Analytical Output
Overarching				
research RQ: How do visual and linguistic modes interact to construct medical meaning and disciplinary knowledge?	Overall multimodal meaning construction through design	Integrated interpretation across visual and intersemiotic strata	Recursive synthesis across Phases 1-3	Explanatory account of multimodal knowledge construction at page-ensemble level
RQ1: How are visual images organised representationally, interactively, and compositionally?	Internal organisation of visual meaning	Visual Semiotic Analysis	Phase 1 – Visual Analytical Coding	Systematic description of visual meaning patterns and epistemic functions
RQ2: How do visual and linguistic modes jointly realise ideational, interpersonal, and textual meanings?	Visual-linguistic coupling mechanisms	Intersemiotic Relations	Phase 2 – Intersemiotic Interpretation	Typology of elaboration, extension, and enhancement across page ensembles

This mapping demonstrates that the analytical process is systematically aligned with the research questions and proceeds through an ordered progression from visual description to intersemiotic interpretation and pedagogic explanation.

In sum, the coding and analytical phases operationalise the IMAF as a coherent explanatory architecture. Visual meaning is first stabilised through disciplined semiotic description; intersemiotic relations are then systematically mapped to demonstrate multimodal interaction; and finally, pedagogic-disciplinary meanings are interpretively synthesised. Through this progression, the analysis moves beyond inventory and description toward theoretically warranted explanation, ensuring that claims about multimodal meaning-making in English medical textbooks are epistemically accountable, analytically transparent, and empirically grounded.

3.6.5 Rubricized Multimodal Codebook

Situated within the interpretive-explanatory paradigm articulated in Section 3.1 and operationalised through the IMAF in Section 3.6, this study adopts a Rubricized Multimodal Codebook as its primary mechanism for translating multimodal theory into analytically accountable evidence. The codebook is not introduced as a technical add-on to analysis; rather, it functions as an epistemic stabiliser that constrains interpretation, renders analytic decisions transparent, and enables systematic comparison across a heterogeneous multimodal corpus.

From a social semiotic perspective, multimodal interpretation without explicit calibration risks collapsing into impressionistic description, particularly in visually dense pedagogic artefacts such as medical textbooks

(Jewitt, 2014; Kress & van Leeuwen, 2021). The rubricised structure of the codebook therefore performs a critical methodological function: it specifies which semiotic features are analytically admissible, how they are to be interpreted within the logic of the IMAF, and how degrees of semiotic salience may be meaningfully distinguished without resorting to statistical abstraction.

Importantly, the organisation of the codebook mirrors the stratal architecture of the IMAF. Each rubric module corresponds to a theoretically distinct but analytically interdependent stratum, ensuring coherence between research questions, analytical focus, and empirical evidence.

3.6.5.1 Module A: Visual Meaning (Visual Semiotic Analysis)

The first module operationalises Stratum 1 of the IMAF, providing a principled framework for analysing visual images as semiotic designs rather than as transparent depictions of biomedical reality. In line with the Grammar of Visual Design (Kress & van Leeuwen, 2021), visual meaning is understood as metafunctionally organised and pedagogically oriented. Images are therefore treated as epistemic artefacts that actively construct disciplinary knowledge by selecting, foregrounding, and stabilising particular representational logics (Lemke, 1998).

To render this visual organisation analytically tractable across the corpus, Table 3.3 formalises the core visual metafunctions – representational, interactive, and compositional – into explicit analytical dimensions and observable indicators.

Table 3.3 Module A: Visual Meaning Rubric

Metafunction	Analytical Dimension	Descriptors / Coding Indicators	Typical Realisation in Medical Textbooks
Representational meaning	Process type	Narrative (actional/reactional) vs. conceptual (classification; analytical/part-whole; symbolic)	Labelled anatomical diagrams; schematic cross-sections; contextual scenes
	Participants & circumstances	Participant roles; relevant circumstances (spatial/functional cues)	Organs, tissues, instruments; arrows, callouts, sectional planes
Interactive meaning	Gaze/contact	Demand vs. offer	Images inviting engagement vs. images positioning learner as observer
	Social distance	Close/medium/long shot	Precision close-ups vs. contextual mid-shots
	Angle	Horizontal (involvement)/vertical (power)	Neutral/objectifying angles vs. learner-inclusive angles
	Modality cues	Colour saturation, detail, realism	Anatomical realism vs. schematic abstraction
Compositional meaning	Information value	Given-New; Ideal-Real; Centre-Margin	Central image + marginal labels; ideal overview + real detail
	Salience	Size, contrast, colour, focus	Highlighting key structures; numbered sequences
	Framing	Connection separation vs.	Hard segmentation vs. integrated text-image layouts

Following its presentation, Table 3.3 should be understood not as a checklist of image features but as a conceptual map of visual meaning potential. It enables the analysis to distinguish, for example, between images that construe medical knowledge through classificatory abstraction and those that enact narrative processes of clinical action; between visuals that position learners as detached observers and those that invite interpersonal engagement; and between layouts that stabilise canonical knowledge and those that scaffold procedural understanding. In this way, Module A establishes the visual baseline required to address RQ1 and prepares the ground for subsequent intersemiotic interpretation.

3.6.5.2 Module B: Linguistic Resources in Intersemiotic Alignment

Consistent with the design-focused scope of the dissertation and the architecture of the IMAF, linguistic analysis is deliberately delimited. Language is not examined as an autonomous semiotic system but as a relational resource whose analytical relevance emerges only through its interaction with visual design. This decision reflects a core assumption of multimodal social semiotics: meaning in multimodal texts arises through orchestration across modes rather than through parallel modal description (Bezemer & Kress, 2016).

Within this constrained but theoretically motivated scope, Table 3.4 specifies those linguistic resources that recurrently participate in visual-linguistic alignment in EMP textbooks.

Table 3.4 Module B: Linguistic Resources in Intersemiotic Alignment

Alignment function	Linguistic resource	What is coded	Why it matters for IMAF/ RQ2
Ideational anchoring	Labels & naming	Technical labels, part names, terminology strings	Enables mapping of visual participants to verbal naming (core for elaboration)
	Definitional wording	“X is/are...”, “refers to...”, classificatory phrasing	Anchors conceptual relations shown visually; supports ideational coupling
Interpersonal alignment	Directive/ advisory wording	Imperatives; “should/must/need to”; advice frames	Can align with viewer positioning (distance/angle) and pedagogic stance
	Evaluative stance (explicit)	“may indicate”, “typically”, warnings/cautions	Supports claims about authority, certainty, and learner positioning
Textual coordination	Figure referencing & callouts	“see Figure...”, “as shown...”, callout links	Provides explicit cross-modal ties (traceable evidence of coupling)
	Sequencing/ procedural markers	Sequencing/ procedural markers	Sequencing/ procedural markers

The analytical value of Table 3.4 lies in its selective logic. By foregrounding definitional structures, technical nominal groups, mood and modality choices, and information flow patterns, the table isolates linguistic features that directly anchor, extend, or frame visual meaning. Linguistic resources not implicated in intersemiotic alignment are intentionally excluded. This ensures that language enters the analysis only insofar as it contributes to multimodal meaning-making, thereby maintaining analytic focus on interaction rather than modal autonomy.

3.6.5.3 Module C: Intersemiotic Coupling (Visual-Linguistic Interaction)

While Modules A and B establish the semiotic resources available within each mode, Module C constitutes the explanatory core of the codebook. It operationalises the overarching research question and RQ2 by specifying how visual and linguistic meanings are coupled within the page ensemble. Drawing on the work of Martinec and Salway (2005), Royce (2007), and Unsworth (2008), intersemiotic relations are treated as structured mechanisms through which multimodal interaction becomes analytically demonstrable.

These mechanisms are formalised in Table 3.5, which delineates recurrent patterns of elaboration, extension, enhancement, and projection across the corpus.

Table 3.5. Module C: Intersemiotic Coupling Rubric

Relation type	Diagnostic question	Operational description	Typical textbook pattern
Elaboration	Does one mode specify/identify the other?	Language names/clarifies what is visually depicted (or vice versa)	Labelled diagrams + definitional captions
Extension	Does one mode add new related content?	One mode contributes additional ideational content not present in the other	Explanatory text adds function/usage beyond the diagram
Enhancement	Does one mode provide circumstantial framing?	Sequence, cause, condition, location, manner that frames interpretation	Stepwise procedures; task instructions framing images
Projection	Is speech/thought represented across modes?	Verbal exchange (spoken/thought) anchored by contextual visuals	Dialogue tasks supported by situational images

Rather than functioning as interpretive labels applied post hoc, the relations specified in Table 3.5 operate as diagnostic categories. They enable the analysis to show, with empirical precision, how images and language jointly realise ideational, interpersonal, and textual meanings. For instance, elaboration foregrounds terminological stabilisation through visual–verbal redundancy, while enhancement captures the procedural and contextual scaffolding characteristic of scenario-based medical discourse. In this way, Module C transforms the abstract notion of “*multimodal interaction*” into a set of traceable analytical relations.

3.6.5.4 Interpretive Salience Calibration Scale

Given the diversity of visual genres and pedagogic designs represented in the forty page ensembles, interpretive consistency constitutes a central methodological concern. To address this, the study incorporates an Interpretive Salience Calibration Scale (0-3), presented in Table 3.6.

Table 3.6. Interpretive Salience Calibration Scale (0-3)

Score	Descriptor	Interpretive meaning in this study
0	Not realised	No interpretable evidence of the feature in the ensemble
1	Minimal / fragmented	Present but weak, partial, inconsistent, or not functionally integrated
2	Functional	Clearly contributes to meaning-making and supports interpretive claims
3	High / integrated	Systematic, prominent, and central to the ensemble’s instructional meaning

The calibration scale does not function as a quantification tool or as a proxy for measurement. Instead, it serves an epistemic role by enabling the analysis to differentiate between incidental semiotic co-presence and

pedagogically consequential multimodal design. By calibrating degrees of salience rather than mere occurrence, the scale supports systematic comparison across page ensembles and across textbooks, while remaining fully compatible with the interpretive orientation of the study.

3.6.5.5 Full Rubricized Multimodal Codebook

The Full Rubricized Multimodal Codebook, presented in Appendix B, constitutes a core methodological component of this study rather than supplementary documentation. It operationalises the IMAF at the level of empirical data, ensuring that all interpretive claims advanced in Chapter 4 are grounded in systematic, transparent, and theoretically constrained analysis.

Within a social semiotic approach, multimodal meaning does not reside in isolated images or texts, but emerges through patterned configurations across a corpus (Kress, 2010; Jewitt, 2014). Accordingly, the full codebook records how each of the forty page ensembles instantiates visual semiotic features (Stratum 1) and intersemiotic relations between image and language (Stratum 2). This corpus-wide documentation prevents impressionistic interpretation and enables explanation to be derived from recurrent semiotic patterns rather than selected examples.

Each page ensemble is coded using the same rubric structure, capturing (i) dominant visual representational, interactive, and compositional features; (ii) linguistic resources insofar as they participate in intersemiotic alignment; and (iii) the type and strength of image-text relations. A calibrated salience scale (0-3) is applied consistently to support interpretive comparability across cases, without reducing qualitative analysis to quantification.

Epistemically, the full codebook performs three key functions. First, it stabilises interpretation by constraining analytical judgement through explicit categories derived from established multimodal theory (Kress & van

Leeuwen, 2021; Unsworth, 2008). Second, it supports analytic generalisation, allowing claims about multimodal meaning-making to be traced across the corpus rather than inferred from isolated illustrations (Miles, Huberman, & Saldaña, 2014). Third, it ensures auditability, enabling readers to verify how specific analytical claims in Chapter IV are grounded in page-level evidence.

For reasons of coherence and readability, the full codebook is located in Appendix B, while Chapter III presents the conceptual logic and rubric architecture. This separation allows the methodology chapter to foreground epistemic justification, while preserving complete transparency of analytical procedures. In this way, the Rubricized Multimodal Codebook functions as the empirical backbone of the study, translating multimodal theory into a disciplined and reproducible analytical practice aligned with the overarching research question and subordinate questions RQ1 and RQ2.

3.7 Reliability, Validity, and Ethical Considerations

Within the interpretive-explanatory paradigm adopted in this study, trustworthiness is understood not as a set of procedural checkpoints but as an emergent property of coherent research design. In contrast to positivist notions of validity and reliability grounded in replication and measurement, interpretive multimodal research locates rigor in the transparency, reflexivity, and theoretical accountability of the analytical process (Lincoln & Guba, 1985; Mikan, 2017). This section therefore consolidates how trustworthiness is established across the study through the alignment of epistemological stance, analytical framework, and procedural safeguards.

Crucially, trustworthiness in this research is distributed across all stages of inquiry - from corpus construction and analytical framework design to coding, interpretation, and representation. As such, it cannot be reduced to isolated techniques. Instead, it arises from the sustained coherence between the study's ontology of meaning (multimodal, socially constructed, and pedagogically

situated), its analytical machinery (Figures 2.1, 2.2, 3.2, and 3.3), and its reflexive methodological practices.

3.7.1 Interpretive Orientation and Epistemic Coherence

The study is grounded in an interpretivist–constructivist ontology, which treats meaning not as an objective property of texts but as semiotically constructed through socially and culturally situated design choices (Jewitt, 2014; Kress, 2010). Within this orientation, images, language, and layout in medical textbooks are understood as epistemic resources that actively shape what counts as disciplinary knowledge and how learners are positioned in relation to that knowledge.

Accordingly, trustworthiness depends on the extent to which analytical claims are theoretically warranted, contextually grounded, and transparently traceable to empirical evidence. The IMAF functions as a central mechanism in this regard, constraining interpretation through established social semiotic categories rather than allowing impressionistic or intuitive readings. This theoretical discipline provides the epistemic backbone upon which subsequent analytical decisions rest.

3.7.2 Reliability as Reflexive Consistency

Within interpretive research, reliability is not equated with inter-rater agreement or procedural replication. Instead, it is conceptualised as reflexive consistency – the degree to which analytical decisions are coherent, stable, and theoretically aligned over time (Tracy, 2010). In this study, reliability is ensured through the systematic application of the IMAF across all page ensembles and through the maintenance of a detailed audit trail documenting coding decisions, category refinement, and interpretive shifts (Miles, Huberman, & Saldaña, 2014).

Throughout the analytical process, reflexive memos were used to record emerging interpretations, uncertainties, and alternative readings. These memos functioned not as ancillary notes but as integral components of the

analytical workflow, enabling continuous scrutiny of the researcher's assumptions and interpretive stance. In this way, reliability is achieved not by eliminating subjectivity but by rendering it visible and accountable.

3.7.3 Validity as Interpretive Sufficiency and Theoretical Saturation

Validity in this study is conceptualised as interpretive sufficiency rather than representational accuracy. That is, an interpretation is considered valid when it provides a theoretically coherent and empirically grounded explanation of how meaning is constructed, without claiming exhaustiveness or generalisability beyond the analytical scope (Lincoln & Guba, 1985).

Methodological triangulation across multimodal strata – visual semiotic analysis, intersemiotic relations, and pedagogic interpretation – supports validity by enabling convergent explanation through multiple semiotic systems (Denzin, 2012). Theoretical saturation was reached when additional page ensembles no longer yielded new patterns of visual organisation or intersemiotic coupling relevant to RQ1 and RQ2. At this point, interpretive categories were considered sufficiently stabilised to support explanatory claims.

3.7.4 Ethical Integrity and Cultural Sensitivity

Ethical considerations in this study extend beyond formal compliance with copyright regulations to encompass intellectual integrity, cultural sensitivity, and responsible representation. Given the pedagogic significance of English medical textbooks in Vietnamese higher education, care was taken to avoid imposing ethnocentric assumptions or evaluative judgments about biomedical content or instructional quality.

All visual materials were accurately attributed, and analytical excerpts were used solely for scholarly purposes. Moreover, interpretations were framed with reflexive awareness of the researcher's positionality, recognising the asymmetries inherent in analysing global textbook materials within local

educational contexts (van Leeuwen, 2005). Ethics, in this sense, functions not merely as a procedural constraint but as an epistemic boundary shaping how meaning is interpreted and represented.

3.7.5 Consolidation: Trustworthiness as an Emergent Property of Design

In aggregate, the study's trustworthiness emerges from the coherence between its ontology of meaning (constructed, multimodal, contextual), its analytic machinery (Figures 2.1, 2.2, 3.2, and 3.3), and its procedural safeguards (audit trail, reflexive consistency, triangulation, and saturation). The result is a theoretically disciplined and empirically anchored account of how English medical textbooks in Vietnam visually and intersemiotically construct disciplinary knowledge.

To visualise how these principles operate as a coherent methodological ecology, the Reflexive Trustworthiness Framework is presented in Figure 3.3.



Figure 3.4 Reflexive Trustworthiness Framework

As illustrated in Figure 3.4, reflexivity occupies the central position of the framework, representing the researcher's continuous awareness of positionality, interpretive responsibility, and epistemological commitments (Mickan, 2017). This central placement signals that meaning-making in multimodal analysis is inseparable from the researcher's theoretically informed interpretive stance.

Surrounding this core are the interdependent dimensions of reliability and validity. Reliability is enacted through reflexive consistency and the maintenance of an audit trail, ensuring that analytical decisions remain traceable and coherent over time (Miles et al., 2014). Validity is conceptualised as interpretive sufficiency achieved through analytic convergence across multimodal strata and through theoretical saturation, rather than through statistical verification.

The outermost layer of the framework represents ethics, which functions as both a moral and epistemic boundary encompassing all stages of the research process. Ethics here governs not only data handling but also interpretive responsibility, cultural sensitivity, and intellectual integrity.

Importantly, the circular configuration of the framework underscores that trustworthiness in multimodal research is dialogic, cyclical, and reflexively sustained. Rather than seeking objectivity in a positivist sense, the framework foregrounds interpretive transparency – the reader's ability to follow the analytical reasoning, evaluate its coherence with theory, and assess its methodological integrity (Lincoln & Guba, 1985). Figure 3.4 thus functions not as an additional analytical model, but as a meta-methodological representation of how rigor is sustained throughout the application of the IMAF.

3.8 Concluding Remarks

This chapter has established a coherent methodological architecture for explaining multimodal meaning-making in English medical textbooks used in Vietnam. Building from the interpretive-explanatory social-semiotic paradigm, the study operationalises a top-down design in which theory guides corpus construction, data selection, and the staged reading of page ensembles as pedagogically designed semiotic constellations. The chapter has also justified the comparative multiple-case focus on *Medical Terminology: A Living Language* and *Career Paths: Medical*, showing how their contrasting multimodal ecologies provide analytic leverage for examining both visual organisation (RQ1) and visual-linguistic interaction (RQ2).

Importantly, the chapter has translated theoretical commitments into analysable procedures. The five-step data collection workflow ensures that the dataset is not only fit for purpose but also fully traceable from page evidence to interpretive claims. The IMAF, presented as a three-stratum analytical apparatus, further disciplines interpretation by specifying what counts as evidence at each level – from the visual semiotic organisation of images, to the mechanisms of intersemiotic coupling, and finally to pedagogic–disciplinary meaning interpretation. In turn, the coding phases and the rubricised codebook provide an explicit operational bridge between theory and analysis, enabling systematic cross-case comparison across the forty selected page ensembles without reducing multimodal meaning to mechanical feature-counting.

Finally, the chapter has consolidated how trustworthiness is sustained as an emergent property of design coherence, supported by reflexive consistency, triangulated warranting across strata, saturation logic, and ethical integrity. Taken together, these methodological provisions ensure that the

analysis reported in Chapter IV can proceed with epistemic accountability: interpretations are not impressionistic, but theoretically constrained, systematically evidenced, and transparent in their analytic reasoning. On this basis, the next chapter turns to the empirical findings, applying the IMAF and the rubricised codebook to reveal patterned ways in which images and language co-construct medical knowledge and pedagogic meaning in the selected textbooks.

CHAPTER IV: VISUAL MEANING IN MEDICAL TEXTBOOK IMAGERY

This chapter examines visual meaning in EMP textbook imagery as an epistemic resource that shapes what counts as medical knowledge for novice learners. Images are treated not as supplementary illustrations but as semiotic designs that stabilise concepts, model practice, and direct attention through representation, viewer positioning, and composition. Grounded in IMAF-Stratum 1 and operationalised through the Grammar of Visual Design, the chapter analyses 40 page ensembles from corpus-level patterns to key meaning mechanisms, establishing the visual foundation for Chapter V's intersemiotic analysis of image-text coupling.

4.1 Visual Meaning as Disciplinary Vision: Analytical Orientation and Scope

4.1.1 Chapter focus and alignment with RQ1

Chapter IV addresses RQ1 by examining how medical textbook images organise meaning within the visual mode (IMAF – Stratum 1). In this chapter, visual images are treated not as ancillary illustrations but as epistemic resources that select, stabilise, and make “*seeable*” disciplinary knowledge for novice learners. This orientation is consistent with multimodal social semiotics, where disciplinary understanding is approached as the outcome of semiotic design rather than the transparent transmission of biomedical facts. Critically, in medical education – where diagrams, cross-sections, charts, and clinical photographs are routinely used to teach professional “*ways of seeing*” – the visual mode does not merely decorate verbal explanation; it participates in constructing what counts as relevant anatomical structure, procedural logic, and clinically salient relations.

4.1.2 Unit of analysis: the page ensemble

The unit of analysis is the page ensemble, defined as a designed constellation of image(s), captions/labels, adjacent verbal text, and layout cues that together realise an instructional message. Importantly, this unit preserves multimodal cohesion: a diagram's meaning potential is typically completed by labels, typographic hierarchy, and spatial alignment, while textual definitions frequently presuppose what is made salient visually. The page ensemble therefore provides the most defensible analytic granularity for answering RQ1, because it enables systematic description of representational, interactive, and compositional visual meanings without fragmenting images from their pedagogic co-text.

4.1.3 Corpus and evidential base: 40 page ensembles across two textbook ecologies

The corpus comprises 40 purposively selected page ensembles logged in Appendix A, including 20 ensembles from *Medical Terminology: A Living Language* (Fremgen & Frucht, 2019) (MT) and 20 ensembles from *Career Paths: Medical* (Evans, Dooley, & Tran, 2012; six-book series) (CPM). Appendix A specifies each ensemble's textbook source, page, brief description of visual-verbal content, and primary visual-pedagogic function – thereby providing corpus traceability and ensuring that all analytic claims in Chapter IV can be mapped to page-level evidence.

However, the two textbooks do not merely provide “*more data*”; they instantiate two distinct visual semiotic ecologies. MT systematically foregrounds conceptual organisation (e.g., classificatory and analytical representations, labelled schemata, and hierarchical layouts), which aligns with a terminology-building pedagogy and encourages learners to treat medical knowledge as taxonomically structured and terminologically

decomposable. In contrast, CPM relies heavily on contextualised images – scenarios, interpersonal encounters, and task-sequenced panels – whose visual meaning tends to be more situational and practice-oriented, supporting communicative rehearsal and professional role familiarisation. These contrasting ecologies are not treated as background; they are analytically productive, because they enable Chapter IV to explain how visual meanings are patterned differently across the corpus before Chapter V accounts for how these patterns are intersemiotically mobilised.

4.1.4 Analytical lens: Grammar of Visual Design operationalised within IMAF - Stratum 1

Analytically, Chapter IV activates IMAF – Stratum 1 (Figure 3.2) by applying the GVD to the textbook images as structured semiotic systems. The analysis proceeds across three interdependent meaning systems:

Representational meaning (A1): how images depict entities, relations, and processes (e.g., conceptual classification vs. processual depiction).

Interactive meaning (A2): how images position viewers through gaze, distance, angle, and modality cues – thereby shaping epistemic relations between learner and knowledge.

Compositional meaning (A3): how layout constructs information value, salience hierarchies, and framing relations that guide reading paths and prioritise what counts as central knowledge.

Importantly, these three systems are not treated as separate “*mini-analyses*” Instead, they form a single explanatory chain: what is represented (A1) is simultaneously made learnable through viewer positioning (A2) and made retrievable through compositional organisation (A3). This integrated treatment is essential in the medical domain, where disciplinary knowledge is frequently visualised through abstraction (schematisation, cross-sectioning,

spatial segmentation) and where comprehension depends on recognising how diagrams and layouts orchestrate attention, hierarchy, and conceptual boundaries.

4.1.5 Coding source and procedural discipline: Visual Codebook + Appendix B-D + interpretive salience calibration

To keep interpretation theory-constrained (rather than impressionistic), Chapter IV draws directly on the Visual Codebook (Appendix B1) and the corpus-wide coding outputs documented in Appendix B4 (auto-coded Medical Terminology ensembles), Appendix B5 (auto-coded Career Paths ensembles), Appendix C (Page-Ensemble Analytical Matrix), and Appendix D (traceability across ensembles). Appendix C explicitly frames the analysis as a systematic coding of all page ensembles across Modules A-C, where Module A operationalises visual meaning through representational (A1), interactive (A2), and compositional (A3) dimensions.

Moreover, Appendix C records not only categorical coding but also calibrated strength using a 0-3 scale and a visual-meaning composite, enabling Chapter IV to argue for corpus-level patterns (e.g., dominance of conceptual representation and observer positioning in MT) with methodological transparency rather than anecdotal selection.

This evidential architecture is complemented by the auto-coded ensemble summaries in Appendix B4/B5, which provide consistent code-level descriptions of how each ensemble realises representational, interactive, and compositional meanings (e.g., MT-P59 coded as V-REP-CON, V-INT-OFF, V-COMP-HIER).

4.1.6 Chapter trajectory and discourse logic

The chapter is structured to move from corpus-wide distributional tendencies to interpretable visual meaning mechanisms, and then to a

synthesising account that prepares the analytic transition to Chapter V. Section 4.2 establishes the overall distribution of visual semiotic resources across the 40 ensembles using Appendix C as the corpus matrix. The subsequent sections (4.3-4.5) provide the core interpretive work of IMAF-Stratum 1: representational meaning (how medical knowledge is visually construed), interactive meaning (how learners are positioned epistemically), and compositional meaning (how knowledge is organised and prioritised). Finally, Section 4.6 consolidates the chapter's findings into visual meaning profiles for the two textbooks, clarifying which visual patterns are likely to produce dense image-text coupling in Chapter V and which patterns suggest a more complementary division of semiotic labour.

Critically, this chapter does not claim that visual meanings automatically translate into learning outcomes. Instead, it advances a defensible epistemic claim consistent with the study's interpretive-explanatory paradigm: visual design makes certain kinds of medical understanding more available than others by stabilising representations, directing attention, and positioning learners in relation to knowledge. This is precisely why Stratum 1 is necessary: before intersemiotic relations can be explained (Chapter V), the analysis must first establish – systematically and corpus-wide – what the images are doing as images.

4.2 Distribution of Visual Semiotic Resources across the Corpus

This section examines how visual semiotic resources are distributed across the full corpus of 40 page ensembles, drawing on the Page-Ensemble Analytical Matrix (Appendix C). Rather than focusing on individual pages, the analysis adopts a corpus-level perspective in order to identify recurrent design tendencies and systematic contrasts between the two textbooks. Such an overview is essential for establishing that the visual patterns discussed in

subsequent sections are not anecdotal but structurally embedded within the instructional design of the materials.

Across the corpus, a clear divergence emerges between MT and CPM. However, this divergence should not be understood as a simple difference in image type. More importantly, the two textbooks repeatedly mobilise visual resources to support distinct epistemic orientations toward medical knowledge.

In MT, visual meaning is predominantly organised around conceptual representation. Anatomical diagrams, classificatory schemata, and labelled cross-sections consistently construe medical knowledge as a stable system of relations. These images tend to abstract away from social action and clinical context, foregrounding instead the internal structure of the body and the taxonomic relations among terms. As a result, visual design supports a view of medical knowledge as something to be mastered through categorisation, naming, and structural understanding.

By contrast, CPM distributes visual meaning more evenly across contextualised and process-oriented representations. Images frequently depict people, tools, and actions situated in recognisable clinical settings, thereby embedding medical concepts within scenarios of use. Rather than presenting knowledge as an abstract system, these visuals frame it as something that is enacted, negotiated, and performed within professional practice.

Importantly, the difference between the two textbooks is not absolute. MT occasionally employs process-based visuals to explain physiological mechanisms, while CPM sometimes uses schematic layouts to support clarity. Nevertheless, the corpus-level pattern remains consistent: conceptual abstraction is the default visual logic in MT, whereas contextualisation and procedural orientation are the default in CPM.

This contrast is reinforced at the level of viewer positioning. In MT, images typically position learners as observers of expert knowledge. Visual distance, neutral angles, and the absence of direct address encourage an epistemic stance in which learners are invited to inspect and classify medical objects rather than participate in interaction. In CPM, however, viewer relations are more variable. Some images still encourage observation, particularly in procedural explanations, but others position learners closer to the depicted activity, aligning them with professional roles and communicative situations.

Compositional organisation further consolidates these tendencies. Pages in MT often employ hierarchical layouts and strong visual segmentation, which stabilise conceptual priorities and guide attention toward what counts as core knowledge. In contrast, CPM frequently adopts balanced or integrated layouts that distribute attention across images, tasks, and verbal prompts, supporting sequential pedagogic action rather than conceptual hierarchy.

Taken together, the distributional patterns documented in Appendix C demonstrate that visual meaning in the corpus is not randomly assembled. Instead, it is systematically aligned with the pedagogic intentions of each textbook. MT uses visual design to stabilise medical knowledge as a conceptual system, whereas CPM uses visual design to situate knowledge within professional action and communication. These corpus-level tendencies provide the necessary foundation for the more fine-grained analyses of representational, interactive, and compositional meaning developed in Sections 4.3-4.5, and they anticipate the intersemiotic interpretations advanced in Chapter V.

4.3 Representational Meaning in Medical Textbook Imagery

Having established the corpus-wide distribution of visual resources in Section 4.2, this section turns to a more fine-grained analysis of

representational meaning in medical textbook imagery. Within the IMAF, representational meaning concerns how images construe the *content of medical knowledge*: what kinds of entities are depicted, what relations are foregrounded, and what kinds of processes are made visible - or rendered invisible.

From a multimodal social semiotic perspective, representational choices are never neutral. They embody disciplinary assumptions about what counts as legitimate knowledge and how that knowledge should be accessed by learners. In medical education in particular, visual representations do not simply illustrate biomedical facts; they actively participate in constructing professional ways of seeing, determining whether the body is apprehended as an abstract system of parts, a site of dynamic processes, or a context for social and clinical interaction (Goodwin, 1994; Lemke, 1998).

Following the GVD, this study distinguishes between conceptual representations, which depict stable relations such as classification, part-whole structure, and symbolic identity, and narrative representations, which depict actions, events, and processes unfolding over time. This distinction is analytically crucial for answering RQ1, because it allows the analysis to explain how different textbooks privilege different epistemic orientations toward medical knowledge.

Accordingly, Section 4.3 is organised into two complementary subsections. Section 4.3.1 examines conceptual representation as the dominant representational mode in *Medical Terminology: A Living Language*, drawing on 20 page ensembles. Section 4.3.2 then examines narrative representation as a salient representational tendency in *Career Paths: Medical*. Together, these analyses demonstrate that representational meaning in EMP textbooks is systematically patterned rather than incidental.

4.3.1 *Conceptual Representation of Medical Knowledge*

Across 0 page ensembles, *Medical Terminology: A Living Language* consistently construes medical knowledge through conceptual representation. Rather than depicting people performing medical actions or interacting in clinical settings, the images predominantly depict anatomical structures, systems, and relations organised for inspection, comparison, and naming. This representational orientation aligns closely with the textbook's pedagogic focus on terminology acquisition, but its epistemic significance extends beyond curricular convenience.

4.3.1.1 *Conceptual abstraction and the stabilisation of biomedical knowledge*

A defining feature of conceptual representation in these ensembles is the abstraction of the body from lived context. Organs, tissues, and systems are presented as stable entities, isolated from individual variation and clinical contingency. By foregrounding structure over action, the images stabilise what counts as *medical knowledge proper*: that which can be categorised, labelled, and re-identified across contexts.

Importantly, such abstraction functions as an epistemic resource. As Kress and van Leeuwen (2021) argue, conceptual images enable learners to apprehend complex domains through relations rather than events. In the MT ensembles, this is realised through repeated emphasis on classificatory structures (e.g. systems and subsystems) and analytical structures (e.g. part-whole relations within anatomical regions). The pedagogic affordance is that learners are invited to treat medical knowledge as a coherent conceptual system, rather than as a set of isolated facts or situational responses.

However, abstraction also entails selection. By removing social actors, tools, and interactional context, conceptual representation foregrounds

internal organisation while backgrounding use, practice, and uncertainty. This selective visibility reflects what Lemke (1998) describes as the characteristic “*thematic condensation*” of scientific discourse, where processes and experiences are reconfigured into stable objects of knowledge.

4.3.1.2 *Classificatory and analytical logics as visual reasoning tools*

Within the conceptual mode, two representational logics recur across the MT page ensembles: classification and analysis. Classificatory representations organise medical entities into taxonomic systems, implicitly teaching learners how medical knowledge is structured hierarchically. Such visuals support definitional reasoning: learners can understand a term not merely by memorising its meaning but by locating it within a broader conceptual field. This aligns with Halliday and Matthiessen’s (2014) observation that scientific knowledge is typically construed through relational and nominal structures, a tendency that is here realised visually rather than solely linguistically.

Analytical representations, by contrast, foreground part–whole relations. Cross-sections, layered diagrams, and segmented schemata invite learners to reason spatially about containment, adjacency, and structural dependency. These images externalise relational knowledge, reducing cognitive load by making abstract relations visually explicit. International research on science and medical textbooks has repeatedly shown that such representations function as cognitive scaffolds for novice learners, enabling them to coordinate terminology with spatial understanding (Bezemer & Kress, 2016; Unsworth, 2008, Peeck, 1993).

The present analysis extends this body of research by demonstrating that these representational logics are not sporadic but systematically

distributed across the MT corpus, as evidenced by the consistent visual profiles recorded in Appendix B4 and synthesised in Appendix C.

4.3.1.3 Canonical biomedical perspective and epistemic authority

Beyond facilitating comprehension, conceptual representation also constructs a canonical biomedical perspective. The images repeatedly adopt idealised viewpoints, standardised orientations, and schematic conventions that naturalise a particular way of seeing the body as authoritative. This does not imply that the representations are inaccurate; rather, they encode disciplinary norms about what is worth seeing and how it should be seen. Such canonical visualisation supports definitional and classificatory discourse by creating an environment in which linguistic statements appear self-evident. When visual space is already organised into discrete, labelled entities, definitional clauses (“X is...”, “X refers to...”) gain epistemic force. From a multimodal perspective, conceptual images therefore function as preconditions for linguistic authority, rather than merely as supplements to text.

4.3.1.4 Relevance to EMP pedagogy in the Vietnamese context

In the Vietnamese EMP context, where instruction has traditionally emphasised terminology mastery, translation accuracy, and reading comprehension, the dominance of conceptual representation in Medical Terminology is pedagogically congruent. Local ESP studies have noted that medical English instruction often prioritises definitional knowledge over situated communication, particularly at early stages of training (Lam & Pham, 2024, 2015; Tran & Lam, 2025). Conceptual visuals therefore align with existing instructional expectations and assessment practices. However, multimodal scholarship cautions that when conceptual representation becomes the sole or dominant visual epistemology, learners’ exposure to the

interactional and procedural dimensions of medical practice may be limited (Unsworth, 2008; O'Halloran, 2008). This observation does not undermine the pedagogic value of conceptual images; rather, it highlights the importance of examining how different representational modes are balanced across instructional materials – a concern that becomes central in Section 4.3.2 and, more explicitly, in Chapter V.

4.3.1.5 Interim synthesis

In sum, the 20 page ensembles analysed in this subsection demonstrate that *Medical Terminology: A Living Language* systematically employs conceptual representation to construe medical knowledge as stable, classifiable, and canonically viewable. Through classificatory and analytical structures, images operate as epistemic resources that stabilise biomedical concepts and support terminology acquisition. At the same time, these representational choices foreground certain dimensions of medical knowledge while backgrounding others. Understanding this representational orientation is therefore essential for interpreting how visual meaning interacts with linguistic explanation and pedagogic intent in subsequent chapters.

4.3.2 Narrative Representation of Medical Processes and Practices

In contrast to the conceptual abstraction that characterises *Medical Terminology*, the page ensembles drawn from *Career Paths: Medical* predominantly construe medical knowledge through narrative representation. Across the 20 ensembles analysed here, images depict actions, reactions, and sequences of practice, frequently involving human participants engaged in recognisable professional activities. Rather than presenting medicine as a stable system of structures to be named, these visuals frame medicine as something that is done – through procedures, interactions, and communicative routines situated in institutional settings.

From a multimodal perspective, narrative representation is epistemically significant because it organises meaning around temporal unfolding and social agency. Images that depict actions and reactions invite learners to understand medical knowledge not only in terms of what it *is*, but also in terms of what it *enables practitioners to do*. In the context of EMP instruction, this shift from conceptual stability to processual enactment has important implications for how learners are positioned in relation to professional practice.

4.3.2.1 *Situated action and the visualisation of medical practice*

A defining characteristic of the CPM ensembles is their consistent grounding of medical processes in situated contexts. Images frequently include human actors, tools, institutional spaces, and interpersonal configurations, thereby embedding medical knowledge within scenes of use. Importantly, these scenes do not merely provide “*realistic*” background; they function as representational anchors that link language learning to professional action.

International research on multimodal professional discourse has shown that such representations play a crucial role in socialising learners into disciplinary practices (Goodwin, 1994; Bezemer & Kress, 2016). By depicting what counts as appropriate action, posture, and interaction in clinical settings, narrative visuals contribute to what has been described as *professional vision*: the ability to recognise relevant cues and respond appropriately within a field. The CPM ensembles analysed here align with this tradition by repeatedly foregrounding actions such as greeting patients, performing routine procedures, documenting information, and coordinating with colleagues.

Moreover, the narrative orientation of these images often works in tandem with task-based verbal prompts, encouraging learners to map

linguistic forms onto visually instantiated actions. In this way, narrative representation functions not simply as illustration but as a semiotic bridge between language and practice – a point that becomes analytically central in Chapter V’s intersemiotic analysis.

4.3.2.2 Actional and reactional processes as epistemic resources

Within narrative representation, two interrelated processes recur across the CPM ensembles: actional processes, where participants are shown performing medical tasks, and reactional processes, where participants respond to others or to clinical situations. Together, these processes visualise medicine as an activity that unfolds through sequences of doing and responding.

Actional representations frequently depict routine and procedural tasks, such as preparing equipment, administering care, or completing documentation. These images externalise procedural logic, making the steps of practice visible and thus learnable. Rather than requiring learners to infer procedures solely from text, narrative visuals offer a scaffolded entry point into procedural reasoning – a pedagogic affordance that has been widely recognised in ESP and vocational education research.

Reactional representations, by contrast, foreground interpersonal orientation: eye lines, body positioning, and spatial proximity depict how practitioners attend to patients, colleagues, and supervisors. Such images are particularly significant in EMP contexts because they visualise communicative norms – who speaks to whom, under what conditions, and with what orientation. In doing so, they prepare learners for the pragmatic dimensions of medical English that extend beyond terminology and syntax.

Critically, the alternation between actional and reactional imagery across the CPM corpus suggests that the textbook does not reduce medical practice to mechanical procedure. Instead, it presents practice as inherently

social and responsive, a view that resonates with sociolinguistic accounts of clinical interaction and patient-centred care.

4.3.2.3 Narrative representation and the enactment of professional roles

Another salient feature of narrative representation in CPM is its role in constructing professional identity. By repeatedly depicting learners' potential future roles – nurse, medical assistant, caregiver – images invite viewers to imagine themselves as legitimate participants in medical practice. This representational strategy aligns with international research on vocational textbooks, which emphasises the importance of visual role modelling in professional socialisation (Bezemer & Kress, 2016; Machin & Mayr, 2023).

Importantly, these images do not simply show “*anyone*” performing actions; they encode normative expectations about professional conduct, including attentiveness, empathy, and procedural competence. In this sense, narrative representation contributes to an epistemic shift: medical knowledge is construed not only as something to be known, but as something to be performed appropriately within institutional norms.

From a Vietnamese EMP perspective, this representational orientation is particularly noteworthy. Local studies have frequently reported challenges in bridging the gap between English-for-medical purposes instruction and actual clinical communication, with learners often feeling confident in terminology but less prepared for interactional demands. Narrative visuals in CPM address this gap by foregrounding communicative practice and professional behaviour, thereby expanding the epistemic scope of EMP learning beyond lexical mastery.

4.3.2.4 Comparative epistemic implications

However, it is important not to overstate the contrast between CPM and MT. While narrative representation is dominant in CPM, conceptual elements still appear, particularly when clarity and standardisation are pedagogically

required. Conversely, as shown in Section 4.3.1, MT occasionally incorporates process-oriented visuals. The crucial point is not mutual exclusivity, but representational priority.

The CPM ensembles prioritise narrative representation as the primary means of meaning-making, thereby construing medical knowledge as situated, procedural, and interactional. This contrasts with the conceptual priority observed in MT, where knowledge is stabilised through abstraction and classification. Together, these representational orientations illustrate how different EMP textbooks mobilise visual meaning to support different pedagogic trajectories: one oriented toward conceptual mastery, the other toward professional enactment.

4.3.2.5 Interim synthesis for Section 4.3

Taken together, Sections 4.3.1 and 4.3.2 demonstrate that representational meaning in medical textbook imagery is systematically patterned at the textbook level. *Medical Terminology* predominantly employs conceptual representation to stabilise biomedical knowledge, whereas *Career Paths: Medical* predominantly employs narrative representation to situate knowledge within practice. These representational choices are not merely stylistic; they encode distinct epistemic assumptions about what it means to “know” medicine in an EMP context.

Understanding this representational contrast is essential for answering RQ1 and provides a necessary foundation for the subsequent analysis of interactive and compositional meanings (Sections 4.4 and 4.5). Moreover, it sets the stage for Chapter V, where the analysis will examine how these visual representational patterns interact with language to produce intersemiotic meanings and, ultimately, multimodal pedagogic orientations.

4.4 Interactive Meaning: Viewer Positioning and Epistemic Relations

Across the 40 page ensembles (Appendix A), interactive meaning is realised through patterned visual choices that regulate how learners are

positioned epistemically in relation to medical knowledge. Importantly, viewer positioning is not an “extra” layer added after representational content; rather, it is built into the semiotic design of images via gaze/contact, social distance, angle, and modality cues – the four interactive systems operationalised in the Grammar of Visual Design and adopted in the IMAF (Figure 3.2; Appendix B1). However, because the two textbooks instantiate different pedagogic orientations (Conceptual Framework – Figure 2.2), their interactive designs systematically diverge: *Medical Terminology* tends to construct learners as detached epistemic observers of canonical biomedical knowledge, whereas *Career Paths: Medical* more often positions learners as professionally oriented participants facing situated practice scenarios. This section substantiates that claim by first examining gaze and visual address (4.4.1), and then consolidating social distance and angle as epistemic relations (4.4.2), drawing on corpus-level patterns in Appendix C and traceability evidence in Appendix D.

4.4.1 Gaze and Visual Address

At corpus level, gaze/contact operates as a primary mechanism for regulating interpersonal demand. In *Medical Terminology*, images overwhelmingly function as offer images – they do not “address” the learner but rather present biomedical entities for inspection. In *Career Paths: Medical*, by contrast, photographs and contextual scenes typically feature indirect gaze (participants look at tasks, tools, or each other), which subtly reorients the learner from “viewer of objects” to “observer of professional action”. Critically, this difference aligns with the textbooks’ distinct pedagogic projects: the former privileges conceptual stability and terminological control, while the latter foregrounds situated communication and workplace routines.

To concretise how “offer” is realised as epistemic distancing, MT-P148 provides a prototypical case.

Anatomy and Physiology of the Cardiovascular System

arteries
blood vessels
capillaries
carbon dioxide
circulatory system
deoxygenated (dee-OK-sih-jen-ay-ted)
heart

oxygen
oxygenated (OK-sih-jen-ay-ted)
pulmonary circulation (PULL-mon-air-ee / ser-kyoo-LAY-shun)
systemic circulation (sis-TEM-ik / ser-kyoo-LAY-shun)
veins

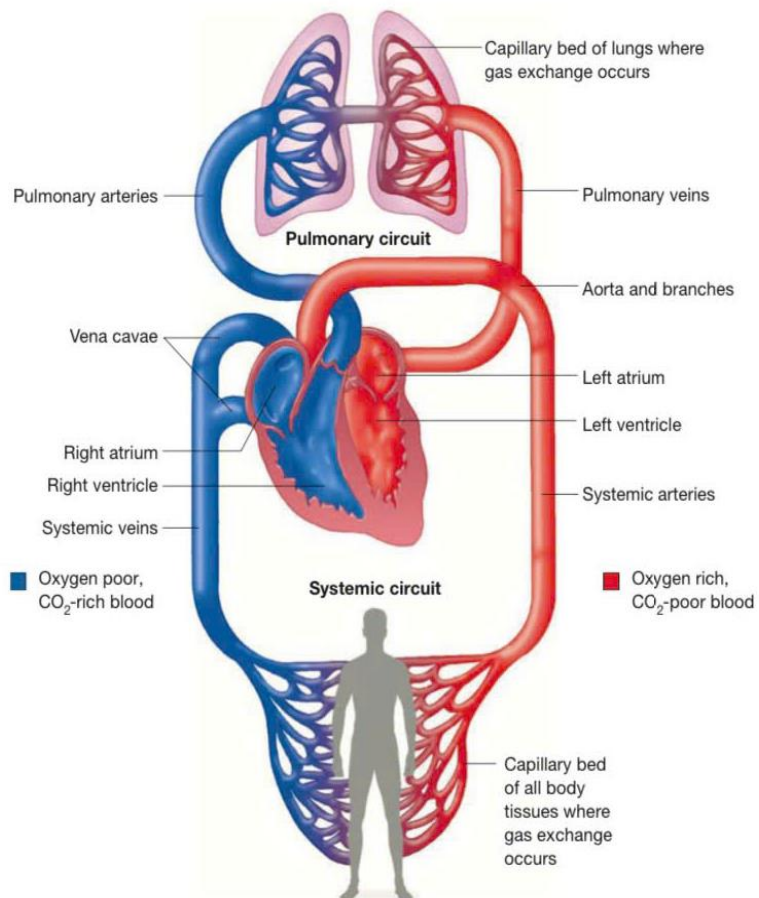
What's In A Name?

Look for these word parts:

ox/o = oxygen
pulmon/o = lung
system/o = system
-ary = pertaining to
-ic = pertaining to
de- = without
di- = two

The cardiovascular (CV) system, also called the **circulatory system**, maintains the distribution of blood throughout the body and is composed of the **heart** and the **blood vessels**—**arteries**, **capillaries**, and **veins**.

The circulatory system is composed of two parts: the **pulmonary circulation** and the **systemic circulation**. The pulmonary circulation, between the heart and lungs, transports **deoxygenated** blood to the lungs to get oxygen, and then back to the heart. The systemic circulation carries **oxygenated** blood away from the heart to the tissues and cells, and then back to the heart (see Figure 5-1 ■). In this way, all the body's cells receive blood and oxygen.



■ **Figure 5-1** A schematic of the circulatory system illustrating the pulmonary circulation picking up oxygen from the lungs and the systemic circulation delivering oxygen to the body.

Within GVD, the absence of visual contact (gaze) is a key cue of “*offer*”, positioning the viewer as an analyst rather than an interlocutor. MT-P148 is dominated by a schematic diagram of pulmonary vs. systemic circulation with colour-coded pathways (red/blue), labelled components, and a neutral presentation of anatomical structures. No represented participant (human face/eyes) is available for eye contact; the human silhouette at the bottom functions as a generic carrier rather than a social actor. The diagram’s communicative burden lies in classificatory clarity and systemic mapping rather than interpersonal engagement. Because the image offers the circulatory system as an abstracted object of study, learners are positioned as epistemic observers – expected to read, trace, and stabilise biomedical relations (flow, separation of circuits, oxygenation contrast) rather than to empathise, respond, or negotiate meaning. Moreover, this “*offered*” anatomy coheres with the textbook’s broader pedagogic logic: disciplined scientific seeing, where knowledge is validated through standardised representation and terminological precision (cf. the thesis’ paradigm in 3.1). This use of “*offer*” to produce objectified biomedical knowledge reflects the GVD claim that gaze regulates interpersonal demand, and resonates with multimodal textbook research noting that scientific diagrams commonly suppress address to sustain an authoritative, decontextualised epistemic stance (e.g., Kress & van Leeuwen; Bezemer & Kress; Lemke; Jewitt).

A complementary “*offer*” pattern is visible in MT-P205, where the interpersonal distance is reinforced by typological display of organs.

Tonsils

adenoids (AD-eh-noydz)

lingual tonsils (LING-gwal)

palatine tonsils (PAL-ah-tyne)

pharyngeal tonsils (fair-IN-jee-al)

pharynx (FAIR-inks)

The tonsils are collections of lymphatic tissue located on each side of the throat or **pharynx** (see Figure 6-11 ■). There are three sets of tonsils: **palatine tonsils**, **pharyngeal tonsils** (commonly referred to as the **adenoids**), and **lingual tonsils**. All tonsils contain a large number of leukocytes and act as filters to protect the body from the invasion of pathogens through the digestive or respiratory systems. Tonsils are not vital organs and can safely be removed if they become a continuous site of infection.



■ **Figure 6-11** The shape of a tonsil.

What's In A Name?

Look for these word parts:

lingu/o = tongue

palat/o = palate

pharyng/o = pharynx

-al = pertaining to

-eal = pertaining to

-ine = pertaining to

Spleen

blood sinuses

macrophages (MAK-roh-fay-jez)

The spleen, located in the upper left quadrant of the abdomen, consists of lymphatic tissue that is highly infiltrated with blood vessels (see Figure 6-12 ■). These vessels spread out into slow-moving **blood sinuses**. The spleen filters out and destroys old red blood cells, recycles the iron, and also stores some of the blood supply for the body. Phagocytic **macrophages** line the blood sinuses in the spleen to engulf and remove pathogens. Because the blood is moving through the organ slowly, the macrophages have time to carefully identify pathogens and worn-out red blood cells. The spleen is also not a vital organ and can be removed due to injury or disease. However, without the spleen, a person's susceptibility to a bloodstream infection may be increased.

What's In A Name?

Look for these word parts:

macro- = large

-phage = to eat



■ **Figure 6-12** The shape of the spleen.

What's In A Name?

Look for these word parts:

lymph/o = lymph

-cyte = cell

Thymus Gland

T cells

thymosin (THIGH-moh-sin)

T lymphocytes

The thymus gland, located in the upper portion of the mediastinum, is essential for the proper development of the immune system (see Figure 6-13 ■). It assists the body with the immune function and the development of antibodies. This organ's hormone, **thymosin**, changes lymphocytes to **T lymphocytes** (simply called **T cells**), which play an important role in the immune response. The thymus is active in the unborn child and throughout childhood until adolescence, when it begins to shrink in size.



■ **Figure 6-13** The shape of the thymus gland.

What's In A Name?

Look for this word part:

-ous = pertaining to

Immunity

acquired immunity

active acquired immunity

bacteria (bak-TEE-ree-ah)

cancerous tumors

fungi (FUN-je-ye)

immune response

immunity (im-YOO-nih-tee)

immunizations (im-yoo-nih-ZAY-shuns)

natural immunity

passive acquired immunity

protozoans (proh-toh-ZOH-anz)

toxins

vaccinations (vak-sih-NAY-shuns)

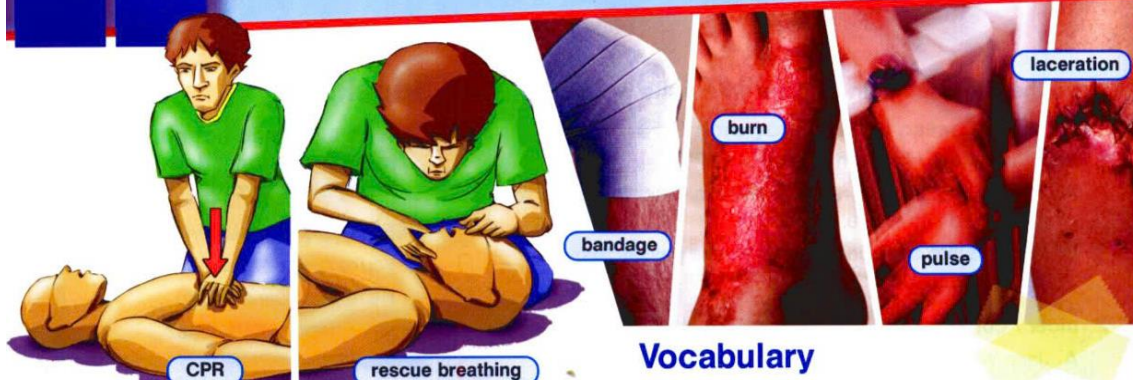
viruses

Immunity is the body's ability to defend itself against pathogens, such as **bacteria**, **viruses**, **fungi**, **protozoans**, **toxins**, and **cancerous tumors**. Immunity comes in two forms: **natural immunity** and **acquired immunity**. Natural immunity, also called *innate immunity*, is not specific to a particular disease and does not require prior exposure to the pathogenic agent. A good example of natural immunity is the

In GVD terms, offer images can be strengthened through isolation and framing, which minimise social presence and maximise classificatory readability. MT-P205 presents multiple organ images (e.g., tonsil, spleen, thymus) as discrete “*specimens*” placed alongside definitional prose and boxed word-part explanations. The organs appear detached from embodied context; there is no face, no interpersonal gaze, and no interaction between represented participants. The layout separates visual objects into stable units, encouraging comparison and recognition. Importantly, this design positions learners not merely as observers, but as classifiers - readers are invited to map labels and word parts onto visually isolated anatomical tokens. That epistemic role supports a particular kind of EMP literacy: decoding and naming, where the learner’s legitimacy comes from correctly aligning form (image) with technical lexis (text). In Vietnamese EMP settings where teaching has often privileged linguistic decoding, such visuals can appear “*supportive*”; however, from a multimodal literacy standpoint, they also risk narrowing learning to identification over interpretive reasoning if teaching does not explicitly mobilise the diagram-text interplay. The strong “*offer + isolation*” pattern exemplifies GVD’s interpersonal system (gaze/contact) and aligns with multimodal scholarship showing that medical/science textbooks frequently use decontextualised displays to stabilise canonical perspectives and reduce interpretive ambiguity (e.g., Lemke; O’Halloran; Bezemer & Kress), a point also relevant to local concerns about linguistically oriented ESP/EMP pedagogy.

While *Medical Terminology* minimises address, CPM-B1P22 illustrates how Career Paths builds professional orientation through indirect gaze and workplace action.

10 First aid



Basic FIRST AID Techniques

In a medical emergency, always call for assistance first. The following are basic steps to take to help an injury victim before medical help arrives.

Assess the victim's condition. Give him or her plenty of space. Prevent **shock** by elevating the legs and covering the victim.

Apply **pressure** to **lacerations** or other open **wounds** to stop bleeding. If bleeding has stopped, wounds should be cleaned and wrapped in a **bandage**. Cool minor **burns** with running water.

If the victim's **pulse** or breathing has stopped, **CPR** or **rescue breathing** can be performed. Be sure to clear the victim's airway before rescue breathing is carried out.

Get ready!

1 Before you read the passage, talk about these questions.

- 1 What are some injuries that first aid can help?
- 2 What are some common first aid techniques?

Reading

2 Read the first aid guide. Then, mark the following statements as true (T) or false (F).

- 1 ___ Bleeding wounds should be wrapped with a bandage.
- 2 ___ Use running water to cool burns.
- 3 ___ Rescue breathing can clear a victim's airway.

Vocabulary

3 Match the words (1-6) with the definitions (A-F).

- | | |
|-------------|------------------------|
| 1 ___ shock | 4 ___ rescue breathing |
| 2 ___ burn | 5 ___ laceration |
| 3 ___ pulse | 6 ___ CPR |

- A a normal, regular heartbeat that can be felt when an artery is pressed
- B a first aid procedure involving chest compression and artificial breathing
- C an open injury caused by a hard impact to soft body tissue
- D a medical emergency in which the circulatory system cannot provide oxygen to the body
- E a first aid procedure involving providing air for someone who has stopped breathing, but still has a heartbeat
- F a physical injury that can be caused by heat

Vocabulary

4 Fill in the blanks with the correct words and phrases from the word bank.

word BANK

wound first aid bandage
pressure assess

- 1 The doctor will need to _____ the victim's condition.
- 2 I am trained in basic _____ techniques.
- 3 Apply _____ to the injury to stop the bleeding.
- 4 Wash the injury and wrap a(n) _____ around it.
- 5 The nurse used stitches to close the _____.

In GVD, indirect gaze typically constitutes “*offer*”, yet in workplace photographs/illustrations it can still position viewers as professionally implicated observers because social actors and practices are present. CPM-B1P22 combines a thematic visual strip (first aid techniques such as CPR/rescue breathing; injury close-ups) with reading/vocabulary tasks organised for classroom performance. Human actors are present (performing first aid), but they do not look at the viewer; their attention is directed to the patient/task. The design thus withholds direct demand while foregrounding procedural action. However, unlike MT’s abstract offer, this page positions learners as observers of practice and potential apprentices to procedures. The indirect gaze draws learners into a “*professional field of action*”: what matters is not merely naming, but recognising actions, risks, and routine responses that can be verbalised and rehearsed through tasks. Moreover, this configuration supports multimodal pedagogy implicitly: images cue what counts as relevant action, while tasks convert visual recognition into communicative performance (reading comprehension → vocabulary control → application). In other words, professional orientation is achieved not by direct address, but by structuring learners’ attention around situated practice. This pattern operationalises the GVD interpersonal system (gaze/contact) while echoing ESP/EMP materials research that task-based visuals often function as contextual anchors enabling learners to rehearse professional discourse without requiring full clinical immersion (e.g., Royce; Unsworth; and broader multimodal pedagogy discussions in Kress/Jewitt).

A second Career Paths example, CPM-B3P8, shows how indirect gaze coexists with interpersonal cues of care and rapport in patient communication.

3

Communicating with patients



To: All Staff

Subject: Communicating with Patients

Clear communication is vital to providing the best care possible. However, recently we have not been getting all the information we need.

If you believe you are not getting adequate information from a patient, remember some of the possible causes of **miscommunication**:

- **Poor Health Literacy** - Uninformed patients may not understand important terms.
- **Negative Preconceptions** - Patients may not trust medical professionals. They may experience **anxiety** when forced to interact with them.
- **Cultural Differences** - Different cultures may have different approaches to dealing with illness and health care.

If you recognize any of these issues, a few simple steps can be all that is required to improve communication:

- Establish a good **rapport** with the patient. Patients heal better in a **therapeutic** environment.
- Show **empathy** and understanding at all times.
- Be aware of how your **verbal** and **nonverbal** cues are affecting the interaction.
- Avoid using **jargon** as much as possible. This may confuse or intimidate the patient.
- Ask **open-ended questions** to signal to patients that you care about their input.

Remember, everyone shares the responsibility of effective communication with patients. The more information we can give and receive, the better care we can provide.

Get ready!

- 1 Before you read the passage, talk about these questions.

- 1 What are some possible causes of miscommunication with patients?
- 2 How can medical professionals improve communication with patients?

Reading

- 2 Read the memo on communicating with patients. Then, choose the correct answers.

- 1 What is the main idea of the article?
 - A Medical professionals must learn to trust patients.
 - B There are easy ways to improve communication with patients.
 - C Verbal and nonverbal communication are equally important.
 - D Poor communication with patients can lead to life-threatening risks.
- 2 Which of the following is NOT a cause of miscommunication?
 - A asking too many open-ended questions
 - B low levels of health education
 - C diverse cultural beliefs
 - D mistrust of medical personnel
- 3 What is an appropriate way to improve communication with patients?
 - A Use a lot of technical words.
 - B Avoid becoming too friendly.
 - C Limit nonverbal communication.
 - D Show proper empathy to patients.

Vocabulary

- 3 Match the words (1-6) with the definitions (A-F).

- | | |
|-------------------------|---------------|
| 1 — verbal | 4 — rapport |
| 2 — therapeutic | 5 — anxiety |
| 3 — open-ended question | 6 — nonverbal |

- A a feeling of understanding and a connection between two people
- B a feeling of nervousness and worry
- C an inquiry that requires an answer longer than a one-word response
- D communication without the use of spoken language
- E spoken
- F helping someone to heal or feel relaxed

GVD emphasises that interpersonal meaning is not only gaze; it emerges from the orchestration of gaze with distance, angle, and modality cues that jointly regulate involvement. CPM-B3P8 uses multiple photos depicting clinician-patient interaction (labels such as *empathy*, *rapprochement*, *miscommunication*, *anxiety*). Participants' gaze is directed at each other or situationally (not at the learner), and the shots are generally medium-range, with facial expression and interactional posture visible. The photographic modality (realism, colour, detail) increases social presence compared to schematic medical diagrams. Consequently, learners are positioned as professionally oriented interpreters of interaction, not simply as readers of biomedical facts. The images cue interpersonal competencies (empathy/rapprochement) as legitimate components of “*medical meaning*”, which is pedagogically consequential in EMP contexts that increasingly emphasise communication skills. Importantly, this positioning creates space for multimodal pedagogy: students can be guided to read nonverbal cues, infer stance, and then rehearse language choices aligned with patient-centred care – an orientation strongly relevant to Vietnamese EMP classrooms seeking to move beyond vocabulary-heavy instruction. This page illustrates GVD's interpersonal resources (contact, distance, modality) in a way consistent with multimodal discourse studies arguing that clinical professionalism is semiotically constructed through images that legitimise affective and relational dimensions of practice (e.g., Goodwin's “professional vision” line of reasoning; Jewitt; and applied multimodal pedagogy scholarship).

4.4.2 Social Distance, Angle, and Epistemic Relations

Beyond gaze, corpus-level evidence (Appendix C; Appendix D) shows a stable contrast in how the two textbooks deploy social distance and angle to regulate epistemic relations. In *Medical Terminology*, visuals frequently adopt

either long-shot schematisation (system-level diagrams) or decontextualised object display, both of which reduce interpersonal involvement and support objectification of knowledge. Moreover, angles tend to be neutral/standardised – consistent with the pedagogic aim of presenting knowledge as stable, canonical, and “seen correctly” through institutionalised representation. In *Career Paths: Medical*, medium shots and eye-level perspectives appear more frequently, producing a sense of engagement: learners are invited to recognise themselves as future participants in communicative and procedural settings rather than solely as observers of anatomy.

However, the epistemic effects of distance and angle are not merely stylistic. When learners repeatedly encounter long-shot schematics such as MT-P148 (Figure 4.1), they are trained into a form of biomedical literacy that privileges systemic relations and definitional stability. Likewise, decontextualised displays such as MT-P205 (Figure 4.2) cultivate classificatory competence – what counts as knowing is recognising and naming. By contrast, medium-shot workplace scenes such as CPM-B1P22 and CPM-B3P8 (Figures 4.3-4.4) normalise a different epistemic stance: knowing involves interpreting actions, interpersonal cues, and communicative appropriacy. Critically, this is where multimodal pedagogy becomes *emergent from design*: the textbook’s interactive choices preconfigure what teachers can plausibly do in classrooms – MT invites explicit instruction in diagram reading and terminology decoding, whereas CPM more readily supports role-play, scenario-based discussion, and communication-focused practice. In Vietnamese EMP settings, this matters because it offers a principled route to rebalancing language-dominant pedagogy toward multimodal disciplinary literacy, without treating images as “*decoration*”.

Taken together, the corpus confirms the GVD proposition that interpersonal meaning is orchestrated through contact, distance, and angle, while also supporting multimodal education research that different visual regimes enact different “*ways of knowing*” (professional vision) and therefore different pedagogic affordances in EMP instruction.

4.4.3 Angle and Authority

Within the Grammar of Visual Design, angle encodes interpersonal relations between viewer and represented content, especially authority, power, and epistemic distance. Vertical angle construes power asymmetry (looking down/up), whereas horizontal angle indicates involvement versus detachment (frontal vs oblique) (Kress & van Leeuwen, 2021).

Across the corpus, MT predominantly relies on neutral and canonical angles, especially in diagrams, cross-sections, and schematic displays. This visual neutrality reduces interpersonal negotiation and naturalises medical knowledge as stable and already-authorised, positioning learners as epistemic observers of a standardised biomedical viewpoint (Lemke, 1998; Chegeni et al., 2016, Kress & van Leeuwen, 2021). In this sense, authority is produced through correct seeing: the image foregrounds an institutionalised way of viewing the body rather than an experience of clinical action.

By contrast, CPM frequently employs eye-level angles in contextual photographs and scenario-based visuals, often combined with flexible horizontal positioning (frontal/oblique). These choices construct authority less as canonical representation and more as practice-oriented legitimacy, inviting learners to align with professional situations and workplace roles rather than merely inspect abstract structures (Goodwin, 1994; Bezemer & Kress, 2016).

Overall, MT establishes authority through canonical visual viewpoints, while CPM constructs authority through situated professional visibility.

4.4.4 Modality Cues and the Construction of Visual Credibility

In GVD, modality cues (e.g., realism, detail, colour saturation, and representational conventions) function as resources for constructing visual credibility, though credibility is always socially and discursively calibrated rather than inherently “*objective*” (Kress & van Leeuwen, 2021).

In MT, credibility is typically grounded in scientific abstraction and standardisation. Diagrammatic clarity, controlled detail, systematic labelling, and conventionalised colouring produce a form of scientific authority: visuals appear trustworthy because they stabilise anatomical relations in a legible, taxonomic, and replicable mode of representation (Unsworth, 2008; Bezemer & Kress, 2016). Here, modality is not aimed at lifelike authenticity but at epistemic control - making structures identifiable and classifiable.

In CPM, credibility is more often constructed through workplace authenticity. Photographic modality, social actors, tools, and institutional settings foreground a recognisable clinical world, legitimising medical meaning as situated practice rather than conceptual system. Such visuals support professional orientation by making interactional cues and procedural context “*seeable*,” which aligns with multimodal accounts of how learners are socialised into disciplinary practices (Goodwin, 1994; Jewitt, 2014).

Thus, MT tends to build credibility through scientific conventionality, whereas CPM builds credibility through contextual realism and professional recognisability.

4.4.5 Interim Synthesis: Interactive Meaning Profiles of MT and CPM

Synthesising Sections 4.4.1 - 4.4.4, interactive meaning across the two textbooks consolidates into two coherent viewer-positioning profiles.

Medical Terminology (MT): an observer-classifier stance.

Offer images, neutral angles, and conventional scientific modality cues repeatedly position learners as detached observers of canonical biomedical knowledge. The learner's role is to inspect, recognise, and classify anatomical entities and relations, supporting a terminology-driven epistemic orientation where knowing is closely tied to identification and definition (Kress & van Leeuwen, 2021; Unsworth, 2008).

Career Paths: Medical (CPM): an apprentice-participant stance.

Contextualised images, frequent eye-level perspectives, and photographic modality cues more often position learners as professionally oriented observers of action, supporting apprenticeship-like engagement with practice scenarios. In this profile, knowing is framed as interpreting workplace activity, roles, and interpersonal cues, thereby expanding medical meaning toward communication and professional conduct (Goodwin, 1994; Bezemer & Kress, 2016).

Importantly, these interactive profiles provide a principled bridge to Chapter V: MT's visual regime is likely to support dense intersemiotic elaboration (naming/defining), whereas CPM's regime is likely to support contextual enhancement in which visuals anchor tasks and situational meaning-making (Unsworth, 2008).

4.5 Compositional Meaning: Visual Information Organisation and Pedagogic Control

Across the entire corpus of 40 page ensembles (Appendix A), compositional meaning functions as a crucial semiotic mechanism through which medical knowledge is ordered, prioritised, and pedagogically staged. Within IMAF–Stratum 1, composition is not treated as a merely aesthetic or technical dimension; rather, it is analysed as an epistemic resource that guides learners in distinguishing between *core disciplinary knowledge*, supporting

explanation, and *pedagogic application*. Drawing on the GVD (Kress & van Leeuwen, 2006), this section focuses on two interrelated compositional principles: information value (how meaning is spatially distributed) and salience and framing (how attention and hierarchy are constructed).

Crucially, the analysis is grounded in corpus-level regularities identified in Appendix C (Page-Ensemble Analytical Matrix) and page-specific traceability documented in Appendix D. At the same time, selected page ensembles are examined in depth to demonstrate how these compositional principles are concretely realised in authentic medical English textbooks.

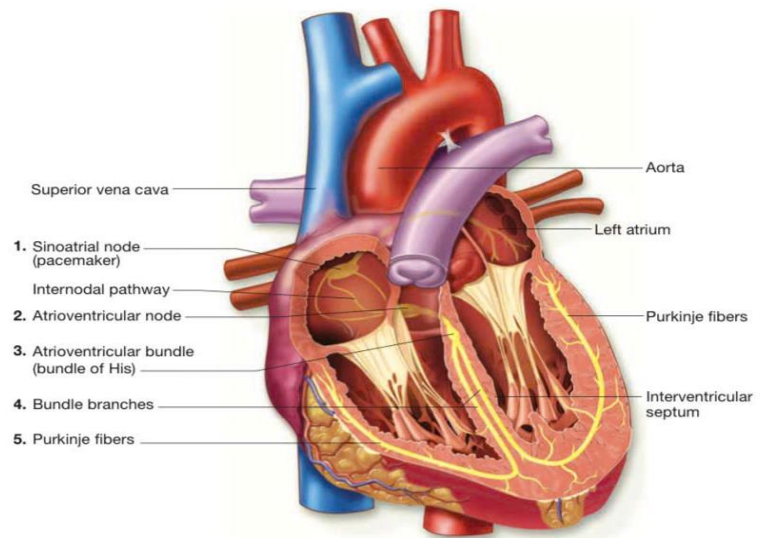
4.5.1 Information Value and Layout Structures

A consistent contrast emerges between the two textbook series in how information value is realised through layout. MT predominantly adopts system-oriented compositional structures, whereas CPM relies on linear, task-based pedagogic sequencing. Importantly, these differences recur systematically across the corpus and therefore reflect underlying pedagogic epistemologies rather than isolated design choices.

Medical Terminology: centre-margin and ideal-real structures

In MT, information value is most frequently organised through centre-margin and ideal-real configurations, which foreground anatomical or physiological systems as stable objects of knowledge before moving toward explanation and practice.

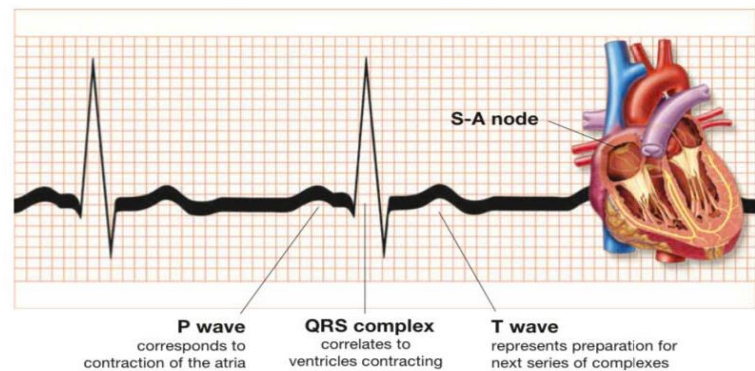
■ **Figure 5-6** The conduction system of the heart; traces the path of the electrical impulse that stimulates the heart chambers to contract in the correct sequence.



Med Term Tip

The electrocardiogram, referred to as an EKG or ECG, is a measurement of the electrical activity of the heart (see Figure 5-7 ■). This can give the physician information about the health of the heart, especially the myocardium.

■ **Figure 5-7** An electrocardiogram (EKG or ECG) wave record of the electrical signal as it moves through the conduction system of the heart. This signal stimulates the chambers of the heart to contract and relax in the proper sequence.



PRACTICE AS YOU GO

A. Complete the Statement

1. The study of the heart is called _____.
2. The three layers of the heart are _____, _____, and _____.
3. The impulse for the heartbeat (the pacemaker) originates in the _____.

Within GVD, ideal-real structures typically position generalised or abstract knowledge in the upper zone of the page, while specific or applied information is placed below, thereby naturalising a top-down epistemic progression (Kress & van Leeuwen, 2006). This pattern is clearly instantiated in MT-P154. On this page, the cardiac conduction system is presented as a large, centrally dominant anatomical diagram occupying the upper and central visual field. The diagram is meticulously labelled, with numbered pathways

guiding the reader through the system. Beneath this, an ECG waveform is placed horizontally, visually subordinated yet directly aligned with the anatomical representation. At the bottom of the page, a framed “*Practice as You Go*” section introduces short completion tasks.

Compositorily, this arrangement constructs medical knowledge as something that must first be seen as a coherent system before it can be interpreted or applied. The learner is positioned epistemically as an observer of biomedical structure, expected to internalise a stable conceptual model prior to engaging in recall or practice. The ECG strip, although practical in nature, remains visually dependent on the anatomical diagram above, reinforcing the primacy of conceptual understanding over procedural engagement.

This pattern closely aligns with international research on science and medical textbooks, which has shown that diagram-centred, ideal-real layouts function to stabilise disciplinary knowledge and legitimise visual explanation as epistemically authoritative (Unsworth, 2008; Bateman, 2014). Vietnamese multimodal textbook research reports comparable tendencies that Vietnamese EFL and science textbooks frequently employ centre-margin layouts to foreground conceptual clarity and reduce interpretive ambiguity. (Tran. & Le., 2021, Pham, 2021) The present analysis shows that MT follows the same compositional logic, despite operating in a specialised medical English context.

A comparable epistemic orientation is evident in MT-P476, further reinforcing the systematic nature of this compositional strategy.

Muscles of the Eye

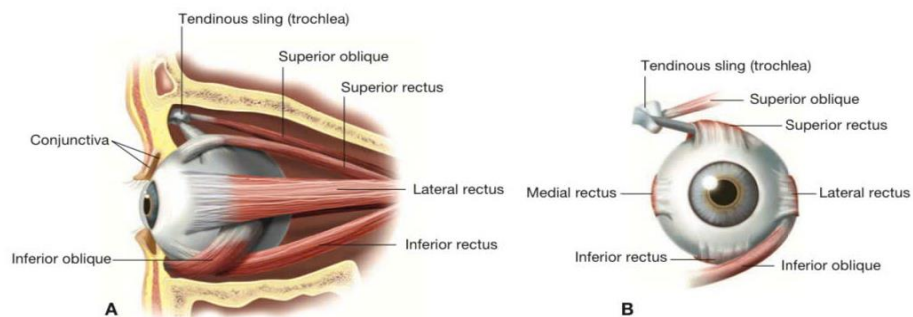
oblique muscles (oh-BLEEK)

rectus muscles (REK-tus)

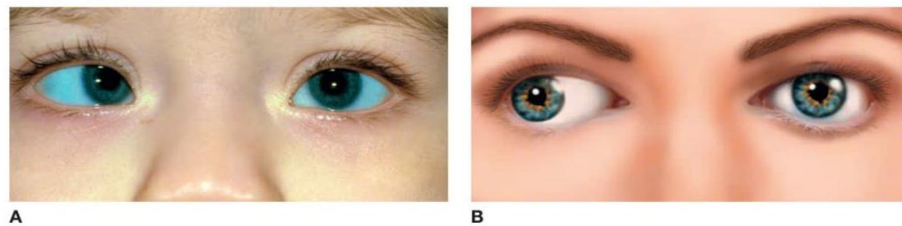
Med Term Tip

Like many other muscles, the names *rectus* and *oblique* provide clues regarding the direction of their fibers, or their *line of pull*. Rectus means *straight* and oblique means *slanted*. Rectus muscles have a straight line of pull. Because the fibers of an oblique muscle are slanted on an angle, they produce rotation.

Six muscles connect the actual eyeball to the skull (see Figure 13-3 ■). These muscles allow for change in the direction of each eye's sightline. In addition, they provide support for the eyeball in the eye socket. Children may be born with a weakness in some of these muscles and may require treatments such as eye exercises or even surgery to correct this problem, commonly referred to as crossed eyes or *strabismus* (see Figure 13-4 ■). The muscles involved are the four **rectus** and two **oblique muscles**. Rectus (meaning *straight*) muscles pull the eye up, down, left, or right in a straight line. Oblique muscles are on an angle and produce diagonal eye movement.



■ **Figure 13-3** The arrangement of the external eye muscles, A) lateral and B) anterior views.



■ **Figure 13-4** Examples of common forms of strabismus. A) Esotropia with the right eye turning inward. (Biophoto Associates/Science Source) B) Exotropia with the right eye turning outward. (Gwen Shockley/Science Source)

The Eyelids

cilia (SIL-ee-ah)

sebaceous glands (seh-BAY-shus)

eyelashes

What's In A Name?

Look for these word parts:

seb/o = oil

-ous = pertaining to

A pair of eyelids over each eyeball provides protection from foreign particles, injury from the sun and intense light, and trauma (see again Figure 13-1). Both the upper and lower edges of the eyelids have **eyelashes**, or **cilia**, that protect the eye from foreign particles. In addition, **sebaceous glands** located in the eyelids secrete lubricating oil onto the eyeball.

Here, eye muscle diagrams occupy the most salient and central visual space, while framed pedagogic boxes (“*Med Term Tip*”, “*What’s in a Name?*”) and photographic examples of strabismus are positioned as secondary elaborations. The layout thus constructs a clear hierarchy in which system knowledge precedes exemplification, guiding learners toward a classificatory and analytical mode of knowing. This resonates strongly with findings from Vietnamese analyses of science textbooks, where labelled

diagrams are shown to function as epistemic anchors that structure learners' engagement with specialised content (Le, 2017).

Career Paths: Medical-linear, task-based layouts

In contrast to MT, CPM pages consistently organise information value through linear pedagogic progression, where meaning unfolds across a sequence of instructional steps rather than around a single conceptual display.

10
Traumatic injuries

Heartland
Hospital

ER TRIAGE GUIDELINES

Heartland Hospital strives to provide the best care it can with limited **resources**. To ensure that each patient is seen at the appropriate time and receives quality services, please follow these guidelines for the **triage** process.

Trauma patients who arrive with obviously **critical** and life-threatening injuries receive the highest **priority**. They should be taken to an available trauma room immediately. Any necessary staff and equipment should be devoted to them until their vital signs are stabilized.

All other patients should report to the **triage station** to undergo an assessment by a nurse. The nurse will utilize the **Emergency Severity Index** to determine the severity of a patient's injuries and when they should be seen. This is a common-sense system that ranks patients according to the danger level of their situation. For example, an open **wound** will rank higher on the Index than a **minor** illness such as a low fever. The nurse will assign an **ESI level** between 1 and 5, with 1 being the most serious, to each patient. In this way, patients will be seen according to the **acuity** of their injuries.

triage

trauma

resources

wound

triage station

Get ready!

1 Before you read the passage, talk about these questions.

- 1 What kind of patients are seen in an emergency room?
- 2 How does the triage process generally work?

Reading

2 Read the poster on treating patients in an emergency room. Then, mark the following statements as true (T) or false (F).

- 1 ☐ Fevers are not ranked on the Emergency Severity Index.
- 2 ☐ A patient with a level one ESI rating has the highest priority.
- 3 ☐ Only an emergency room doctor should assign a patient an ESI rating.

Vocabulary

3 Match the words (1-7) with the definitions (A-G).

1 ☐ priority	5 ☐ triage
2 ☐ resource	6 ☐ minor
3 ☐ wound	7 ☐ critical
4 ☐ ESI level	

- A a unit of measurement of a trauma patient's injuries
- B the process of determining the order to treat trauma patients based on the severity of their injuries
- C something that is more important than other things
- D requiring immediate medical attention
- E an injury in which a part of the body is damaged
- F not very serious
- G anything a person uses to achieve a goal

22

From a GVD perspective, linear layouts distribute information value temporally rather than hierarchically, aligning page design with pedagogic action and classroom workflow (Kress & van Leeuwen, 2006). This is clearly observable in CPM-B3P22. The page is structured around a poster-style reading text (“*ER Triage Guidelines*”) positioned on the left, visually marked by bold colour blocks and headline typography. On the right, a vertically ordered sequence of sections (“*Get ready*”, “*Reading*”, “*Vocabulary*”) guides learners through a clear task pathway. Supporting images are embedded within the poster but do not dominate the page visually.

This compositional design positions the learner not as a detached observer of medical systems, but as a participant in professional activity. Knowledge is construed as something to be used and enacted through tasks, rather than merely contemplated. Such layout strategies are widely documented in international ESP research, where vocational textbooks are shown to simulate workplace practices through activity sequencing and task salience (Hyland, 2006; Bhatia, 2008). In the Vietnamese context, similar observations have been made regarding the growing prominence of task-oriented layouts in communicative English materials, reflecting shifts in pedagogic expectations and classroom practice (Le, 2017).

4.5.2 Salience Patterns and Attention Management

Beyond information placement, compositional meaning is further realised through salience – what draws immediate attention – and framing – how elements are grouped or separated into meaningful units. Across the corpus, these resources construct distinct visual hierarchies of importance in MT and CPM.

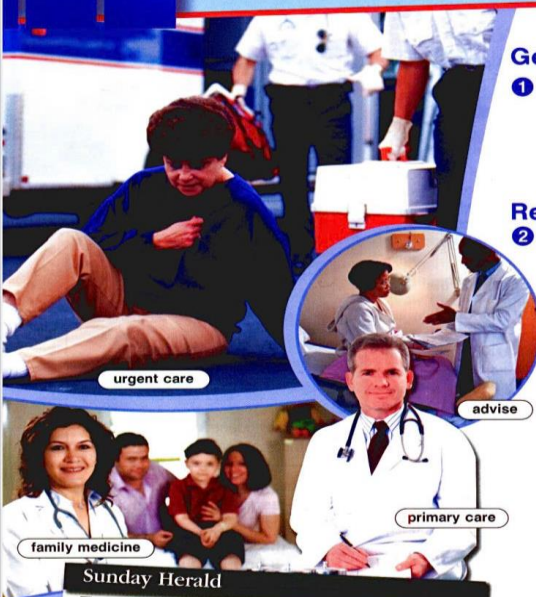
In MT, salience is achieved primarily through diagram size, colour saturation, and labelling density, while framing devices such as boxes, margins, and sectional boundaries regulate how specialised knowledge is segmented. Revisiting MT-P154 and MT-P476, anatomical diagrams clearly

function as the most salient elements, visually outweighing text and exercises. Framed pedagogic boxes act as authorised interpretive guides, narrowing the learner's reading path and reinforcing terminological precision.

This compositional hierarchy aligns with an epistemology that values classification, accuracy, and conceptual stability, a tendency widely recognised in international medical discourse studies and Vietnamese analyses of science and EFL textbooks (Tran, 2024; Nguyen, 2020).

In CPM, however, salience is redistributed toward instructional headings, numbered tasks, and response spaces. Images, although visually engaging, are often framed as contextual cues rather than focal objects of sustained analysis.

13 Family medicine



Sunday Herald

Family Medicine — by Dr. John Olson, MD

The Heart of Primary Care

For decades the family medical clinic has been the first point of **primary care** for millions of patients. However, today patient demand threatens to overcome the supply of family medicine **practitioners**. Thus it is important to emphasize the need for these vital health care professionals.

Family practitioners provide a variety of health care services to all patients, regardless of age. Not only do they treat injuries and illnesses, they also provide **preventative** care such as routine checkups and immunizations. They **counsel** patients about their health care options and **advise** a recommended course of action. They are a vital source of **health education**. If needed, they **refer** patients to specialists for specific diseases and injuries. Even though they do not work in an emergency room, many family practitioners provide **urgent care** to patients requiring immediate attention.

Family medicine will always be at the core of the health care system. Medical professionals should be encouraged to follow this rewarding career.

Get ready!

① Before you read the passage, talk about these questions.

- 1 How is family medicine structured in your country?
- 2 What are some of the duties of a family practitioner?

Reading

② Read the newspaper article on family medicine. Then, choose the correct answers.

- 1 What is the main idea of the article?

A Family medicine is not as important as in the past.

B More medical professionals should pursue a career in family medicine.
- C Family practitioners should not provide urgent care.
- D Family practitioners are the fastest growing group of medical professionals.

- 2 Which of the following do family practitioners NOT provide?

A urgent care

C specialized care
- B checkups and immunizations
- D health education

- 3 Which of the following is part of preventative care?

A regular checkup

B treating an injury
- C urgent care
- D prescribing medicine

Vocabulary

③ Match the words (1-5) with the definitions (A-E).

1 — primary care
 2 — health education
 3 — family medicine

4 — urgent care
 5 — preventative

- A medicine that provides long-term, general health care for all individuals
- B focusing on preventing diseases from occurring rather than curing them
- C the act of teaching people about their bodies and good health habits
- D the first and main point of medical assistance for patients in a health care system
- E health care provided to immediately respond to an injury or illness, but not in an emergency room

28

On CPM-B3P28 (“*Communicating with patients*”), the most salient elements are the section header and boxed task instructions. Photographic images depicting doctor-patient interaction are arranged in a uniform grid and labelled with keywords such as *empathy*, *anxiety*, and *rappport*. These images support the pedagogic message but remain subordinate to the textual tasks that structure learner activity.

Epistemically, this hierarchy signals that professional competence in medical English is demonstrated through appropriate communicative response, rather than mastery of visual detail. This finding is consistent with international multimodal analyses of vocational ELT materials and with Vietnamese research highlighting the communicative orientation of contemporary English language teaching (Nguyen, 2000; Le, 2017).

4.5.3 *Framing, Segmentation, and Integration*

In GVD, framing signals whether elements belong to one meaning unit or should be processed separately (Kress & van Leeuwen, 2006). Across the corpus, MT and CPM deploy framing to enact different pedagogic epistemologies. MT tends toward strong framing (boxes, margins, clear separation of diagram/prose/practice), which modularises medical knowledge into discrete units, reducing interpretive variability and supporting a classificatory, accuracy-driven logic typical of scientific discourse (Lemke, 1998; Unsworth, 2008). CPM more often uses integrated ensembles, embedding images within task sequences and reading activities so that meaning is distributed across modes and staged as pedagogic action - an orientation consistent with vocational/ESP materials that simulate workplace routines through multimodal task design (Hyland, 2006; Bhatia, 2008). Framing therefore functions as boundary-making: MT frames knowledge as inspectable objects; CPM frames knowledge as enactable practice.

4.5.4 Reading Paths and Multimodal Navigation

Composition also scripts reading/learning routes by organising attention through placement, hierarchy, and sequencing cues (Kress & van Leeuwen, 2006; Bateman, 2014). In MT, navigation is typically diagram-anchored: learners move from a salient schematic to labels/captions and then to definitional prose and controlled exercises. This supports a top-down conceptual pathway, aligning with research showing diagrams often operate as epistemic nuclei in science/medical textbooks, with text stabilising visual meaning (Lemke, 1998; Unsworth, 2008). In CPM, navigation is more workflow-like: headings, numbered prompts, and staged tasks guide learners through “*prepare-read-practise-perform*” sequences. Images mainly contextualise tasks rather than demand prolonged visual inspection, matching ESP accounts of task progression as a key organising principle in professional English materials (Hyland, 2006; Bhatia, 2008).

4.5.5 Interim Synthesis

Compositional resources in the corpus do not merely “present” content; they control access to knowledge by delimiting units (framing) and scripting routes (reading paths). MT’s profile (strong framing + diagram-centred navigation) supports conceptual extraction and terminological consolidation in a canonised “scientific seeing” mode (Kress & van Leeuwen, 2006; Lemke, 1998). CPM’s profile (integrated ensembles + task-sequenced navigation) supports procedural/interactional enactment characteristic of vocational ESP design (Hyland, 2006; Bateman, 2014). These compositional contrasts specify the conditions under which Chapter V can explain how image–text coupling either stabilises conceptual authority (MT) or scaffolds situated professional practice (CPM).

4.6 Integrated Visual Meaning Profiles: Consolidation of Findings for RQ1

This section consolidates the findings of Chapter IV by integrating the three visual meaning systems operationalised in IMAF–Stratum 1, namely representational (A1), interactive (A2), and compositional meaning (A3). In line with multimodal social semiotics, the purpose is not to treat images as supplementary decoration, but to specify how visual design functions as an epistemic resource that organises medical knowledge and positions learners in relation to it (Kress & van Leeuwen, 2006, 2021; Bezemer & Kress, 2016). By drawing on corpus-level patterns established in Appendix C and the traceability procedures documented in Appendix D, the analysis moves beyond page-by-page commentary to formulate two textbook-level “visual meaning profiles”. These profiles capture stable design regularities across the corpus and provide an empirically grounded answer to RQ1, while also preparing the analytical transition to Chapter V, where visual-verbal coupling will be explained intersemiotically (Unsworth, 2008; Bateman, 2014).

4.6.1 Visual Meaning Profile of MT: “Terminological-Conceptual Regime”

The visual semiotic profile of Medical Terminology: A Living Language can be characterised as a terminological-conceptual regime, insofar as its images recurrently construe medical knowledge as a stable system of relations designed for classification, inspection, and naming. At the representational level, MT consistently privileges conceptual representation, particularly classificatory and analytical structures, in which anatomical entities are displayed as detachable objects and organised into taxonomies or part-whole configurations. Such visual choices enact an epistemology of medical knowledge as something that can be stabilised through categorisation,

segmentation, and definitional alignment, a tendency that resonates with the wider semiotic condensation typical of scientific discourse, where dynamic processes are reconfigured into stable, nameable objects of knowledge (Lemke, 1998; Kress & van Leeuwen, 2006). Importantly, this conceptual logic is not merely stylistic; it functions pedagogically by externalising relations that learners must master in order to acquire medical terminology, thereby making classification and spatial decomposition visually “*learnable*” (Unsworth, 2008; Bezemer & Kress, 2016).

At the interactive level, MT further consolidates this regime through patterned viewer positioning. The corpus shows that images overwhelmingly operate as offer visuals, with minimal interpersonal demand and limited social presence. The reduced prominence of gaze and direct address, together with the prevalence of neutral angles and controlled distance, positions learners as detached observers of canonical biomedical knowledge rather than as participants in situated practice (Kress & van Leeuwen, 2006). In this sense, interactive meaning reinforces representational abstraction: the viewer is invited to inspect medical objects as epistemically stable entities, and authority is naturalised through standardised viewpoints that present “the body” as a universal, discipline-legitimate object rather than an embodied, context-dependent experience. This supports Lemke’s (1998) argument that scientific meaning-making often relies on semiotic strategies that suppress contingency in order to stabilise epistemic certainty.

At the compositional level, MT sustains strong pedagogic control through diagram-centred hierarchy and robust framing. Across ensembles, diagrams tend to occupy the most salient positions and carry the principal meaning load, while supporting verbal text and boxed prompts function as authorised pathways that regulate interpretation. These compositional

strategies are consistent with research showing that science and medical textbooks often use layout to stabilise disciplinary knowledge, guide attention, and minimise interpretive ambiguity by structuring reading routes around privileged explanatory anchors (Unsworth, 2008; Bateman, 2014). Taken together, MT's representational abstraction, interpersonal distancing, and hierarchical layout converge into a coherent semiotic regime that trains learners to "see" medicine as terminologically decomposable and conceptually structured. Within this regime, what counts as successful engagement is largely the ability to recognise, label, and organise biomedical entities within stable conceptual frameworks (Kress & van Leeuwen, 2006, 2021; Bezemer & Kress, 2016)

4.6.2 Visual Meaning Profile of CPM: "Contextual-Professional Regime"

In contrast, the visual semiotic profile of Career Paths: Medical can be characterised as a contextual-professional regime, in which images recurrently construe medical knowledge as enacted practice embedded in workplace activity and interpersonal interaction. At the representational level, CPM tends to foreground narrative representation, depicting sequences of action, procedural routines, and communicative encounters involving human participants. Through these choices, medical meaning is framed not primarily as anatomical structure to be classified, but as professional activity to be recognised, interpreted, and rehearsed. This orientation aligns with Goodwin's (1994) account of professional vision, where members of a field learn to perceive relevant cues and organise perception in ways that are socially and institutionally legitimate. CPM images do not only visualise content; they model practice, offering novices semiotic access to professional

scenarios that would otherwise remain inaccessible in a language classroom context (Bezemer & Kress, 2016).

This contextual orientation is strengthened by CPM's interactive design. While direct gaze towards the viewer may remain limited, the repeated presence of human actors, workplace settings, and communicative configurations increases interpersonal immediacy and positions learners as professionally implicated observers rather than detached classifiers. The viewer is invited to notice not only medical objects but also roles, stances, and interactional expectations, thereby expanding the epistemic scope of "medical meaning" to include pragmatic competence and professional conduct (Kress & van Leeuwen, 2006; Bezemer & Kress, 2016). Modality cues often intensify this effect: compared with schematic diagrams, workplace photographs and contextualised visuals carry higher realism and situational specificity, which supports a sense of authenticity and legitimises interpersonal action as a learnable domain of medical English discourse.

At the compositional level, CPM consistently organises meaning through linear task sequencing. Headings, numbered steps, and response spaces guide learners along an anticipated pedagogic route in which visual cues are coordinated with reading, vocabulary, and communicative performance. This layout logic corresponds closely to ESP scholarship describing vocational materials as staged instructional simulations that convert situational recognition into language practice through task progression (Hyland, 2006; Bhatia, 2008). In CPM, images rarely function as isolated conceptual anchors; instead, they are integrated into activity structures that prioritise doing over contemplating, rehearsal over classification. As a result, representational narrativity, interpersonal implication, and compositional sequencing converge into a contextual–

professional regime that trains learners to notice practice-relevant cues and to rehearse professional communication as part of medical meaning-making (Bezemer & Kress, 2016; Hyland, 2006).

4.6.3 Cross-textbook Comparison: Disciplinary Vision and Pedagogic Orientation

Juxtaposing the two integrated profiles reveals that MT and CPM instantiate two distinct visual epistemologies, each cultivating a different form of disciplinary vision and learner positioning. MT privileges a terminological-conceptual mode of knowing, in which medical knowledge is stabilised through abstraction, categorisation, and hierarchical organisation. CPM, by contrast, privileges a contextual-professional mode of knowing, in which medical knowledge is made visible through enacted procedures, workplace roles, and communicative encounters. These epistemological differences are not reducible to image type alone; they are systematically realised across representational content, interactive stance, and compositional control, confirming that visual meaning is patterned at the level of the textbook ecology rather than produced by isolated design decisions (Kress & van Leeuwen, 2006; Unsworth, 2008).

The two textbooks also diverge in the epistemic identities they offer to learners. MT repeatedly positions learners as observers and classifiers of canonical biomedical objects, thereby privileging recognition, naming, and conceptual mastery as the dominant form of competence. CPM, in contrast, more often positions learners as apprentices to professional practice, where competence involves interpreting situations, aligning with professional roles, and rehearsing appropriate communication. In multimodal terms, these are not simply different pedagogic “styles”; they are different semiotic pathways into

what counts as medical knowledge and medical English literacy (Bezemer & Kress, 2016; Goodwin, 1994).

Finally, the comparison clarifies the design affordances each regime makes available for instruction. MT's semiotic organisation strongly supports diagram reading, terminology decoding, and conceptual organisation, but it may require additional pedagogic work to connect conceptual knowledge to situated communication. CPM's design, conversely, more readily affords scenario-based discussion, role-play, and pragmatic rehearsal, although visual detail is often subordinated to task flow rather than developed as an object of sustained analytic attention. These contrasting affordances matter for Vietnamese EMP contexts, where pedagogy has historically emphasised terminological mastery but increasingly faces demands to address professional communication and workplace interaction. In this respect, the two regimes should not be treated as mutually exclusive alternatives but as complementary emphases that preconfigure different trajectories of multimodal medical literacy (Hyland, 2006; Kress & van Leeuwen, 2006; Bezemer & Kress, 2016).

4.7 Concluding Remarks

This chapter has answered RQ1 by providing a corpus-wide account of how medical textbook images construct meaning within the visual mode (IMAF-Stratum 1). Across 40 page ensembles, the findings confirm that images are not decorative supplements but epistemic resources that stabilise disciplinary knowledge, position learners in relation to it, and regulate access through compositional design.

At the representational level, Medical Terminology predominantly employs conceptual structures that construe medical knowledge as stable, classifiable, and terminologically decomposable, whereas Career Paths:

Medical favours narrative representations that situate knowledge in workplace actions and communicative routines. These representational priorities are reinforced interactively: MT typically positions learners as detached observers of canonical biomedical objects, while CPM more often constructs an apprentice-like stance through contextualised professional scenes. Compositional choices further consolidate these regimes, with MT relying on diagram-centred hierarchy and strong framing, and CPM foregrounding linear task sequencing and integrated page ensembles.

Overall, the analysis identifies two coherent visual meaning profiles - a terminological-conceptual regime in MT and a contextual-professional regime in CPM - showing that visual meaning is systematically patterned at textbook level rather than determined by isolated images. These visual profiles establish the necessary foundation for Chapter V, which explains how visual and verbal resources couple intersemiotically to produce multimodal pedagogic meaning.

CHAPTER V: INTERSEMIOTIC MEANING IN MEDICAL TEXTBOOKS

Chapter V shifts the analysis from visual meaning in isolation to intersemiotic meaning, examining how images and language are systematically orchestrated within page ensembles to construct medical knowledge. Building on Chapter IV's visual profiles of *Medical Terminology* and *Career Paths: Medical*, this chapter explains how patterned image-text relations stabilise concepts, scaffold procedures, and enact professional roles across the corpus. Using an intersemiotic framework grounded in multimodal social semiotics, the analysis demonstrates how multimodal design transforms visual and verbal resources into learnable medical knowledge, thereby establishing the bridge to the pedagogical synthesis in the final sections.

5.1 Analytical Orientation and Intersemiotic Scope

Chapter V builds directly on the findings of Chapter IV by shifting the analytical focus from visual meaning alone to intersemiotic meaning, that is, how images and language work together within the textbook page ensemble to construct medical knowledge. Chapter IV has established that MT and CPM exhibit distinct visual meaning profiles – conceptual abstraction in the former and practice-oriented contextualisation in the latter. However, these visual patterns acquire pedagogic force only through their systematic coupling with verbal text.

Accordingly, Chapter V adopts an intersemiotic perspective, treating image-language relations not as simple redundancy but as a designed division of semiotic labour. This position is well established in international multimodal research, which demonstrates that educational meaning is produced through patterned relations between modes rather than through

isolated visual or verbal resources (Kress & van Leeuwen, 2021; Bezemer & Kress, 2016). More specifically, image-text relations are approached as structured logico-semantic relations, drawing on established frameworks for intersemiosis in multimodal discourse (Martinec & Salway, 2005; Royce, 2007).

The unit of analysis remains the page ensemble, as this unit preserves the integrity of multimodal meaning-making: intersemiotic relations become analytically visible only when images, labels, running text, and layout are examined as a coordinated whole. Evidence for this chapter is drawn systematically from all 40 page ensembles (Appendix A), with coding guided by the Intersemiotic Codebook (Appendix B3) and corpus-wide patterns documented in Appendix C, while Appendix D ensures full traceability of claims.

This analytical orientation is also consistent with Vietnamese discourse-analytic scholarship, which has emphasised the importance of applying international discourse frameworks through methodologically controlled, text-based analysis rather than impressionistic interpretation (Nguyen, 2000). Within this framework, Chapter V aims to explain how different intersemiotic relations – elaboration, extension, enhancement, and projection – systematically shape what kinds of medical knowing are made available to learners, thereby addressing RQ2 and preparing the synthesis developed later in the chapter.

5.2 Distribution of Intersemiotic Relations across the Corpus

This section provides a corpus-level map of how images and language relate across the 40 page ensembles (Appendix A), answering the question: Which intersemiotic relations are mobilised in these textbooks, and how are they distributed? Rather than analysing individual pages, the aim here is to

identify systematic tendencies that characterise each textbook ecology and to establish the empirical ground for the focused analyses in Sections 5.3-5.5.

Intersemiotic relations are coded using the Intersemiotic Codebook (Appendix B3), which operationalises four logico-semantic relation types – elaboration, extension, enhancement, and projection – drawing on established multimodal frameworks for image–text relations (Martinec & Salway, 2005; Royce, 2007). Corpus-wide frequencies and calibrated strengths are synthesised in the Page-Ensemble Analytical Matrix (Appendix C), while Appendix D ensures traceability from each aggregate claim to page-level evidence.

At the corpus level, a clear distributional contrast emerges between the two textbook series. In MT, elaboration is the dominant intersemiotic relation. Images (typically diagrams and schemata) and language (labels, captions, definitions) repeatedly realise same-meaning relations with functional specialisation, whereby visual structure is anchored and stabilised by terminological text. This pattern aligns with international findings that scientific and medical textbooks rely heavily on elaborative image-text coupling to objectify knowledge and reduce interpretive ambiguity (Martinec & Salway, 2005; Kress & van Leeuwen, 2021).

By contrast, CPM shows a markedly higher incidence of enhancement and projection. Here, images frequently add procedural, temporal, or conditional meaning to tasks and explanations (enhancement), or visually stage dialogue and role enactment that is realised linguistically through speech bubbles, scripts, or task prompts (projection). Such distributions are consistent with international ESP research showing that vocational materials favour intersemiotic relations that support action, sequencing, and professional interaction (Royce, 2007; Bezemer & Kress, 2016).

Importantly, this contrast is not absolute: all four relation types occur across the corpus. However, the relative dominance of elaboration in Medical Terminology and enhancement/projection in Career Paths is robust across all 40 ensembles, as evidenced by Appendix C. This distributional mapping demonstrates that intersemiotic relations are systematically patterned, not incidental. It therefore provides the necessary analytic foundation for Sections 5.3-5.5, where each relation type is examined in depth to explain how different multimodal designs support distinct epistemic and pedagogic orientations.

5.3 Intersemiotic Elaboration and Conceptual Stabilisation

This section examines how intersemiotic elaboration functions as a key mechanism for conceptual stabilisation across the corpus, with particular emphasis on MT. Within the IMAF, elaboration refers to image–text relations in which one semiotic mode restates, specifies, or exemplifies the experiential meaning of another, thereby reducing ambiguity and consolidating disciplinary understanding (Royce, 2007).

Building on the distributional patterns identified in Section 5.2, the analysis demonstrates that elaboration operates systematically and strongly in MT, while appearing selectively and instrumentally in CPM. This contrast is not incidental; rather, it reflects two distinct epistemic orientations toward what it means to “*know*” medicine in EMP contexts.

5.3.1 Systematic Elaboration and Conceptual Stabilisation in Medical Terminology

In MT, intersemiotic elaboration functions as the dominant strategy for stabilising medical knowledge as a conceptual system. Images, labels, and definitions are tightly coupled in a redundant yet specialised configuration that anchors terminology visually and linguistically, thereby constructing

biomedical knowledge as stable, classifiable, and retrievable. This stabilising logic is clearly instantiated in MT-P122.

Anatomy and Physiology of the Muscular System

Med Term Tip

The term *muscle* is the diminutive form of the Latin word *mus* or "little mouse." This is thought to describe how the skin ripples when a muscle contracts, like a little mouse running.

What's In A Name?

Look for these word parts:
cardi/o = heart
-ac = pertaining to
in- = not

muscle fibers

muscles

Muscles are bundles of parallel **muscle fibers**. As these fibers contract (shorten in length) they produce movement of or within the body. The movement may take the form of bringing two bones closer together, pushing food through the digestive system, or pumping blood through blood vessels. In addition to producing movement, muscles also hold the body erect and generate heat.

Types of Muscles

cardiac muscle

involuntary muscles

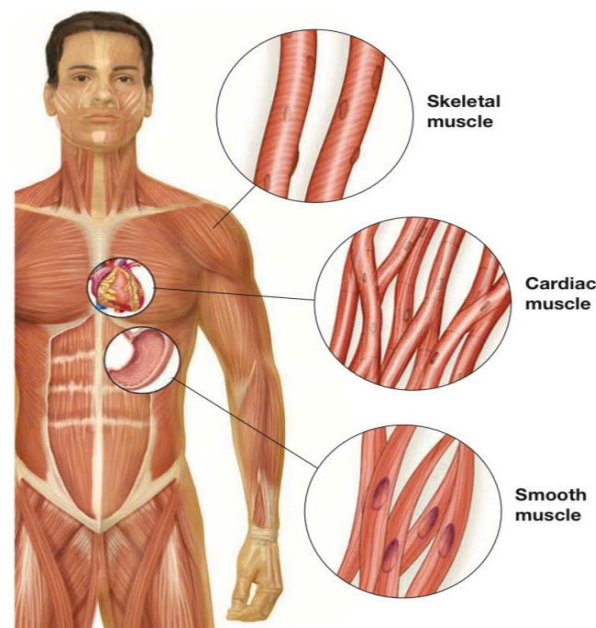
skeletal muscle

smooth muscle

voluntary muscles

The three types of muscle tissue are **skeletal muscle**, **smooth muscle**, and **cardiac muscle** (see Figure 4-23 ■). Muscle tissue may be either voluntary or involuntary. **Voluntary muscles** are those muscles for which a person consciously chooses to contract and for how long and how hard to contract them. The skeletal muscles of the arm and leg are examples of this type of muscle. **Involuntary muscles** are the muscles under the control of the subconscious regions of the brain. The smooth muscles found in internal organs and cardiac muscles are examples of involuntary muscle tissue.

■ **Figure 4-23** The three types of muscles: skeletal, smooth, and cardiac.



On this page, the diagram of the muscular system visually classifies three muscle types (skeletal, cardiac, smooth), while adjacent verbal text defines and differentiates them in terms of control and function. Labels, colour coding, and zoomed-in insets restate the same experiential meanings already expressed verbally. From an intersemiotic perspective, image and text do not divide semiotic labour; instead, they elaborate each other through restatement

and exemplification. Appendix B3 codes this relation as Elaboration: Expository-Classificatory, with high strength recorded in Appendix C. A similar pattern appears in MT-P170.

Diagnostic Procedures (continued)		
Term	Word Parts	Definition
Holter monitor		Portable ECG monitor worn by patient for a period of a few hours to a few days to assess heart and pulse activity as person goes through activities of daily living; used to assess patient who experiences chest pain and unusual heart activity during exercise and normal activities
stress testing		Method for evaluating cardiovascular fitness; patient is placed on treadmill or bicycle and then subjected to steadily increasing levels of work; EKG and oxygen levels are taken while patient exercises; test is stopped if abnormalities occur on EKG; also called <i>exercise test</i> or <i>treadmill test</i>



■ **Figure 5-18** Man undergoing a stress test on a treadmill while physician monitors his condition. (Serafino Mozza/Shutterstock)

Therapeutic Procedures		
Term	Word Parts	Definition
Medical Procedures		
cardiopulmonary resuscitation (CPR) (kar-dee-oh-PULL-mon-air-ee / ree-suss-ih-TAY-shun)	cardi/o = heart pulmon/o = lung -ary = pertaining to	Procedure to restore cardiac output and oxygenated air to lungs for person in cardiac arrest; combination of chest compressions (to push blood out of heart) and artificial respiration (to blow air into lungs) is performed by one or two CPR-trained rescuers
defibrillation (dee-fib-rih-LAY-shun)	de- = without	Procedure that converts serious irregular heartbeats, such as fibrillation, by giving electric shocks to heart using instrument called defibrillator; also called <i>cardioversion</i> ; automated external defibrillators (AEDs) are portable devices that automatically detect life-threatening arrhythmias and deliver appropriate electrical shock; designed to be used by nonmedical personnel and found in public places such as shopping malls and schools



■ **Figure 5-19** An emergency medical technician positions defibrillator paddles on the chest of a supine male patient. (Floyd Jackson/Pearson Education, Inc.)

Here, procedural terms such as stress testing are embedded within a definitional table, accompanied by a photographic image of a patient undergoing the procedure. Importantly, the photograph does not introduce new narrative meaning; it visually specifies what the definition already states. This redundancy is pedagogically productive rather than wasteful: the visual

mode concretises an abstract term, while the text constrains interpretation, ensuring terminological precision.

Across MT ensembles, elaboration serves an epistemic stabilising function. By repeatedly aligning diagram ↔ label ↔ definition, the textbook minimises interpretive drift and constructs medical knowledge as something that can be securely “*fixed*” through multimodal alignment. This finding aligns with international multimodal research showing that scientific textbooks rely heavily on elaborative image-text relations to stabilise abstract knowledge domains (Kress & van Leeuwen, 2006; Bezemer & Kress, 2016).

From a Vietnamese discourse-analytic perspective, Le et al. (2021) similarly notes that disciplinary discourse often privileges semiotic redundancy to consolidate specialised meaning. The MT corpus exemplifies this principle in visual-verbal form: elaboration is not merely supportive, but structurally central to meaning-making.

5.3.2 Selective Elaboration in Career Paths: Medical

In contrast, CPM employs intersemiotic elaboration selectively rather than systematically. Elaboration is activated only at critical conceptual junctures where minimal stabilisation is required to support subsequent procedural or communicative activity. As a result, elaboration in CPM is locally dense but globally constrained. This selective logic is evident in CPM-B3P22.

4

Challenges





caseload



loan



paperwork

The Ailing Medical Profession

Hospitals today are facing economic hardships. While most doctors and nurses are competent, hardworking individuals, they encounter various challenges that can interfere with patient care.

The **demand** for medical professionals has increased in recent years. With rapidly rising populations, communities need good health care now more than ever. At the same time, however, the economic downturn has led to shortfalls in the **budget** of health care providers, so hospitals cannot afford to hire new staff. Consequently, doctors and nurses are paid less to work longer **shifts** and take on larger **caseloads**. The added **paperwork** forces many to work unpaid overtime just to maintain the necessary records. More and more of these workers are suffering from **exhaustion**.

The **shortage** of medical professionals is partly due to individual financial hardships as well. Increases in medical school costs have forced students to take out **loans** and incur substantial **debt** before they can start practicing. And today's doctors must also pay hefty insurance premiums each month in case they are sued for medical **malpractice**. For many would-be doctors, these expenses are too high, and so students are pursuing more affordable career paths.

Get ready!

1 Before you read the passage, talk about these questions.

- How can economic hardships affect hospitals?
- How can economic hardships affect individual medical professionals?

Reading

2 Read the article on challenges in the medical profession. Then, choose the correct answers.

- What is the main idea of the article?
 - A how to improve medical practices
 - B poor hospital management strategies
 - C shortfalls in medical education
 - D difficulties faced by medical workers
- Which is NOT a challenge mentioned in the article?
 - A shortages of medical supplies
 - B rising medical school costs
 - C longer shifts for doctors
 - D increased paperwork
- According to the article, what might prevent a student from pursuing medical school?
 - A substantial loans
 - B increased caseloads
 - C too much paperwork
 - D exhaustion

Vocabulary

3 Write a word that is similar in meaning to the underlined part.

- Dr. Bainbridge was sued for neglectful or illegal action after the failed surgery.
_ _ l p _ _ c t _ _ e
- I have to file this official documentation in the patient's medical record. p _ _ e r _ _ r k
- Our available amount of money decreased this year by 15%. b _ _ g e _
- Henry cannot afford to pay off his amount of money owed. _ _ e b _

The page combines visual hierarchy with explanatory text to clarify patient prioritisation. Image and text elaborate the same experiential meaning: how severity determines order of treatment. However, the elaboration is bounded. The page does not systematise triage as a medical taxonomy; instead, it provides just enough conceptual clarity to enable learners to complete reading comprehension and vocabulary tasks. Appendix D records this as Elaboration: Expository, but with moderate strength, reflecting its preparatory role. A parallel pattern occurs in CPM-B3P10.

4 Read the sentence and choose the correct word.

- 1 Please organize these files in **chronological order** / status.
- 2 Nurse Wilson takes over for me at the **plan of care** / shift change.
- 3 Since the doctor's writing was not **comprehensive** / legible, the pharmacist could not read it.

5 Listen and read the email about staff communications again. What is a danger of poor communication?

Listening

6 Listen to a conversation between two nurses. Choose the correct answers.

- 1 What is wrong with the medical record?
 - A The wrong medication is listed.
 - B The notes are not in chronological order.
 - C The patient's name is not clearly marked.
 - D The correct dosage amount is missing.
- 2 How will the nurses handle the problem?
 - A They will ask the patient about his medication.
 - B They will review the medical record again.
 - C They will find out who wrote the patient's notes.
 - D They will contact a nurse from the previous shift.

7 Listen again and complete the conversation.

Nurse 1: Hey Wendy, is Mr. Williams 1 _____ medication?

Nurse 2: I don't know. What does his 2 _____ say?

Nurse 1: Well, I'm not sure. I 3 _____ these notes.

Nurse 2: Let me see ... Wow, this doesn't make any sense. I think he's supposed to get his meds every day at nine AM, but it doesn't look like 4 _____ dosages for the last three days.

Nurse 1: Really? That 5 _____. How could that happen?

Nurse 2: Wait, I think I see what happened. These notes aren't in 6 _____. Someone entered last week's dosages first, after this week's. It's all mixed up.

Nurse 1: So did he get his medication today or not?

Nurse 2: I'm not sure. Who was on duty at nine AM?

Nurse 1: I'll check the shift board ... Tom was here this morning.

Nurse 2: These notes are 7 _____. We'd better give him a call.

Nurse 1: That's a good idea. We certainly don't want to give Mr. Williams the wrong dosage.

Speaking

8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

I can't read ...
It doesn't look like ...
These notes are ...

Student A: You are a nurse. Talk to Student B about:

- a patient's medication
- a problem with a medical record
- what to do next

Student B: You are a nurse. Talk to Student A about a problem with a medical record.

Writing

9 Use the conversation from Task 8 and the email to write a memo to the hospital director. Include a description of a communication problem and your steps to fix it.



Here, photographs labelled empathy, anxiety, and rapport visually exemplify interpersonal concepts referenced in the text. The images elaborate the meanings of key lexical items, but they do not aim to construct a conceptual system of medical communication. Instead, elaboration functions as conceptual calibration, ensuring shared understanding before learners engage in speaking and role-play activities.

What defines CPM's approach is not the absence of elaboration, but its instrumental containment. Elaboration is used sparingly, targeting threshold concepts that must be aligned before action can occur. This pattern supports findings in international ESP research that vocational materials prioritise enactment over conceptual consolidation, relegating elaboration to a supportive role (Hyland, 2006; Bhatia, 2008).

Intersemiotically, CPM treats images not as anchors of disciplinary knowledge, but as transitional scaffolds that bridge understanding and performance. This explains why enhancement and projection emerge as more salient relations in later sections.

Taken together, the analysis shows that intersemiotic elaboration is epistemically asymmetrical across the corpus. In MT, elaboration functions as a primary mechanism for conceptual stabilisation, tightly binding visual and verbal modes to construct authoritative biomedical knowledge. In CPM, elaboration is deployed selectively to support task progression and professional enactment rather than long-term conceptual consolidation.

This contrast substantiates the claim that intersemiotic relations are not merely formal patterns, but reflections of underlying pedagogic epistemologies. Understanding how elaboration operates differently across the two textbooks provides a crucial foundation for analysing enhancement (Section 5.4) and projection (Section 5.5), where procedural meaning and professional roles come to the fore.

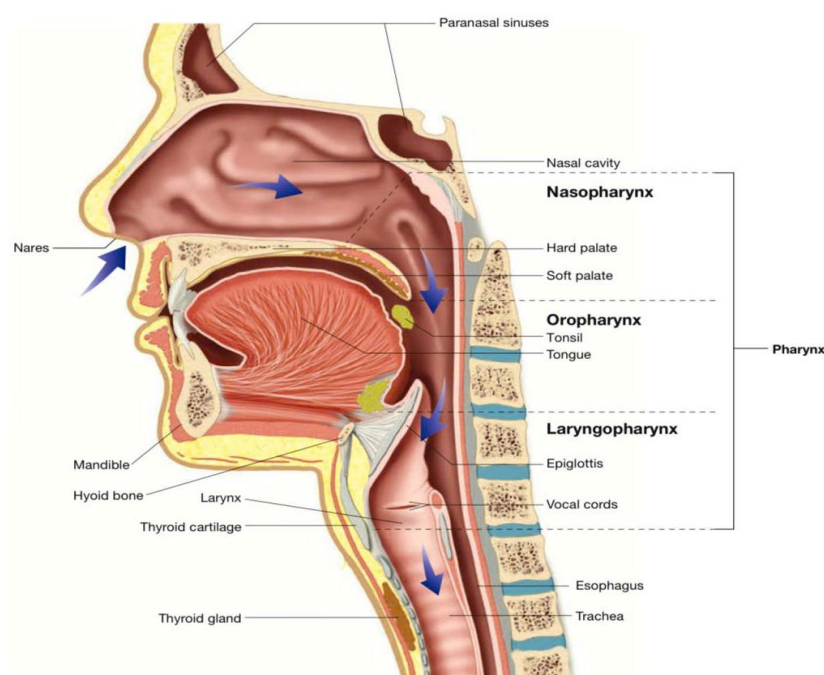
5.4 Intersemiotic Enhancement and Procedural Meaning-Making

This section examines how procedural meaning is constructed through intersemiotic enhancement, where visual resources do not merely restate verbal information but add meanings of sequence, manner, and condition to guide learners' understanding of medical processes. Within the IMAF, enhancement is realised when images expand linguistic meaning by specifying how, when, or under what conditions actions occur (Appendix B3). Across the corpus, enhancement is present in both textbooks; however, it is systematically stronger and more pedagogically central in CPM than in MT.

5.4.1 Procedural enhancement in Medical Terminology: constrained but epistemically functional

Although MT is primarily oriented toward conceptual stabilisation, selected page ensembles demonstrate that procedural meaning is not absent, but is tightly constrained by biomedical explanation.

Respiratory System 231



■ **Figure 7-1** Sagittal section of upper respiratory system illustrating the internal anatomy of the nasal cavity, pharynx, larynx, and trachea.

and filter out large dirt particles before they can enter the lungs. Capillaries in the mucous membranes warm inhaled air as it passes through the airways. Additionally, several **paranasal sinuses**, or air-filled cavities, are located within the facial bones. The sinuses act as an echo chamber during sound production and give resonance to the voice.

Pharynx

adenoids (AD-eh-noydz)

auditory tube

eustachian tube (yoo-STAY-shee-en)

laryngopharynx (lah-ring-goh-FAIR-inks)

lingual tonsils (LING-gwal)

nasopharynx (nay-zoh-FAIR-inks)

oropharynx (or-oh-FAIR-inks)

palatine tonsils (PAL-ah-tyne)

pharyngeal tonsils (fair-IN-jee-al)

Air next enters the pharynx, also referred to as the *throat*, which is used by both the respiratory and digestive systems. At the end of the pharynx, air enters the trachea while food and liquids are shunted into the esophagus.

Word Watch

The term *cilia* means *hair*, and there are other body systems that have cilia or cillialike processes. For example, when discussing the eye, *cilia* means *eyelashes*.

What's In A Name?

Look for these word parts:

audit/o = hearing

lingu/o = tongue

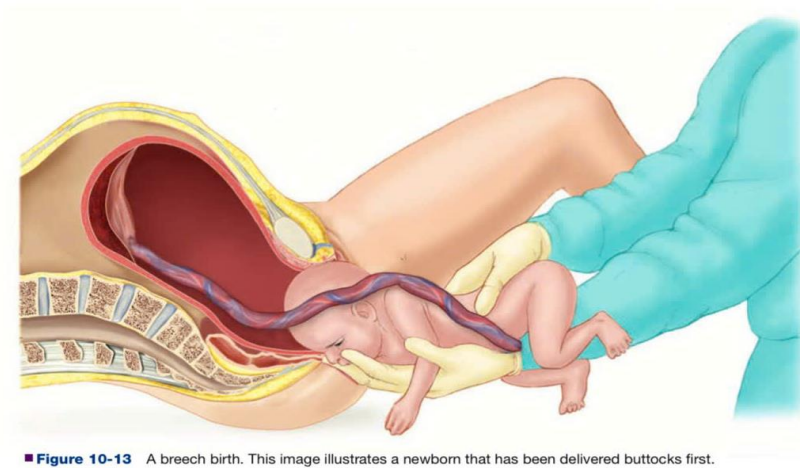
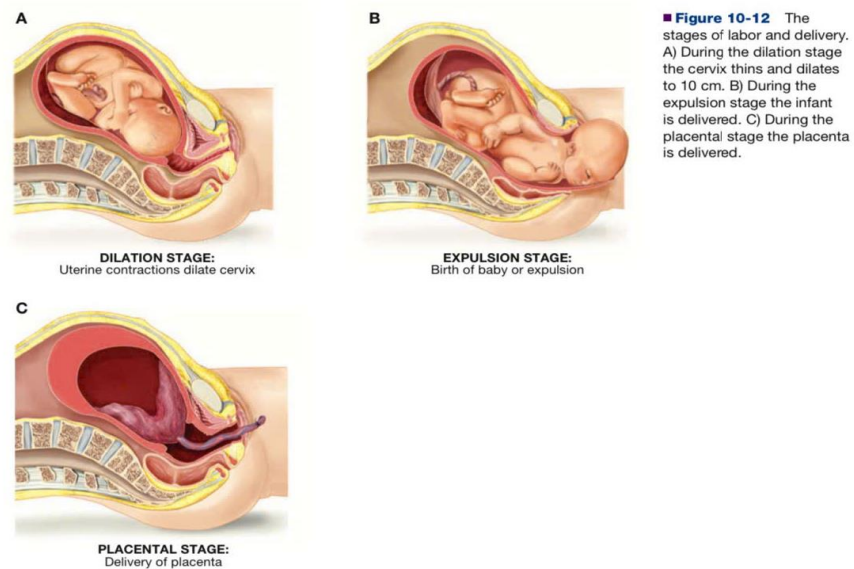
-al = pertaining to

-ory = pertaining to

MT-P231 illustrates enhancement through process sequencing. The verbal text describes physiological steps, while the accompanying diagram visually orders these stages through directional arrows and spatial progression. Here, the image enhances the text by making temporal sequence visually explicit, reducing inferential load for the learner. However, the procedure remains

framed as biological process rather than professional action. The learner is positioned as an observer of internal mechanisms, not as an agent performing steps.

A similar pattern occurs in MT-P355, where childbirth stages are explained linguistically and enhanced visually through a series of anatomical cross-sections.



The images specify manner and progression (e.g., dilation → expulsion → placental stage), but do not invite interaction or decision-making. Procedural meaning is therefore epistemic rather than performative: learners are guided to understand processes, not enact them.

These findings align with international multimodal research showing that science and medical textbooks often use enhancement to clarify naturalised processes, maintaining disciplinary authority while limiting learner agency (Unsworth, 2008; Bezemer & Kress, 2016).

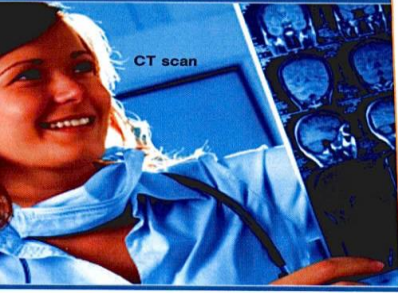
5.4.2 Procedural enhancement in Career Paths: Medical: action-oriented and performative

In CPM, enhancement functions differently. Rather than explaining internal mechanisms, images work intersemiotically with tasks to stage professional action. In CPM-B3P26, verbal instructions outline task steps, while photographs and layout structure enhance meaning by specifying manner and condition.

12 Diagnostic equipment



ultrasound



CT scan

Community Health Clinic Medical Imaging Services

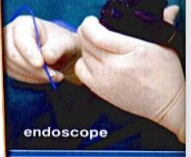
At the Community Health Clinic, we have a wide range of **medical imaging** techniques available. Which technique to use depends on the situation and needs of the patient.

Radiography is the use of **X-rays** to take pictures of the internal body. It is a common imaging technique used in many different situations. A **CT scan** is an advanced form of radiography. It uses many different two-dimensional X-rays to create a three-dimensional image.

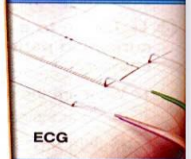
An **ultrasound** uses the reflection of sound waves to create images. A special **gel** placed on the skin aids this process. Unborn children are often examined through ultrasounds.

Recently the use of **endoscopes** during examinations and surgery has become quite common. A doctor **feeds** a tube with a tiny camera on its end directly into the organ or cavity to be examined.

Advanced imaging techniques include **MRIs** and **ECGs**. An MRI utilizes the magnetic properties of atoms to create very detailed pictures. An ECG measures the electrical activity of the heart.



endoscope



ECG



X-ray

Get ready!

1 Before you read the passage, talk about these questions.

- What are some different kinds of medical imaging?
- Why should a health professional be familiar with all kinds of diagnostic equipment?

Reading

2 Read the clinic's website about medical imaging. Then, choose the correct answers.

- What is the website mainly about?

A when to use advanced imaging

B the cost of medical imaging techniques

C different kinds of medical imaging

D how X-rays are being replaced by other imaging techniques
- Which of the following is NOT an advanced medical imaging technique?

A MRI

C CT scan

B X-ray

D ECG
- Who is most likely to be examined by an ultrasound?

A a patient in surgery

B a patient with a broken bone

C a patient who is pregnant

D a patient with abnormal heart activity

Vocabulary

3 Match the words (1-5) with the definitions (A-E).

- endoscope
- CT scan
- radiography
- ultrasound
- MRI

- a form of medical imaging that uses many two-dimensional X-rays to create a three-dimensional image of an object
- a thin, tubular instrument used to examine the inside of an organ or body cavity
- a medical imaging technique that uses magnetic forces on atoms to produce an image of the body
- the use of x-rays to view images of the internal human body
- a medical imaging technique that uses the reflection of sound waves to produce an image of the body

Visuals show how procedures are carried out (posture, tools, interactional orientation), while task sequencing directs learners through a simulated workflow. The image does not duplicate the text; instead, it anchors abstract instructions in visible practice, transforming procedural language into enactable knowledge.

This enhancement becomes more pronounced in that page ensemble, where task-based activities are supported by contextual images depicting workplace interaction. Here, enhancement operates at multiple levels: Sequence: ordered tasks mirror real-time professional routines; Manner: images show appropriate handling, positioning, and interaction; Condition: contextual cues signal when and why actions are required.

Crucially, the learner is positioned not as an external observer but as a prospective participant. Procedural meaning emerges intersemiotically through the alignment of instruction, image, and task design.

This pattern corresponds with ESP and multimodal pedagogy research demonstrating that vocational materials rely on enhancement to simulate professional practice and scaffold procedural competence (Bateman, 2014; Bezemer & Kress, 2016; Halliday & Matthiessen, 2014).

5.4.3 Comparative interpretation

Taken together, the contrast is clear. MT uses intersemiotic enhancement to clarify biomedical processes, reinforcing conceptual understanding within a controlled epistemic frame. CPM, by contrast, mobilises enhancement to enable procedural enactment, integrating visual cues with task-based language to support professional action. This difference explains why procedural meaning, while present in both textbooks, is pedagogically consequential only in CPM.

Importantly, this finding complicates any simplistic division between “*conceptual MT*” and “*procedural CPM*”. MT does construct procedural meaning, but it does so in a descriptive and explanatory mode, whereas CPM operationalises procedure as something to be done. This distinction becomes critical in Chapter V’s synthesis, where intersemiotic patterns are linked to contrasting multimodal pedagogies: epistemic stabilisation versus enactment-oriented learning.

5.5 Projection and the Enactment of Professional Roles

This section explains how professional roles and interactional routines are “*rehearsed*” intersemiotically – that is, how language and image co-operate to project speech/role relations and make “*who is speaking/acting*” pedagogically actionable within the page ensemble. In systemic-functional terms, projection concerns meaning relations that construe locutions (what is said) and ideas (what is thought) as reportable, quotable, and attributable; in multimodal ensembles, this is often realised through scripted dialogue, role labels, and scenario framing that align visual participants with textual voices.

Crucially, projection is not a decorative add-on: it is an epistemic staging device. It reorganises medical English from “*knowledge-to-recognise*” (terminology/definitions) into “*practice-to-perform*” (speaking as a nurse, questioning as a clinician, responding as a colleague). This aligns with multimodal educational research showing that images and layouts can preconfigure participation frameworks – how learners are invited to take up roles and actions rather than only interpret content.

5.5.1 Strong projection in Career Paths: Medical

To illustrate how projection operates most explicitly, the following ensemble is analytically treated as a high-projection page because the page’s central learning work depends on voicing (who says what, to whom, and with what consequence).

8

Taking a history



family medical history



Assist-A-Med

Guide to Taking Medical Histories

As a medical professional, you should be prepared to take thorough medical histories. The **data** you collect reveals important details about a patient. Doctors use these details to uncover likely causes of medical conditions and establish diagnoses.

First, get as many details as possible about a patient's **complaint**. Find out if the patient has a **history of** the condition. Ask when the patient experienced the **onset** of symptoms. Check what makes it worse and what makes it better.

Then find out about the patient's **past medical history**. Record any previous medical problems, especially those that could relate to the **current** condition.

Other details can also contribute to a useful medical history. A good **social history** includes information about the patient's personal habits and lifestyle. A **family medical history** reveals conditions that are more likely to occur in a particular family.

Before any medical **procedure**, make sure you have an updated record of the patient's **allergies**. This way, you avoid exposing patients to medications or substances that might harm them.

Get ready!

- 1 Before you read the passage, talk about these questions.

- 1 What kind of information is included in a patient's medical history?
- 2 How do doctors use medical histories?

Reading

- 2 Read the guide to taking medical histories. Then, complete the table using information from the passage.

Type of History	Information Included
Past Medical History	_____
Social History	_____
Family Medical History	_____

Vocabulary

- 3 Match the words (1-6) with the definitions (A-F).

- | | |
|-------------|----------------------------|
| 1 — data | 4 — complaint |
| 2 — onset | 5 — procedure |
| 3 — allergy | 6 — family medical history |

- A the first instance of something
 B a record of a patient's family's conditions
 C a collection of information
 D a condition that causes a reaction or illness
 E a medical treatment
 F a pain or illness reported by a patient

CPM-B2P18 uses dialogue-as-text plus scenario imagery to project professional interaction, thereby positioning the learner as a role performer (e.g., nurse/colleague) rather than a detached reader.

On this page, the left column stages listening/gap-fill work through a clearly delimited conversation frame (e.g., alternating speaker turns labelled

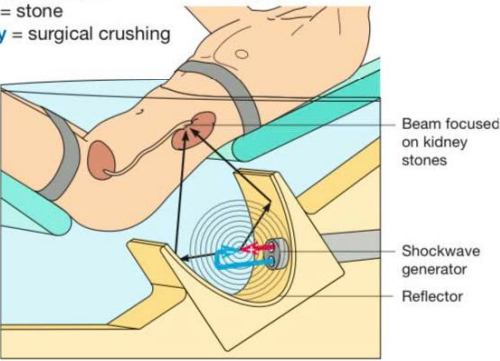
“Nurse 1” / “Nurse 2” with blanks to complete). The right column then upgrades this into overt enactment via “Speaking” tasks that explicitly assign roles (“Student A / Student B”) and provide a functional language bank (e.g., prefabricated phrases learners should use). A contextual photograph at the lower right – showing a caregiving interaction – works as situational anchoring, supplying an interpersonal frame that makes the projected talk socially plausible (care, responsibility, risk).

Importantly, the page does not merely describe workplace communication; it simulates a participation structure: (i) learners first reconstruct projected talk (gap-fill), then (ii) re-animate it as performable speech (role-play), and (iii) evaluate it against workplace stakes (e.g., medication/record accuracy implied by the prompts). In other words, the image does not carry the dialogue content; instead, it stabilises the scene of practice in which the dialogue becomes professionally meaningful. This is consistent with image-text relation models that treat projection as a meaning relation where one mode (verbal dialogue) is licensed and naturalised by the other (visual scenario), creating a coherent enactment space rather than two parallel streams.

The decisive point is attribution: the ensemble repeatedly answers “*who speaks?*” through role labels and turn-taking structure. That is exactly the semiotic work of projection: voices are staged as reportable/performable, not merely as information to comprehend.

5.5.2 *Weaker/constrained projection in Medical Terminology*

The second ensemble is analytically important because it clarifies a boundary condition: MT can evoke professional practice, but it typically does so through conceptual/terminological control rather than through dialogic role projection.

Therapeutic Procedures		
Term	Word Parts	Definition
Medical Treatments		
catheter (KATH-eh-ter)		Flexible tube inserted into body for purpose of moving fluids into or out of body; most commonly refers to tube threaded through the urethra into the bladder to withdraw urine (see again Figure 9-9)
catheterization (cath) (kath-eh-ter-ih-ZAY-shun)		Insertion of tube through the urethra and into the urinary bladder for purpose of withdrawing urine or inserting dye
extracorporeal shockwave lithotripsy (ESWL) (eks-trah-kor-POR-ee-al / shockwave / LITH-oh-trip-see)	extra- = outside of corpor/o = body -eal = pertaining to lith/o = stone -tripsy = surgical crushing	Use of ultrasound waves from outside the body to break up stones; process does not require invasive surgery
■ Figure 9-12 Extracorporeal shockwave lithotripsy, a noninvasive procedure using high-frequency sound waves to shatter kidney stones.		
hemodialysis (HD) (hee-moh-dye-AL-ih-sis)	hem/o = blood	Use of artificial kidney machine that filters the blood of a person to remove waste products; use of this technique in patients who have defective kidneys is lifesaving
■ Figure 9-13 Patient undergoing hemodialysis. Patient's blood passes through hemodialysis machine for cleansing and is then returned to the body. (Gopixa/Shutterstock)		

MT-P330 constructs professional action largely as described procedure (definitions + classificatory organisation), with limited projection: the page presents clinical activity as objectified knowledge rather than as voiced interaction.

The page is organised as a structured instructional table (e.g., “*Diagnostic Procedures*”/“*Therapeutic Procedures*”) with columns for Term, Word Parts, and Definition, accompanied by clinical photographs and captions (e.g., a patient monitored during a stress test; an emergency technician positioning defibrillator paddles). Here, the verbal mode does not allocate speaking turns; instead, it names, decomposes, and defines. Images function as indexical witnesses of practice (someone doing X), but the discourse is not dialogically attributed to a speaker-role (no nurse/doctor/patient voice is staged as quoted talk).

Intersemiotically, the ensemble produces an epistemic stance closer to professional observation: learners are oriented to recognise procedures as terminological objects (“*Holter monitor*”, “*defibrillation*”) and to align them with canonical descriptions. The photos add credibility and situational concreteness, but they do not convert the page into a rehearsal space for interaction. In projection terms, MT-P330 therefore exemplifies constrained projection – the page “*projects*” professional practice mainly as seen and defined rather than as spoken and enacted. This contrast matters because it shows why, pedagogically, CPM more readily supports role-play and interactional competence, while MT more strongly supports literacies of classification and definitional precision (a pattern consistent with multimodal accounts of how disciplinary texts train different “*ways of seeing/doing*”).

5.5.3 Synthesis: *projection as the bridge to pedagogy*

Taken together, the two ensembles show that projection is the most direct semiotic bridge from textbook discourse to pedagogy: In CPM-B2P18, projection is overt and performative: roles are named, turns are structured, and language is offered as reusable professional action. This creates a designed pathway from comprehension to enactment. In MT-P330, projection

is attenuated and observational: professional action is visualised and defined, but not voiced as interaction; role identity remains implicit and secondary to terminological control.

This asymmetry is not a deficiency; it is a genre-consistent design outcome. CPM's projection-heavy ensembles build an enactment-oriented pedagogy because the page itself scripts participation. MT's projection-light ensembles sustain an epistemic pedagogy of disciplined recognition and definitional mastery, where professional authority is modelled through standardised description rather than through simulated talk.

5.6 Intersemiotic Density and Cross-Textbook Synthesis

This section synthesises intersemiotic findings across all 40 page ensembles in order to explain how different configurations of image–text relations cluster into distinct pedagogic logics in the two textbooks. Functionally, it operates as a bridge from Stratum 2 (intersemiotic relations) to Stratum 3 (multimodal pedagogy) by shifting the analysis from page-level mechanisms to corpus-level density and epistemic patterning.

At the level of method, the synthesis draws explicitly on Appendix C, where intersemiotic relations are coded for type (Elaboration, Extension, Enhancement, Projection) and strength (0-3), and on Appendix D, which ensures traceability from aggregate claims back to individual ensembles. Importantly, “*intersemiotic density*” is not treated as a purely quantitative measure; rather, it is interpreted as an epistemic indicator of how tightly visual and verbal modes are coupled to stabilise, extend, or enact medical knowledge – an approach consistent with theory-driven multimodal research (Halliday, 1994; Martinec & Salway, 2005; Bezemer & Kress, 2016).

Across the corpus, a clear cross-textbook asymmetry emerges. MT exhibits high intersemiotic density dominated by Elaboration, whereas CPM

displays more diversified density profiles, with Enhancement and Projection recurring at moderate to high strength. However, this contrast should not be understood as a simple opposition; rather, it reflects different epistemic priorities that are systematically realised through intersemiotic design.

In MT, the consistently high scores recorded in Appendix C are largely driven by redundant but specialised elaboration: diagrams, labels, captions, and definitions repeatedly construe the same medical concept across modes, while adding increasing degrees of terminological precision. This results in dense conceptual coupling, where image and language mutually constrain interpretation and minimise ambiguity. From an epistemic perspective, such density supports what Halliday and Matthiessen (2014) describe as technicalisation – the transformation of experience into stable, nominalised knowledge. International multimodal studies of science and medical textbooks have shown that this kind of elaborative density is a hallmark of disciplinary literacy training, particularly for novice learners who must align visual schemata with specialised lexis (Unsworth, 2008; Bezemer & Kress, 2016). Vietnamese discourse-analytic scholarship similarly notes that educational texts often privilege stabilisation and definitional clarity as a means of legitimising knowledge (Nguyen, 2012). The MT corpus fits squarely within this tradition.

However, high elaboration density does not equate to pedagogic uniformity. Critically, Appendix D shows that even when MT introduces procedural content (e.g. physiological stages, diagnostic processes), the intersemiotic logic remains predominantly elaborative: images clarify what a process is, rather than how or by whom it is enacted. Thus, MT's intersemiotic density consolidates an epistemic pedagogy of recognition and

classification, where knowing medicine means being able to identify, name, and conceptually organise biomedical phenomena.

By contrast, CPM demonstrates lower elaborative density but greater relational diversity. While some CPM ensembles do rely on elaboration to anchor key terminology, Appendix C shows that Enhancement (sequence, manner, condition) and Projection (dialogue, role attribution) recur with higher relative frequency and strength. This redistribution of intersemiotic weight has important epistemic consequences. Enhancement allows images to scaffold processual understanding – for example, by visually sequencing actions that are only partially specified in language while Projection explicitly stages who speaks and acts within professional scenarios. Multimodal research in vocational and ESP contexts has repeatedly emphasised that such relations enable learners to move from conceptual knowledge toward situated practice and role enactment (Kress & van Leeuwen, 2006; Royce, 2007; Bezemer & Kress, 2016).

From a Vietnamese EMP perspective, this pattern is particularly significant. Applied linguistics research in Vietnam has long observed a tension between knowledge-heavy instruction and the need for communicative, practice-oriented competence (Bui & Nguyen, 2016; Lam, 2024). CPM's intersemiotic density profile – less concentrated on elaboration, more distributed across enhancement and projection – can therefore be interpreted as a designed response to this pedagogic gap, using multimodal resources to simulate professional participation within the textbook itself.

Importantly, the synthesis also reveals that intersemiotic density is qualitative, not merely additive. MT's dense elaboration produces strong conceptual coherence but limits opportunities for enactment; CPM's diversified density produces weaker conceptual redundancy but stronger

affordances for procedural and interactional rehearsal. This finding aligns with Martinec and Salway's (2005) argument that different image–text relations realise different semantic work, and that pedagogic meaning emerges from their patterned distribution rather than from any single relation type.

In sum, the cross-textbook synthesis demonstrates that intersemiotic density functions as an epistemic signature. MT's high elaboration density realises a pedagogy of conceptual stabilisation, while CPM's mixed enhancement-projection density realises a pedagogy of professional enactment. This contrast does not imply superiority of one design over the other; rather, it explains how each textbook mobilises multimodal resources to privilege particular forms of medical knowing. By establishing these corpus-level density profiles, Section 5.6 prepares the ground for Stratum 3, where these intersemiotic patterns are reinterpreted as distinct multimodal pedagogies with implications for EMP teaching and learning in the Vietnamese context.

5.7 Emergent Multimodal Pedagogy in Medical English Instruction

Synthesising the intersemiotic patterns identified across all 40 page ensembles, this section argues that two distinct but systematic multimodal pedagogies emerge from the corpus. These pedagogies are not inferred impressionistically; rather, they arise from recurring configurations of intersemiotic relations (Appendix C, D) and can be situated within established multimodal and ESP scholarship, both internationally and in Vietnam.

5.7.1 Epistemic multimodal pedagogy in Medical Terminology

Across the MT corpus, intersemiotic relations are dominated by Elaboration, particularly label-diagram-definition coupling. This pattern supports what can be characterised as an epistemic multimodal pedagogy, in

which medical knowledge is construed as stable, classifiable, and conceptually self-contained.

Internationally, this finding aligns closely with social semiotic accounts of scientific discourse, where diagrammatic representation and definitional language function together to stabilise disciplinary knowledge (Halliday & Martin, 1993; Lemke, 1998; Kress & van Leeuwen, 2006). In such texts, meaning is not advanced through action or interaction, but through conceptual condensation – a process whereby complex phenomena are rendered as inspectable, nameable entities.

Vietnamese multimodal research reports a comparable tendency in textbook-based instruction. Trinh (2010) notes that Vietnamese academic textbooks, particularly in science-related domains, privilege classificatory discourse and definitional closure, positioning learners as epistemic observers rather than participants. Seen in this light, MT's multimodal design is not pedagogically neutral. It cultivates a form of disciplinary literacy in which knowing means recognising canonical representations and aligning them with precise medical terminology. This epistemic orientation corresponds directly to the “*knowledge-consolidation*” axis of the Conceptual Framework (Figure 2.2) and explains why MT is particularly effective for terminological accuracy, yet limited in modelling professional interaction.

5.7.2 Enactment-oriented multimodal pedagogy in Career Paths: Medical

By contrast, CPM exhibits a markedly different intersemiotic profile, characterised by the recurrent use of Enhancement (sequence, condition, manner) and Projection (dialogue-image relations). Together, these relations underpin what can be described as an enactment-oriented multimodal pedagogy.

This orientation is well documented in international ESP and professional communication research. Studies by Hyland (2006) and Bhatia (2008) show that vocational English materials increasingly construct learning as participation in simulated professional practices, where language is learned through task performance rather than conceptual exposition. From a multimodal perspective, Bezemer and Kress (2016) argue that such materials rely on images and layout to script action, while language specifies roles, intentions, and interactional moves.

Vietnamese ESP research echoes this shift. Lam & Pham (2024) observes a growing emphasis on task-based and scenario-driven instruction in professional English courses, particularly in healthcare education. Although not framed explicitly as multimodal analysis, this work highlights the same pedagogic logic identified here: learners are socialised into professional identity through doing rather than defining.

Within CPM, intersemiotic projection is crucial. Images provide situational grounding (who is involved, where, doing what), while dialogue boxes, role labels, and task prompts project learners into enacted professional roles. Consequently, medical English is construed as a mode of action and interaction, not merely a body of knowledge. This corresponds to the “*practice-enactment*” axis of Figure 2.2 and explains CPM’s strong pedagogic affordances for communicative competence.

5.7.3 Integrative implications

Critically, the study does not argue for the superiority of one pedagogy over the other. Instead, the corpus demonstrates that different intersemiotic architectures enable different kinds of learning. MT supports epistemic control and conceptual precision; CPM supports procedural fluency and professional enactment.

This duality resonates with both international multimodal theory (Jewitt, 2014; O'Halloran, 2008) and Vietnamese discourse-analytic scholarship (Nguyen & Tran, 2023), which emphasise that pedagogic meaning emerges from semiotic design choices rather than from language alone. Accordingly, Chapter VI will argue that effective Medical English instruction in Vietnam requires a principled integration of these two multimodal pedagogies, rather than reliance on either terminology-heavy or communication-only approaches.

5.7.4 Implications for Vietnamese EMP classroom practice

The intersemiotic patterns identified in this chapter suggest that Vietnamese EMP classrooms will benefit most from teaching students to read and use page ensembles as multimodal systems, rather than treating images as optional “illustrations” or treating language as the sole carrier of medical knowledge. This implication aligns with social semiotic research showing that disciplinary understanding is produced through designed relations between modes (Kress & van Leeuwen, 2021; Bezemer & Kress, 2016) and with intersemiotic frameworks emphasising structured logico-semantic relations between image and text (Martinec & Salway, 2005; Royce, 2007). It also directly addresses a recurring concern in Vietnamese EMP/EAP contexts: terminology-heavy instruction often stabilises vocabulary but under-prepares students for procedural reasoning and interactional competence (Nguyen, 2000; Le, 2017).

First, MT-type materials (elaboration-dense) should be taught through explicit diagram-text literacy routines. Because MT frequently couples diagram/label/definition via strong elaboration, effective teaching should foreground how meaning is stabilised: (i) trace labels to visual parts, (ii) map each label to definitional clauses, and (iii) articulate the

classificatory/analytical relation being constructed (e.g., type-of, part-of). This approach treats elaboration as a learnable strategy rather than passive redundancy, supporting disciplinary literacy as conceptual control (Unsworth, 2008; Bezemer & Kress, 2016). In Vietnamese classrooms where definitional learning is common, the key shift is methodological: teachers should move from “*translate-and-memorise*” to “*read the ensemble*,” thereby making students’ biomedical understanding more robust and less dependent on rote vocabulary lists - consistent with discourse-analytic calls in Vietnam for controlled, text-based interpretation rather than impressionistic teaching (Nguyen, 2008).

Second, CPM-type materials (enhancement/projection-heavy) should be leveraged to teach procedural and interactional competence through multimodal rehearsal. Where enhancement specifies sequence, manner, and condition, teachers can operationalise images as procedural scaffolds: students describe steps from images, align them with imperative/task language, and justify why an action is appropriate under a given condition. Where projection stages roles and dialogue, teachers can structure role-play that explicitly links visual participation frameworks (who is doing/speaking) to linguistic choices (requests, reassurance, clarification), reflecting ESP research on vocational materials as simulated professional practice (Hyland, 2006; Bhatia, 2008) and multimodal accounts of apprenticeship-like learning through scenario design (Bezemer & Kress, 2016). This is particularly relevant to Vietnamese EMP courses seeking to strengthen communication skills alongside terminology (Le, 2017).

Third, curriculum design should integrate the two regimes rather than selecting one. The corpus indicates a principled pathway for Vietnamese EMP sequencing: use MT-style elaboration to build conceptual stability and

terminological precision, then use CPM-style enhancement/projection to convert that stability into enactable professional practice. Such integration avoids a false dichotomy between “*knowledge*” and “*communication*” by recognising that different intersemiotic relations make different kinds of medical knowing available (Martinec & Salway, 2005; Royce, 2007). In practical terms, this means designing lessons that deliberately shift students from naming and classifying (elaboration) to explaining and doing (enhancement) and finally to speaking as a role (projection), thereby aligning classroom activity with the multimodal pedagogies emergent from the textbooks themselves.

In Vietnamese EMP classrooms, the pedagogical implications of this study may be translated into guided multimodal learning activities that encourage students to examine how visual and linguistic resources interact within textbook page ensembles. Through such activities, students may be guided to analyse the representational, interactive, and compositional meanings of visual images and to explore how these meanings are integrated with linguistic resources to communicate medical concepts and disciplinary knowledge. By engaging students in this process of multimodal interpretation, EMP instruction may support the development of multimodal disciplinary literacy and enhance learners' ability to interpret specialised medical discourse across multiple semiotic modes.

Overall, the implication for Vietnamese EMP is clear. Improving learning outcomes is not only a matter of selecting “*better*” textbooks, but of teaching the intersemiotic work that textbooks presuppose so that students can interpret, operationalise, and perform medical meanings across modes in ways consistent with contemporary multimodal and ESP research.

5.8 Concluding Remarks

This chapter has addressed the study's intersemiotic focus by demonstrating that medical meaning in EMP textbooks is produced through systematic image-text relations rather than through isolated modes. Across 40 page ensembles, the distributional mapping confirms a stable cross-textbook contrast: *Medical Terminology* is dominated by intersemiotic elaboration, while *Career Paths: Medical* shows stronger and more frequent enhancement and projection, with all relation types occurring but patterned differently at textbook level.

At an epistemic level, these relations generate distinct forms of medical knowing. In MT, elaboration tightens diagram-label-definition coupling to stabilise conceptual categories and terminological precision, constructing knowledge as canonical, classifiable, and retrievable. In CPM, enhancement adds procedural meanings of sequence, manner, and condition, while projection stages dialogue and role attribution, converting medical English from “*knowledge-to-recognise*” into “*practice-to-perform*”. As a result, MT primarily supports conceptual consolidation, whereas CPM more readily scaffolds procedural reasoning and professional enactment.

Synthesising these patterns through intersemiotic density further shows that multimodal pedagogy is an emergent property of design. MT's elaboration-heavy architecture produces an epistemic pedagogy of disciplined recognition and definitional mastery; CPM's diversified enhancement - projection architecture produces an enactment-oriented pedagogy grounded in workplace simulation and communicative rehearsal. Importantly, the chapter does not treat either regime as inherently superior. Instead, it clarifies how different intersemiotic configurations enable different pedagogic affordances and constrain what can be made learnable within EMP classrooms.

Overall, Chapter V provides the necessary analytic bridge between Chapter IV's visual meaning profiles and the thesis' pedagogical claims: it explains how visual and verbal resources are distributed, coupled, and mobilised to construct multimodal medical knowledge. These findings directly inform the synthesis of multimodal pedagogy and the broader implications for Vietnamese EMP instruction developed at the end of this chapter and extended in Chapter VI.

CHAPTER VI: CONCLUSIONS

This chapter synthesises and consolidates the thesis's major insights, integrating empirical findings, theoretical contributions, and Vietnam-relevant implications derived from the comparative multimodal analysis of two EMP textbook series. Central to the thesis is the argument that medical meaning and pedagogic orientation are not simply contained in textbook content but are semiotically designed through patterned visual and intersemiotic orchestration across page ensembles. Accordingly, this chapter restates the research purpose and methodological stance, provides a consolidated answer to the research questions, articulates the study's theoretical and pedagogical contributions, acknowledges controlled limitations, and outlines a coherent agenda for future research extending from designed semiotic affordances to classroom enactment, assessment, localisation, and curriculum development.

6.1 Restatement of Research Purpose and Methodological Orientation

This thesis set out to explain how medical meaning and pedagogic orientations are constructed through the visual and intersemiotic design of EMP textbooks used in Vietnam. Rather than treating images as decorative supplements to verbal explanation, the study has consistently approached the textbook as a designed multimodal artefact, where meaning emerges from patterned orchestration across modes (Kress, 2010; Jewitt, 2014; Cresswell, 2009). In this view, medical knowledge in textbooks is not simply “*presented*”; it is semiotically organised – made visible, stabilised, contextualised, and rendered pedagogically actionable through specific visual-verbal configurations (Kress & van Leeuwen, 2021; Bezemer & Kress, 2016).

Within that overarching aim, the research questions operate as a single explanatory logic, not a set of unrelated prompts. RQ1 frames the broad explanatory task: how visual and linguistic modes interact to construct

medical meaning and disciplinary knowledge in EMP textbooks used in Vietnam. RQ1.1 then specifies the first analytic move by examining how images are employed and organised representationally, interactively, and compositionally (Kress & van Leeuwen, 2021). Building on this, RQ1.2 advances to the second analytic move by investigating how visual and linguistic modes jointly realise meaning through intersemiotic relations and metafunctional alignment (Unsworth, 2006, 2008; Martinec & Salway, 2005; Royce, 2007). Importantly, this RQ structure mirrors the thesis' empirical argument: visual meaning establishes what is epistemically available to learners, while intersemiotic design determines how that availability is pedagogically shaped within the page.

Methodologically, the thesis is grounded in an interpretive-explanatory qualitative design consistent with multimodal social semiotics, where meaning is treated as socially produced through semiotic choice rather than as a transparent reflection of reality (Halliday, 1978; van Leeuwen, 2005; Hodge & Kress, 1988). However, interpretivism here does not imply impressionistic reading. The analysis is theory-driven and operationalised through the IMAF, structured into three interdependent strata: (i) visual semiotic analysis based on the GVD; (ii) intersemiotic relations and multimodal metafunctional alignment; and (iii) pedagogicdisciplinary interpretation of emergent multimodal pedagogy (Kress & van Leeuwen, 2021; Unsworth, 2008). Moreover, this framework is applied comparatively across two textbook cases in order to explain how different multimodal architectures systematically privilege different epistemic priorities in EMP learning.

A key methodological commitment of the study is the use of the page ensemble as the unit of analysis. This choice is not merely technical; it is epistemic. Intersemiotic meanings do not become analytically visible when

images or clauses are isolated, because textbook knowledge is designed through the co-presence of diagrams/photos, labels, running text, task prompts, and layout pathways within a single ensemble (Bateman, 2014; Bateman, Wildfeuer, & Hiippala, 2017). By analysing ensembles, the thesis can account for how semiotic labour is distributed across modes and how learners are guided through stabilisation (e.g., diagram-label-definition coupling) or enactment (e.g., image-supported task sequences and role-based dialogue).

Critically, the thesis does not claim to measure learning outcomes or classroom effectiveness. Instead, it examines the semiotic affordances and constraints produced by textbook design – what kinds of disciplinary literacy, professional reasoning, and role positioning are made available by the multimodal architecture of the materials (Airey, 2016; Unsworth, 2006, 2008). This distinction is central: outcomes depend on teaching practice, institutional conditions, and learner histories, whereas this thesis contributes by explaining how textbooks preconfigure pedagogic possibilities through design. Importantly, this orientation also speaks directly to the Vietnamese context where EMP materials are often imported or adapted, and where multimodal literacy has been increasingly foregrounded as a strategic educational priority (MOET, 2018; Hoang, 2015; Lam & Pham, 2024). In this sense, the thesis positions multimodal analysis not as a descriptive add-on, but as an essential explanatory lens for understanding how EMP textbooks mediate biomedical knowledge and professional practice in Vietnamese medical education.

6.2 Summary of Key Findings in Relation to Research Questions

This section consolidates the thesis findings by answering the research questions directly and comparatively. Across both textbooks, the analysis confirms a central claim of multimodal social semiotics: meaning in educational discourse is not carried by language or image alone but emerges from systematic semiotic design, where modes are coordinated to stabilise knowledge, guide attention, and position learners in particular ways (Kress, 2010; Jewitt, 2014; van Leeuwen, 2005). However, the two textbooks do not realise this multimodal principle in the same way. Instead, they instantiate two contrasting but internally coherent semiotic regimes, which in turn underpin two distinct pedagogic orientations in EMP learning.

6.2.1 Key Findings for RQ1: Visual Meaning in EMP Textbooks

The findings for RQ1 establish that the two textbooks operate through two systematically different visual semiotic regimes. In MT, visual meaning is predominantly conceptual, diagrammatic, and classificatory, construing medicine as a system of stable entities and relations. In CPM, visual meaning is predominantly contextualised, narrative, and interaction-oriented, construing medicine as situated practice, routine action, and interpersonal engagement. This contrast is not superficial style difference; it reflects different epistemic assumptions about what counts as “*medical knowing*” for learners and how such knowing should be accessed (Kress & van Leeuwen, 2021; Goodwin, 1994).

Representationally, MT prioritises “*what is made visible*” through conceptual structures – especially part-whole and classificatory relations – so that anatomical and terminological knowledge becomes inspectable, nameable, and retrievable. Such images function as epistemic scaffolds for

disciplinary categorisation, aligning with accounts of scientific discourse where abstraction and systematic representation enable technical knowledge to be stabilised (Halliday, 1998; Lemke, 1998). CPM, by contrast, makes visible not primarily systems but events and scenes: bodies in contexts, practitioners in action, tools in use, and patient-provider relations. In this narrative regime, medical meaning is rendered as processual and socially organised, resonating with the notion of professional vision, where what counts as relevant knowledge is shaped through attention to situated cues and routine practices (Goodwin, 1994). Importantly, the thesis does not treat MT as “*better*” for knowledge nor CPM as “*better*” for practice; rather, it shows how each representational regime selectively foregrounds particular dimensions of the medical domain while backgrounding others.

Interactively, the textbooks diverge in how learners are positioned. MT’s diagram-centred design tends to construct learners as epistemic observers, with “*offer*” images and neutralised viewpoint conventions that support inspection rather than interpersonal engagement (Kress & van Leeuwen, 2021). CPM’s photographic and scenario-based imagery, while often still indirect in gaze, increases social presence via human participants, workplace settings, and interactional cues, thereby positioning learners as professionally oriented interpreters of practice and communication. This interactive orientation coheres with ESP perspectives that view vocational learning as induction into discourse communities and routine genres of professional action (Belcher, 2009; Dudley-Evans & St John, 1998). Within Vietnam, the findings connect to local discussions that EMP teaching has often remained vocabulary – and text-heavy, creating a gap between terminological competence and communicative readiness; CPM’s interactive visual regime is therefore pedagogically consequential because it normalises communication

as part of “*medical meaning*” rather than an optional add-on (Nguyen & Tran, 2023, Lam & Pham, 2024).

Compositionally, MT and CPM organise knowledge through different logics of salience and reading path. MT typically constructs a hierarchy where diagrams and labels occupy high salience and stable locations, with text and practice often subordinated as secondary supports – an arrangement consistent with the role of diagrammatic display in producing epistemic authority and conceptual closure in science education (Bateman, 2014; Kress & van Leeuwen, 2021). CPM, conversely, redistributes salience toward task progression, headings, activity sequencing, and response spaces, so that the page reads as a staged workflow rather than a single conceptual tableau. Moreover, this compositional regime aligns with ESP coursebook traditions where learning is structured as guided performance across tasks rather than accumulation of definitional knowledge (Hyland, 2006; Bhatia, 2008). Vietnamese textbook scholarship similarly emphasises that visual layout is not neutral: it guides what counts as “main knowledge” and what is treated as practice or application, shaping classroom enactment (Hoang, 2015; Unsworth & Ngo, 2014). The thesis extends these insights to EMP by showing that compositional design is a primary mechanism through which medical knowledge is prioritised and pedagogically staged.

6.2.2 Key Findings for RQ2: Intersemiotic Meaning

Findings for RQ2 show that the two textbooks carry distinct intersemiotic signatures – recurring ways in which image and language are coupled – and that these signatures systematically shape what kinds of medical meaning become learnable. In MT, intersemiotic meaning is characterised by high-density elaboration, where diagram-label-definition configurations repeatedly

restate and specify the same experiential content across modes. This produces conceptual stabilisation: learners are guided to align technical terms with canonical visual schemata, reducing interpretive ambiguity and consolidating disciplinary categorisation (Royce, 2007; Unsworth, 2008; Martinec & Salway, 2005). This mechanism resonates with SFL accounts of technicalisation, where knowledge becomes stabilised through regrammaticised relations and controlled meaning (Halliday & Matthiessen, 2014; Halliday, 1998).

CPM, by contrast, is intersemiotically marked by enhancement and projection. Enhancement makes procedural meanings actionable by adding sequence, manner, and conditional cues through images and layout that verbal text alone does not fully specify. Projection stages voice, dialogue, and role allocation so that learners can rehearse professional interaction as part of the page's designed meaning. These signatures align with intersemiotic complementarity research showing that educational texts frequently allocate different "*semiotic labour*" to different modes – one mode anchors scene and action, another specifies role, stance, and linguistic performance (Royce, 2007; Unsworth, 2006, 2008). Moreover, CPM's projection-heavy ensembles connect strongly to ESP genre and professional discourse accounts where learning involves simulated participation in workplace communicative events (Bhatia, 1993, 2017; Hyland, 2006).

Critically, across both textbooks, the thesis demonstrates that multimodal meaning does not arise because images "*illustrate*" language. Instead, meaning emerges from a designed division of semiotic labour, where modes are strategically coordinated to produce different kinds of medical knowing – classification and definitional control in MT; procedure and interactional rehearsal in CPM (Kress, 2010; Bezemer & Kress, 2016). This insight is

consistent with international multimodal scholarship on textbooks and multiple representations, which cautions that the learning value of images depends on how representations are coordinated rather than merely co-present (Ainsworth, 2006; Mayer, 2014). In Vietnam, the point is equally important: studies on multimodal literacy in local classrooms underline that images become pedagogically consequential only when teachers and materials activate their relation with text, rather than treating visuals as peripheral decoration (Unsworth & Ngo, 2014; Nguyen, 2000; Nguyen & Tran, 2023).

6.2.3 Integrated Answer to the overarching research question

Taken together, the findings provide an integrated answer to the overarching research question: visual and linguistic modes interact to construct medical meaning by realising two contrasting epistemic-pedagogic orientations, mapped explicitly onto the study's Conceptual Framework (Fig. 2.2) and operationalised through the IMAF's stratal logic (Fig. 3.2). MT primarily realises an epistemic, knowledge-stabilising pedagogy, where medical meaning is organised as canonical biomedical knowledge: visible through conceptual diagrams, secured through elaborative image-text coupling, and reinforced through compositional hierarchies that privilege system over situation (Kress & van Leeuwen, 2021; Lemke, 1998; Halliday & Matthiessen, 2014). CPM, in contrast, realises an enactment-oriented, practice-based pedagogy, where medical meaning is organised as situated professional action and interaction: made visible through narrative contextualisation, operationalised through enhancement, and rendered performable through projection and task sequencing (Bezemer & Kress, 2016; Royce, 2007; Bhatia, 2008).

Importantly, this integrated conclusion should not be read as a binary evaluation. Rather, it clarifies the thesis' core epistemic argument: different multimodal architectures make available different forms of disciplinary literacy. MT strengthens learners' access to the medical domain through classificatory recognition and terminological precision, while CPM strengthens access through procedural reasoning and communicative role performance. This duality resonates with disciplinary literacy accounts that treat professional competence as a coordinated repertoire of ways of seeing, naming, reasoning, and acting (Airey, 2016; Goodwin, 1994). It also speaks directly to the Vietnamese EMP context, where both knowledge-heavy traditions and newer communicative imperatives coexist: the findings suggest that curriculum effectiveness depends not on choosing one orientation but on integrating them in principled ways (MOET, 2008; Le, 2017; Lam & Pham, 2024). In this sense, Chapter VI proceeds from a grounded empirical conclusion: EMP textbook meaning is not merely in content, but in design – and design is where epistemology becomes pedagogy.

6.3 Theoretical Contributions

This thesis contributes theoretically by showing that multimodal analysis in EMP can move beyond “*applying a framework*” toward explaining how disciplinary knowledge is semiotically designed. Situated within multimodal social semiotics, the study treats textbooks as epistemic artefacts: they do not merely carry medical content but recontextualise biomedical and professional meanings through patterned visual-verbal orchestration (Kress, 2010; van Leeuwen, 2005). Importantly, the thesis demonstrates that EMP textbook meaning-making is not adequately captured through linguistic genre analysis alone, nor through visual description in isolation. Instead, it requires an integrated account of how modes co-function to stabilise concepts, scaffold

procedures, and stage professional participation – an argument consistent with foundational multimodality scholarship (Jewitt, 2014; Bateman, 2014; Bezemer & Kress, 2016), yet still under-realised in EMP research and particularly under-developed in Vietnam-based EMP studies (Lam & Pham, 2024; Ton, 2020).

6.3.1 Advancing Multimodal Theory in EMP Research

The first theoretical contribution is the thesis’ extension of multimodal theorisation into Medical English textbooks as a disciplinary-professional domain. While multimodal discourse analysis has been widely applied to general ELT materials and school textbooks, EMP remains comparatively less examined through systematic, corpus-based multimodal frameworks, despite the field’s high dependence on visual-verbal synthesis (e.g., diagrams, procedural images, tables, labelled schemata). By locating EMP textbook analysis within multimodal social semiotics, the study strengthens the argument that multimodal literacy is not an optional enrichment in medical education but a foundational competence for interpreting biomedical communication (Jewitt, 2014; Kress, 2010; Mukundan & Nimehchisalem, 2012).

Moreover, the study shows that “*medical meaning*” in EMP is not monolithic. It is semiotically constructed through different epistemic pathways – conceptual classification versus situated enactment – and these pathways are materially realised through different multimodal regimes. This finding links multimodal theory to disciplinary literacy, where the focus is not only on decoding texts but on acquiring disciplinary ways of seeing, reasoning, and legitimising knowledge (Airey, 2016; Goodwin, 1994). In EMP terms, this means the visual organisation of anatomical systems and the

staging of professional interaction should be theorised as distinct but complementary forms of medical knowing, rather than treated as different “*types of illustration*”

In the Vietnamese context, this theoretical move matters because local EMP discussions have repeatedly highlighted a structural tension between knowledge-heavy instruction and communicative/professional competence demands (Nguyen & Tran, 2023; Nguyen, 2000, Tran, 2024). However, those debates often remain at the level of pedagogy or curriculum. The present thesis contributes a semiotic explanation: it shows how the tension is built into the multimodal architecture of imported coursebooks, and therefore how “*knowledge vs. enactment*” is not only a teaching-method issue but also a design-property of materials. This positions multimodality as a necessary theoretical lens for EMP research in Vietnam, rather than an optional analytical add-on (Unsworth & Ngo, 2014; Trinh, 2010; Lam & Pham, 2024).

6.3.2 Operationalising IMAF as an Integrated Explanatory Framework

The second theoretical contribution lies in demonstrating the explanatory value of the IMAF as a workable integration of major strands in multimodal theory. The thesis does not claim to invent a new theory; instead, it contributes by showing how an integrated framework can be operationalised rigorously across a full EMP corpus to generate defensible explanations, not impressionistic commentary. In doing so, it responds directly to a known methodological problem in multimodal research: analyses can become either (i) richly descriptive but theoretically fragmented, or (ii) theoretically ambitious but operationally thin. The IMAF addresses this by disciplining analysis through a stratal logic that makes interpretation traceable from

semiotic description to pedagogic inference (Kress & van Leeuwen, 2021; Unsworth, 2008; Jewitt, 2014).

Concretely, the study demonstrates the IMAF's effectiveness across three linked explanatory layers. First, at the level of visual meaning, the GVD provides a systematic metalanguage for representational, interactive, and compositional structures, enabling the thesis to explain not merely what images show but how they make particular biomedical realities "*seeable*" and authoritative (Kress & van Leeuwen, 2006, 2021). Second, at the level of intersemiotic relations, the thesis operationalises image-text relations as structured meaning relations rather than vague notions of "*support*" drawing on established intersemiotic frameworks (Martinec & Salway, 2005; Royce, 2007; Unsworth, 2006, 2008). Third, at the level of pedagogic interpretation, the framework enables the thesis to connect multimodal configurations to disciplinary literacy and professional reasoning – without collapsing into outcome claims – aligning with social semiotic principles that texts offer semiotic affordances rather than measurable learning effects (Bezemer & Kress, 2016; van Leeuwen, 2005).

Critically, the theoretical contribution here is not simply integration as a tidy structure, but integration as an epistemic control mechanism. By defining the page ensemble as the unit of analysis, the IMAF operationalises the social semiotic claim that meaning emerges from orchestration across modes and across layout, not from isolated features (Bateman, 2014; Bezemer & Kress, 2016). This methodological-theoretical alignment is particularly important for EMP, where diagrams, labels, tables, tasks, and captions form tightly coupled ensembles that cannot be explained through a single-mode lens.

6.3.3 *Re-conceptualising Intersemiotic Density as an Epistemic Indicator*

The third theoretical contribution is the thesis' reconceptualisation of intersemiotic density/signature as an epistemic indicator of pedagogy rather than a descriptive metric. Intersemiotic research often identifies relations such as elaboration, extension, and enhancement; however, the present study advances the argument that what matters most in textbook discourse is not merely the presence of relation types but their patterned density and distribution across a textbook ecology – because those patterns function as semiotic evidence of what the material treats as “*legitimate knowing*”

In this thesis, MT's high-density elaboration (diagram-label-definition coupling) is theorised not as redundancy but as a mechanism of conceptual stabilisation: medical knowledge is produced as bounded, classifiable, and terminologically controlled. This aligns with social semiotic and SFL accounts of disciplinary discourse that highlight how scientific knowledge is stabilised through condensation, nominalisation, and controlled relations – here realised multimodally through repeated visual-verbal alignment (Halliday, 1998; Halliday & Matthiessen, 2014; Lemke, 1998). CPM's contrasting intersemiotic signature – where enhancement and projection are more prominent – functions as a mechanism of procedural enactment and role rehearsal, consistent with intersemiotic complementarity approaches that treat modes as dividing labour to realise different dimensions of meaning and participation (Royce, 2007; Unsworth, 2008).

Moreover, positioning density as epistemic indicator strengthens multimodal theory in a way that is analytically productive for ESP. It provides a principled bridge between micro-level semiotic analysis and macro-level

pedagogic explanation, which is precisely the move required if multimodal studies are to inform discourse theory rather than remain at the level of case illustration (Bateman et al., 2017; O'Halloran, 2008). Importantly, the thesis also resists treating density as “*more is better*” Instead, it theorises density as orientation: dense elaboration tends to privilege recognition and classification literacies, while diversified density (enhancement/projection) tends to privilege enactment and participation literacies. This provides a semiotic basis for understanding why different EMP textbooks feel pedagogically different even when they share similar lexical targets.

Finally, this reconceptualisation connects meaningfully to Vietnam-based scholarship on multimodal literacy and materials use. Vietnamese studies have argued that images in textbooks are often under-utilised pedagogically because they are treated as decoration rather than as meaning-bearing resources; the notion of intersemiotic density offers a theoretical tool for identifying where meaning is most tightly designed and therefore where pedagogic attention is most warranted (Unsworth & Ngo, 2014; Hoang, 2015; Le et al., 2021). In other words, density functions as a theoretically grounded diagnostic: it helps explain where the textbook is doing heavy semiotic work and what kind of disciplinary knowing that work is designed to produce.

In sum, the thesis' theoretical contributions lie in (i) extending multimodal theory into EMP as a medical disciplinary domain, (ii) demonstrating IMAF as an integrated explanatory apparatus capable of moving from description to interpretation with traceable logic, and (iii) advancing “*intersemiotic density/signature*” as an epistemic indicator that links multimodal design to pedagogic orientation. Together, these contributions enable the thesis to “*stand*” theoretically: the study not only uses multimodal theory, but also

refines how it can be mobilised to explain disciplinary knowledge-making in EMP materials.

6.4 Pedagogical Contributions

Pedagogically, the thesis contributes at the level of EMP teaching principles and learning objects, rather than lesson planning. It argues that what medical students need to learn from EMP textbooks is not reducible to “*medical vocabulary plus reading*” Instead, learners are apprenticed into disciplinary ways of meaning-making that are inherently multimodal: they must learn to read diagrams as conceptual systems, to interpret procedures as sequenced action, and to recognise how professional roles are semiotically staged and made speakable. This aligns with the broader claim in multimodal social semiotics that pedagogy is always mediated by semiotic design - what a text makes learnable depends on how modes are orchestrated (Kress, 2010; Jewitt, 2014; Bezemer & Kress, 2016). Importantly, the thesis does not infer learning outcomes; it identifies semiotic affordances that enable (or constrain) particular forms of EMP learning, and it clarifies what those affordances demand from pedagogy if they are to be realised in Vietnamese medical classrooms (van Leeuwen, 2005; Hoang, 2015; Le, 2017).

6.4.1 Multimodal Disciplinary Literacy in Medical English

A central pedagogical contribution is the thesis’ reframing of EMP learning as multimodal disciplinary literacy. In disciplinary literacy terms, to “know” a field is to access its characteristic ways of representing, reasoning, and validating knowledge (Airey, 2016). In medicine, this includes the ability to interpret multiple representations and to coordinate them under professional constraints – precisely because medical knowledge is routinely externalised through images, charts, and procedural formats alongside

technical language. Therefore, medical English pedagogy cannot be confined to lexical decoding; it must cultivate learners' competence in reading medical semiotic systems.

However, this is not a generic claim about “*visuals help learning*” Rather, the thesis specifies three discipline-relevant literacy demands that the analysed textbooks systematically foreground:

Reading diagrammatic knowledge as conceptual structure. In MT-type ensembles, labelled schemata and classificatory diagrams function as epistemic technologies that stabilise part–whole relations and taxonomies. Pedagogically, this means students must learn how diagrams mean – how salience, framing, and labelling constrain interpretation and legitimise a canonical biomedical viewpoint (Kress & van Leeuwen, 2021; Lemke, 1998). This strengthens a form of literacy that is closer to conceptual mapping and definitional control than to conversational fluency (Halliday, 1998; Halliday & Matthiessen, 2014).

Understanding procedures as staged, conditional action. In CPM-type ensembles, procedural meaning is frequently distributed across task instructions, contextual photographs, and stepwise layouts. Students therefore need literacy in procedural semiosis: recognising sequencing, manner, and condition across modes, not only following text instructions. This resonates with research on learning with multiple representations, which emphasises that understanding emerges when learners coordinate representations rather than treat them as duplicates (Ainsworth, 2006; Mayer, 2014).

Positioning professional roles as semiotically enacted. The thesis shows that role relations are not only taught by dialogue lines; they are projected through the alignment of scenario images, role labels, and interactional

prompts. In other words, professional identity becomes learnable when the page ensemble offers a participation framework that specifies “*who speaks/acts*” “*to whom*” and “*with what stance*” This connects to the notion of professional vision – how novices learn to see, interpret, and act within a professional field (Goodwin, 1994) – and to multimodal accounts of how texts script participation and social relations (Jewitt, 2014; Royce, 2007).

Moreover, this reframing is particularly consequential in Vietnam. Vietnamese EMP research has repeatedly noted persistent gaps between terminology-heavy instruction and interactional/procedural competence in clinical contexts (Nguyen, 2000; Nguyen & Tran, 2023). At the same time, Vietnamese scholarship on teaching materials and multimodal literacy indicates that visuals are often treated as peripheral rather than systematically taught as meaning-making resources (Unsworth & Ngo, 2014; Hoang, 2015; Tran & Lam, 2025). The thesis contributes pedagogically by offering a principled correction: it identifies what students must learn to do with diagrams, procedures, and role enactments, and thereby clarifies multimodal disciplinary literacy as a legitimate target of EMP teaching in medical universities.

6.4.2 Two Complementary Multimodal Pedagogies

A second pedagogical contribution is the thesis’ articulation of two complementary multimodal pedagogies emerging from the corpus, each with distinct epistemic value. Crucially, the thesis does not rank them; it explains what each pedagogy is designed to make learnable, and why both are necessary for comprehensive EMP competence.

Epistemic multimodal pedagogy (MT)

MT's dominant design logic is conceptual stabilisation through dense elaboration: diagrams, labels, and definitions repeatedly converge to reduce interpretive ambiguity and produce terminological precision. Pedagogically, this constitutes an epistemic pedagogy because it trains learners into the disciplinary habit of construing the body as a systematic object of knowledge – classifiable, nameable, retrievable. This is consistent with multimodal social semiotic accounts of how scientific knowledge is stabilised through semiotic condensation and controlled representation (Kress & van Leeuwen, 2021; Lemke, 1998; Halliday & Matthiessen, 2014). In Vietnamese classrooms, where early-stage EMP often emphasises reading and terminology mastery, this pedagogy aligns with prevailing instructional expectations; nevertheless, the thesis shows it should be understood as a designed epistemic apprenticeship, not merely “*vocabulary support*” (Nguyen, 2000; Nguyen & Tran, 2023).

Enactment-oriented multimodal pedagogy (CPM)

CPM's design logic foregrounds enhancement and projection: images specify procedural contingencies and stage interactional roles that learners are invited to rehearse. Pedagogically, this is enactment-oriented because it treats medical English as professional action – language-as-doing within a simulated workplace. This aligns with ESP perspectives that highlight vocational materials as sites of participation in professional practices rather than only exposure to specialist lexis (Belcher, 2009; Bhatia, 2017; Dudley-Evans & St John, 1998). Multimodally, it reflects the claim that meaning emerges through division of semiotic labour across modes: images do not decorate; they anchor procedures and social roles so that tasks become enactable (Bezemer & Kress, 2016; Unsworth, 2008; Royce, 2007). In Vietnam, this pedagogy speaks directly to documented needs for stronger

clinical communication and scenario-based learning in EMP training (Le et al., 2021; Ton, 2017).

Importantly, the pedagogical payoff of theorising these two pedagogies together is integrative: the thesis clarifies why a “*terminology-first*” course and a “*communication-first*” course often feel pedagogically misaligned. They are built on different multimodal epistemologies. MT privileges recognition and classification literacies; CPM privileges procedural and interactional rehearsal. Once this is made explicit, the route to improvement is no longer framed as choosing one approach over the other, but as designing principled integration across a curriculum – an issue that becomes central in the next section on implications for EMP materials and instruction in Vietnamese medical universities (Tomlinson, 2011; McGrath, 2013; MOET, 2018; Lam & Pham, 2024).

In sum, Section 6.4 positions the thesis’ pedagogical contribution as conceptual clarity: it specifies multimodal disciplinary literacy as the learning target, and it theorises two complementary multimodal pedagogies as the means through which EMP textbooks scaffold different forms of medical knowing. This conceptual consolidation directly prepares the move from “*what the textbooks make possible*” to “*what institutions and teachers can do with that possibility*” in the Vietnamese context.

6.5 Practical and Institutional Implications for the Vietnamese Context

Section 6.4 established that the two textbooks make available two distinct but complementary pedagogic logics: conceptual stabilisation (MT) and enactment-oriented practice (CPM). The practical question for Vietnam, however, is not whether either logic is “*better*” but how Vietnamese medical

universities can institutionalise a coherent pathway from terminology-based disciplinary literacy to communicative and procedural competence in clinical contexts. This matters because Vietnamese EMP scholarship has repeatedly documented a persistent gap between students' lexical knowledge and their readiness for interactional demands in healthcare settings, alongside constraints in time, assessment regimes, and classroom traditions (Ton, 2020; Tran, 2024; Trinh, 2010). Moreover, national policy directions emphasise upgrading foreign-language competence and learning orientations, which implicitly requires moving beyond text-only literacies toward multimodal competence (MOET, 2008). Against that backdrop, the thesis offers three institution-relevant implications.

6.5.1 Implications for EMP Curriculum Design

A first implication is curricular: Vietnamese EMP programmes should be designed as a sequenced integration of conceptual stabilisation and enactment, rather than as a binary choice between “*terminology courses*” and “*communication courses*” The thesis shows that MT and CPM do not simply differ in topic or difficulty; they embody different epistemic commitments about what counts as medical knowing. If curricula treat these commitments as interchangeable – or if they privilege one while neglecting the other – students are likely to develop uneven competence: strong decoding of specialised lexis but weak participation in clinical talk, or conversely fluent classroom role-play with fragile conceptual grounding.

However, integration does not mean mixing activities randomly. It means structuring a pathway in which conceptual resources enable practice, and practice recontextualises conceptual resources. This is consistent with ESP curriculum thinking that emphasises alignment among needs, genres, and

professional tasks (Dudley-Evans & St John, 1998; Belcher, 2009), and with vocational genre research arguing that professional competence is constituted through interdiscursive performance across knowledge and action (Bhatia, 2017). It also aligns with research on learning with multiple representations: learners benefit when representational systems are coordinated and when instruction supports translating meaning across them (Ainsworth, 2006; Mayer, 2014).

Concretely, the curriculum-level implication is to organise EMP progression around two complementary strands:

Foundation strand (epistemic/terminological): diagram literacy, word-part reasoning, definitional precision, and assessment literacy – competences that MT strongly affords through dense diagram-label-definition coupling.

Enactment strand (procedural/interactional): scenario reasoning, interactional routines, and professional role projection – competences that CPM affords through enhancement and projection in task-based ensembles.

Importantly, such integration speaks directly to Vietnamese conditions where imported coursebooks often carry Western pedagogic assumptions that do not automatically fit local assessment cultures and classroom expectations (Tran & Le, 2021). The thesis therefore supports curriculum design that treats materials as resources to be sequenced and recontextualised, not as scripts to be followed verbatim (van Leeuwen, 2005; Tomlinson, 2011).

6.5.2 Implications for Teaching Practice

A second implication concerns pedagogy in the classroom: Vietnamese EMP teaching should explicitly cultivate students' ability to read and use images, layout, and task structures as learning resources, rather than treating visuals as decorative support. This implication follows directly from the

thesis' central finding that meaning is produced through division of semiotic labour across modes (Kress, 2010; Jewitt, 2014; Bezemer & Kress, 2016). If teachers focus primarily on vocabulary lists and translated definitions, they risk leaving a large part of the textbook's meaning-making architecture pedagogically unused.

However, the implication is not that teachers must “*teach art*” or become visual-design specialists. Rather, teaching practice can become more effective by making a small set of multimodal reading routines explicit:

Diagram-guided reasoning: guiding students to infer part-whole relations, classificatory logics, and information hierarchies before naming items – supporting conceptual stabilisation and reducing rote memorisation (Kress & van Leeuwen, 2021; Lemke, 1998).

Procedure tracing: asking students to follow sequence, manner, and condition across image + instruction + layout, so procedural meaning becomes visible and discussable, not left to implicit inference (Unsworth, 2008; Ainsworth, 2006).

Role framing and stance practice: treating scenario images and dialogue/task frames as participation structures, then rehearsing role-appropriate language choices and pragmatic stance (Royce, 2007; Goodwin, 1994).

Critically, this teaching implication aligns with Vietnamese work calling for stronger communicative competence and for more principled exploitation of coursebooks in ESP classrooms (Hoang, 2015; Lam & Pham, 2024). It also responds to reported constraints such as limited time and uneven teacher preparation for EMP: the thesis does not demand a radical methodological

overhaul, but proposes a focused shift in what counts as “text” in EMP instruction – namely, the page ensemble as a coherent semiotic unit.

6.5.3 Implications for Materials Selection and Adaptation

A third implication is institutional: departments should develop clearer criteria for selecting and adapting EMP materials in ways that match programme outcomes and local constraints. Vietnam often relies on imported textbooks; the issue is therefore not access, but fit – including fit to assessment, classroom culture, and clinical communication needs (Tran & Lam, 2025; Le et al. 2023). The thesis provides an evidence-based rationale for using MT and CPM as complementary resources with different strengths.

Using MT effectively in Vietnam. MT is well-suited as a foundation resource because it scaffolds terminological decoding and conceptual organisation through highly stabilised visual-verbal coupling. In Vietnamese contexts where early EMP courses often align with knowledge-based assessment, MT can support “*assessment literacy*” by clarifying what counts as correct identification, naming, and definitional precision. However, the thesis also implies a caution: if MT is used as the sole core material, students may receive limited support for role enactment and interactional reasoning. Therefore, MT should be adopted with an institutional expectation that its conceptual stabilisation will be recontextualised into practice elsewhere in the programme (van Leeuwen, 2005; Halliday & Matthiessen, 2014).

Using CPM effectively in Vietnam. CPM is well-suited for communication-oriented modules because it scripts participation through tasks, staged dialogues, and scenario visuals, supporting role-play and clinical interaction. This is directly relevant to Vietnamese concerns about pragmatic failures and gaps in doctor–patient communication competence (Nguyen,

2010). However, CPM's enactment strength can become a weakness if students lack conceptual grounding: role-play risks becoming formulaic performance without stable understanding of biomedical concepts. Therefore, CPM is most effective when institutions ensure that foundational conceptual resources – whether from MT or other sources – are available and actively connected to practice.

An institutional adaptation principle: rather than “*following the book*” programmes should adopt a deliberate adaptation approach: MT pages can be used to establish conceptual anchors and diagram-reading routines, while CPM pages can be used to rehearse communicative action and procedural sequencing. This principle is consistent with materials development scholarship that stresses contextual adaptation and principled evaluation over textbook fidelity (Cunningsworth, 1995; Tomlinson, 2011; McGrath, 2013), and with Vietnamese discussions of how English teaching materials must be aligned with local pedagogic realities (Hoang, 2015; Tran, 2024).

In sum, Vietnam “*gets*” from this thesis a practical, institutionally actionable map: it clarifies what each textbook is semiotically designed to make learnable, and it converts that explanation into three coherent implications – curriculum sequencing, classroom practice focused on multimodal resources, and evidence-based material selection/adaptation. These implications provide a grounded basis for Vietnamese medical universities to move from imported materials to locally coherent EMP programmes without assuming that imported multimodal designs will automatically produce desired learning in local classrooms.

6.6 Limitations of the Study

This study acknowledges several limitations that arise from its research scope, methodological choices, and theoretical commitments. These limitations are not shortcomings of execution; rather, they define the epistemic boundaries within which the findings are valid and meaningful.

6.6.1 Corpus Scope and Analytic Generalisability

The corpus consists of 40 page ensembles drawn from two EMP textbook series (Medical Terminology: A Living Language and Career Paths: Medical). While this corpus is sufficiently robust for comparative, theory-driven qualitative analysis, it does not aim to represent all EMP materials, nor ESP textbooks in general. This limitation is consistent with interpretive qualitative research, where analytic generalisation and theoretical transferability are prioritised over statistical representativeness (Lincoln & Guba, 1985; Creswell & Poth, 2018).

Accordingly, the findings should be read as explanatory patterns of multimodal design, not as claims about the typicality of all medical English textbooks. The contribution lies in clarifying how meaning is constructed in these materials, not in quantifying how often similar designs occur elsewhere.

6.6.2 Focus on Semiotic Affordances Rather than Learning Outcomes

A second limitation concerns the absence of classroom observation, learner interviews, or outcome-based measures. This study does not investigate how students actually learn from these textbooks, nor does it evaluate instructional effectiveness. Instead, it analyses the semiotic affordances designed into textbook page ensembles.

This is a deliberate methodological choice grounded in multimodal social semiotics, which distinguishes between designed meaning potential and actual uptake in use (Kress, 2010; Bezemer & Kress, 2016). While this choice restricts pedagogic claims to the level of design rather than impact, it preserves theoretical coherence and avoids conflating semiotic analysis with pedagogical evaluation – a distinction emphasised in both international multimodal research and Vietnamese applied linguistics scholarship (Nguyễn Hòa, 2012).

6.6.3 Domain-Specificity of EMP

The study is situated explicitly within ESP, a domain characterised by high conceptual density, strong visual traditions (e.g. diagrams, schemata), and professional communication involving institutional authority and risk (Candlin & Candlin, 2002; Ferguson, 2013). As a result, the intersemiotic patterns identified such as dense elaboration for conceptual stabilisation or projection for role enactment – should not be uncritically extended to other ESP domains.

This limitation reflects a broader consensus in ESP and genre research that disciplinary practices shape semiotic design, and that cross-domain generalisation must be theoretically justified rather than assumed (Hyland, 2006; Bhatia, 2017).

6.6.4 Theoretical Constraints of Multimodal Frameworks

No single multimodal framework offers a complete account of meaning-making. While this study productively draws on Grammar of Visual Design, intersemiotic complementarity, and IMAF, it also recognises their theoretical limits.

In particular, Unsworth's (2008) model of image-text relations, although influential in educational multimodality, has not fully theorised issues such as intersemiotic gradience, dynamic pedagogic sequencing, or epistemic weighting across modes. These limitations have been noted in subsequent multimodal scholarship (Bateman, 2014; Bezemer & Kress, 2016). In this thesis, Unsworth's framework is therefore used selectively and critically, in conjunction with Martinec and Salway (2005) and Royce (2007), rather than being treated as an exhaustive or closed system.

This theoretical positioning demonstrates an explicit awareness that theories are analytic resources to be mobilised under specific conditions, not total explanations – a stance consistent with constructivist and interpretivist research traditions (Sunderland & McGlashan, 2012; Schwandt, 2007).

6.6.5 Interpretive Mediation in Multimodal Analysis

Finally, as with all multimodal discourse analysis, the findings are mediated by the researcher's interpretive decisions. Although systematic coding procedures, analytic matrices, and traceability mechanisms were employed (Appendices B-D), multimodal meaning is not mechanically recoverable. This limitation is intrinsic to multimodal research and is addressed through methodological transparency, theoretical explicitness, and analytic reflexivity, rather than claims of objectivity (Norris, 2019; Miles et al., 2019).

Taken together, these limitations clarify what the study does and does not claim. The thesis does not attempt to predict learning outcomes, generalise across all ESP domains, or offer a totalising multimodal theory. Instead, it offers a controlled, theory-driven explanation of how visual and intersemiotic design in EMP textbooks constructs distinct epistemic and pedagogic

orientations. Within these clearly defined boundaries, the findings remain both valid and theoretically significant. duce desired learning in local classrooms.

6.7 Directions for Future Research

Building directly on the limitations identified in Section 6.6, future research should extend this thesis from designed semiotic affordances to enacted pedagogic effects, and from a bounded corpus to broader, locally grounded EMP ecologies. Importantly, these directions are not add-ons; they are the most coherent “*afterlife*” of an interpretive-explanatory multimodal study that has established what textbooks make possible but has not claimed to show what classrooms make actual (Kress, 2010; Bezemer & Kress, 2016).

6.7.1 Classroom Enactment of Textbook Multimodality

Because this thesis did not include classroom observation or learner data (Section 6.6.2), a first priority is to investigate how teachers and students activate, transform, or bypass textbook multimodal designs in real pedagogic time. Such work would examine how page ensembles become interactional events – how diagrams are pointed to, how tasks are staged, how images are verbalised, and how learners are positioned through talk, gesture, and material handling (Norris, 2004, 2019). Moreover, this direction responds to the social semiotic distinction between meaning potential and meaning in use: multimodal resources do not “*teach*” by themselves; they are mobilised through situated pedagogic action (Kress, 2010). In Vietnamese EMP contexts – where teachers often report constraints on communicative teaching and uneven access to clinical simulation – classroom-based multimodal research can clarify the conditions under which imported materials actually

support, or fail to support, intended professional learning (Ton, 2017; Trinh, 2010; Tran, 2024).

6.7.2 Assessing Multimodal Medical Literacy

A second direction follows from the thesis's theoretical claim that EMP learning involves multimodal disciplinary literacy, not vocabulary alone. However, without outcome measures (Section 6.6.2), the field still lacks principled ways to assess whether learners can (i) interpret diagrammatic representations, (ii) integrate image-text evidence, and (iii) perform role-based clinical communication grounded in multimodal cues. Future studies should therefore develop and validate assessment constructs for multimodal medical literacy, drawing on established work on disciplinary literacy (Airey, 2016) and multimodal literacy (Jewitt, 2014; Mills & Unsworth, 2018). Critically, assessment should not reduce multimodality to “*picture description*”; rather, it should target the epistemic practices the thesis has identified – conceptual stabilisation through elaboration, procedural reasoning through enhancement, and role enactment through projection – so that evaluation aligns with how medical meaning is semiotically organised in EMP materials (Royce, 2007; Unsworth, 2008). This direction is also practically consequential for Vietnam, where assessment regimes can silently steer pedagogy back toward translation and terminology recall unless multimodal competencies are made visible and examinable (MOET, 2018; Hoang, 2015).

6.7.3 Comparing Imported Textbooks with Vietnamese-Produced EMP Materials

The corpus limitation (two textbook series; Section 6.6.1) and the Vietnam-specific rationale of the thesis together motivate a third direction:

systematic comparison between internationally published EMP textbooks and Vietnamese-produced or locally adapted materials. Such comparative work can test whether the two intersemiotic signatures identified here – conceptual stabilisation (MT) and enactment-oriented rehearsal (CPM) – are typical of imported materials, or whether locally produced textbooks embed different epistemic assumptions about what counts as “*medical English competence*”. Moreover, this line of research speaks to wider debates on how global ELT/ESP coursebooks carry implicit cultural and institutional orientations, which may align unevenly with local pedagogic expectations (Gray, 2010; Fairclough, 2001; Le, 2017). Vietnamese studies of textbooks and multimodal resources provide a concrete foundation for this agenda, including analyses of multimodality in Vietnamese ELT materials (Hoang, 2015; Pham, 2021) and early EMP-focused multimodal work (Lam & Pham, 2024). A sharper empirical map of local-versus-imported multimodal design would also allow future researchers to specify what kinds of adaptation are needed – lexical, genre-based, visual, or task-structural – rather than treating “*localisation*” as a general slogan.

6.7.4 *Longitudinal Studies of EMP Curriculum Design and Sequencing*

Finally, because this thesis is corpus-based and cross-sectional (Section 6.6.1), it cannot show how multimodal pedagogies accumulate over time across a programme. Future research should therefore examine longitudinal curriculum designs in which conceptual stabilisation and enactment-oriented learning are not competing options but staged as a coherent developmental trajectory. Here, the thesis’s core comparative finding becomes a curriculum hypothesis: early phases may privilege diagrammatic and classificatory literacies (supporting terminology and biomedical concept formation), while

later phases increasingly require simulation, interactional competence, and risk-sensitive communication – especially in clinical placements (Ferguson, 2013; Candlin & Candlin, 2002). Importantly, longitudinal research would also provide an ideal site to refine the thesis’s theoretical tools: it can test how well IMAF captures sequencing effects, intersemiotic gradience, and shifts in epistemic weighting across modes – precisely the areas where existing frameworks (including Unsworth’s model) remain under-specified (Unsworth, 2008; Bateman, 2014; Bezemer & Kress, 2016). In the Vietnamese context, such studies would directly support institutional decision-making about programme outcomes, materials pathways, and the integration of multimodal literacy into EMP curriculum standards (MOET, 2008; Hoang, 2015).

Taken together, these directions follow a disciplined logic: the thesis has established what multimodal textbook designs afford; subsequent research should demonstrate how those affordances are enacted, assessed, localised, and sequenced. However, the agenda is not merely methodological expansion. More critically, it positions multimodality as an epistemic problem in EMP: how medical knowledge is made learnable through semiotic orchestration, and how Vietnamese institutions can design curricula that develop not only medical vocabulary but also the multimodal literacies required for contemporary biomedical communication (Jewitt, 2014; Kress, 2010; Swales, 2002).

6.8 Concluding Remarks

This study set out to examine how visual and linguistic modes interact to construct medical meaning and disciplinary knowledge in English medical textbooks used in Vietnam. Through the development and application of the IMAF, the study has shown that meaning in these textbooks is realised

through systematic intersemiotic relationships rather than through isolated verbal or visual resources.

The findings indicate that images are not merely supplementary illustrations accompanying written texts. Instead, visual and linguistic modes operate as complementary semiotic resources that jointly realise representational, interpersonal, and textual meanings, contributing to the construction of medical concepts, professional roles, and disciplinary relationships. In this sense, medical knowledge represented in English medical textbooks can be understood as a multimodal and intersemiotic achievement, emerging from the coordinated use of multiple semiotic resources.

The study also demonstrates the analytical value of IMAF as an integrated framework that synthesises visual grammar and multimodal metafunctional analysis. By bringing together these complementary perspectives, IMAF provides a systematic means of explaining how different semiotic modes interact to produce disciplinary meanings in specialised educational discourse.

Overall, the study highlights the importance of approaching English medical textbooks not only as linguistic teaching materials but also as multimodal environments for disciplinary learning. This perspective may contribute to further developments in EMP research, multimodal pedagogy, and the design of educational materials that better reflect the increasingly multimodal nature of contemporary medical communication.

REFERENCES

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>
- Airey, J. (2016). Contexts for EAP. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 69-105). Routledge.
- Azkiyah, I., Hidayat, D. N., Alek, & Dewi, R. S. (2021). A multimodal discourse analysis of Disney+ Hotstar Indonesia TV advertisement. *Journal of English Language and Language Teaching (JELLT)*, 5(1), 14-24. <https://doi.org/10.36597/jellt.v5i1.10023>
- Bateman, J. A. (2014). *Text and image: A critical introduction to the visual/verbal divide*. Routledge.
- Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis*. De Gruyter Mouton.
- Belcher, D. (2009). *English for specific purposes in theory and practice*. University of Michigan Press.
- Bezemer, J., & Kress, G. (2016). *Multimodality, learning and communication: A social semiotic frame*. Routledge.
- Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Longman.
- Bhatia, V.K. (2008). Genre analysis, *ESP and professional practice*. *English for specific purposes*, 27(2), 161-174, <https://doi.org/10.1016/j.esp.2007.07.005>
- Bleakley, A. (2015). *Medical humanities and medical education: How the medical humanities can shape better doctors*. Routledge.
- Bosher, S., & Smalkoski, K. (2002). From needs analysis to curriculum development: Designing a course in health-care communication for immigrant students in the USA. *English for Specific Purposes*, 21(1), 59-79. [https://doi.org/10.1016/S0889-4906\(01\)00002-3](https://doi.org/10.1016/S0889-4906(01)00002-3)

- Bui H., (2004). Foreign language structure in Vietnam education, *Bulletin of Europe Study*, (4), 79-83.
<http://thuvienso.bvu.edu.vn/handle/TVDHBRVT/7767>
- Bui, T. T. N., & Nguyen, H. T. M. (2016). Standardizing English for educational and socio-economic betterment: A critical analysis of English language policy reforms in Vietnam. In R. Kirkpatrick (Ed.), *English language education policy in Asia* (pp. 363-388). Springer.
https://doi.org/10.1007/978-3-319-22464-0_17
- Burns, A. (1999). *Collaborative action research for English language teachers*. Cambridge University Press.
- Canagarajah, A. S. (2013). *Translingual practice: Global Englishes and cosmopolitan relations*. Routledge.
- Candlin, C. N., & Candlin, S. (2002). Discourse, expertise and the management of risk in healthcare settings. *Research on Language and Social Interaction*, 35(2), 115-137.
https://doi.org/10.1207/S15327973RLSI3502_1
- Carney, R., & Levin, N. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14(1), 5-26.
<https://doi.org/10.1023/A:1013176309260>
- Chegeni, N., Kamali, B., Noroozi, A., & Chegeni, N. (2016). General English textbook evaluation: A closer look at "Four Corners". *Theory and Practice in Language Studies*, 6(12), 2325-2330.
<https://doi.org/10.17507/tpls.0612.13>
- Creswell, J.W., (2009). *Research design: Qualitative, quantitative, and mixed methods approach*. (3rd ed.). Sage.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage.
- Cunningsworth, A. (1995). *Choosing your coursebook*. Heinemann.
- Davies, A., & Elder, C. (Eds.). (2004). *The handbook of applied linguistics*. Blackwell Publishing. <https://doi.org/10.1002/9780470757000>

- Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research*, 6(2), 80-88. 10.1177/1558689812437186
- Dimopoulos, K., Koulaidis, V., & Sklaveniti, S. (2003). Towards an analysis of visual images in school science textbooks and press articles about science and technology. *Research in Science Education*, 33(2), 189-216. <https://doi.org/10.1023/A:1025006310503>
- Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Dudley-Evans, T., & St. John, M. (1998). *Developments in English for specific purposes: A multidisciplinary approach*. Cambridge University Press.
- Eggins, S. (2004). *An introduction to systemic functional linguistics* (2nd ed.). Continuum.
- Evans, V., Dooley, J., & Tran, M. T. (2012). *Career paths: Medical*. Express Publishing.
- Ewer, J. R., & Latorre, G. (1969). *A course in basic scientific English*. Longman.
- Fairclough, N. (1995). *Critical discourse Aanalysis: The critical study of language* (1st ed.). Longman.
- Fairclough, N. (2001). *Language and power*. (2nd ed.) Longman.
- Febrianti, Y., & Vu, T. (2023). Understanding the roles of images and intermodal relationships for optimized use of visual and verbal resources in Vietnam's textbooks for lower secondary levels, *Bahasa dan Seni: Jurnal Bahasa, Sastra, Seni, dan Pengajarannya*, 51 (1), 54-72. <https://doi.org/10.17977/um015v51i12023p54>
- Ferguson, G. (2013). English for medical purposes. In: Paltridge, B., Starfield, S. Chichester (Eds.), *The handbook of English for specific purposes*. (pp. 243-261). Wiley-Blackwell.
- Fremgen, B. F., & Frucht, S. S. (2019). *Medical terminology: A living language* (7th ed.). Pearson.

- Foucault, M. (1973). *The birth of the clinic: An archaeology of medical perception*. Tavistock Publications.
- Goodwin, C. (1994). Professional vision. *American anthropologist*, 96(3), 606-633. <https://doi.org/10.1525/aa.1994.96.3.02a00100>
- Gray, J. (2010). The branding of English and the culture of the new capitalism. *Applied Linguistics*, 31(5), <https://doi.org/714-733>. 10.1093/applin/amq034
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.
- Halliday, M. A. K. (2004). *An introduction to functional grammar* (3rd ed.), Hodder Arnold.
- Halliday, M. A. K. (1998). Things and relations: Regrammaticizing experience as technical knowledge. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp. 185-235). Routledge.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
- Hsu, P., & Yang, W. (2007). Print and image integration of science texts and reading comprehension: A systemic functional linguistics perspective. *International Journal of Science and Mathematics Education*, 5(4), 639-659. <https://doi.org/10.1007/s10763-007-9091-x>
- Hoang, V. V. (2010). The current situation and issues of the teaching of English in Vietnam. *Ritsumikan Studies in Language and Culture*, 22(1), 7-18.
- Hoang, V. V. (2015). The development of the ten-year English textbook series for Vietnamese schools under the national foreign language 2020 project: A cross-cultural collaborative experience. *VNU Journal of Foreign Studies*, 31(3), 1-17.
- Hodge, R., & Kress, G. (1988). *Social semiotics*. Cornell University Press.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.

- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.
- Hutchinson, T & Torres, E. (1994). The textbook as agent of change. *ELT Journal*. 4(48), 315-328.
- Jewitt, C., Kress, G., Ogborn, J., & Tsatsarelis, C. (2001). Exploring learning through visual, actional and linguistic communication: The multimodal environment of a science classroom. *Educational Review*, 53 (1), 5-18. <https://doi.org/10.1080/00131910123753>
- Jewitt, C. (2009a). An introduction to multimodality. In: C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 14-27). Routledge.
- Jewitt, C. (2009b). Different approaches to multimodality. In: C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 28-39). Routledge.
- Jewitt, C. (Ed.). (2014). *The Routledge handbook of multimodal analysis* (2nd ed.). Routledge.
- Jewitt, C. & Oyama, R. (2004). Visual meaning: A social semiotic approach. In: T. van Leeuwen, T. & C. Jewitt (Ed.), *The handbook of visual analysis* (pp. 134-156). SAGE Publications. <https://doi.org/10.4135/9780857020062.n7>
- Jewitt, C., & Henriksen, B. (2016). Social semiotic multimodality. In N. M. Klug & H. Stöckl (Eds.), *Handbuch Sprache im multimodalen Kontext: Handbook of Language in Multimodal Contexts* (pp. 145-164). De Gruyter. <https://doi.org/10.1515/9783110296099-007>
- Kress, G. (2000). Multimodality: Challenges to thinking about language. *TESOL Quarterly*, 34(3), 337-340. <https://doi.org/10.2307/3587959>
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold. <https://doi.org/10.2307/1512155>

- Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design* (1st ed.). Routledge.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.
- Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge.
- Lam, Q. D., & Pham, B. N. (2024). A multimodal discourse analysis of a medical textbook in Vietnam: Preliminary findings. *VNU Journal of Foreign Studies*, 40(3), 21-34. <https://doi.org/10.63023/2525-2445/jfs.ulis.5220>
- Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp. 87-113). Routledge.
- Le, T. T., Phan, Q. T. N, & Mai, L. X. (2021). Medical students' evaluation of an English for specific purposes coursebook: A case in Vietnam. *LET: Linguistics, Literature and English Teaching Journal*, 11(2), 1-23.
- Le, V. C. (2017). English language education in Vietnamese universities: National benchmarking in practice. In E. S. Park & B. Spolsky (Eds.), *English education at the tertiary level in Asia: From theory to practice* (pp. 223-246). Routledge.
- Levin, J. R., & Mayer, R. E. (1993). Understanding illustrations in text. In B. K. Britton, A. Woodward, & M. R. Binkley (Eds.), *Learning from textbooks: Theory and practice* (pp. 95-113). Lawrence Erlbaum Associates.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Lim, F. V., O'Halloran, K. L., & Podlasov, A. (2012). Spatial pedagogy: Mapping meaning in the use of classroom space, *Cambridge Journal of Education*, 42(2), 235-251. <https://doi.org/10.1080/0305764X.2012.676629>
- Lupton, D. (2013). *The social worlds of the unborn*. Palgrave Macmillan.

- Machin, D., & Mayr, A. (2023). *How to do critical discourse analysis: A multimodal introduction*. Sage. <https://doi.org/10.4135/9781036212933>
- Machin, D. (2013). *Introduction to multimodal analysis* (2nd ed.). Bloomsbury.
- Makhuble, T., Kekana, T., Mogoboya, M. (2022). Evaluation of English for specific purposes textbooks by medical sciences students at the University of Limpopo: A needs analysis approach. *International Journal of Special Education*, 37(3), 14150-14162.
- Martinec, R., & Salway, A. (2005). A system for image-text relations in new (and old) media. *Visual Communication*, 4(3), 337-371. <https://doi.org/10.1177/1470357205055928>
- Mayer, R. E. & Fiorella, L. (Eds.). (2022). *The Cambridge handbook of multimedia learning* (3rd ed.). Cambridge University Press.
- McGrath, I. (2013). *Teaching materials and the roles of EFL/ESL teachers: Practice and theory*. Bloomsbury.
- Mickan, P. (2017) Text-based research and teaching from a social semiotic perspective: Transformative research and pedagogy. In P. Mickan & E. Lopez (Eds.), *Text-based research and teaching: A social semiotic perspective on language in use* (pp. 15-35). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-59849-3_2
- Ministry of Education and Training. (2008). *Decision No. 1400/QĐ-TTg: Teaching and learning foreign languages in the national education system, period 2008-2020*. http://www.chinhphu.vn/portal/page/portal/chinhphu/hethongvanban?class_id=1&_page=1&mode=detail&document_id=78437
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.
- Mukundan, J. (2011). Developing an English language textbook evaluation checklist. *Contemporary Issues in Education Research (CIER)*. 4(6), 21-28.

- Mukundan, J., & Nimehchisalem, V. (2012). Evaluative criteria of an English language textbook evaluation checklist. *Journal of Language Teaching and Research*, 3(6), 1128-1134. <https://doi.org/10.4304/jltr.3.6.1128-1134>
- Myers, G. (1990). *Writing biology: Texts in the social construction of scientific knowledge*. University of Wisconsin Press.
- Nguyen, H. (2000). *An introduction to discourse analysis*. Vietnam National University Press.
- Nguyen, T. Q. H., & Tran, T. V. A. (2023). Resumés in English from the perspective of multimodal discourse analysis, *The University of Danang-Journal of Science and Technology*, 21 (12.2), 61-66. <https://doi.org/10.31130/ud-jst.2023.300E>
- Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. Routledge. <https://doi.org/10.4324/9780203379493>
- Norris, S. (2019). *Systematically working with multimodal data: Research methods in multimodal discourse analysis*. Wiley-Blackwell.
- O'Halloran, K. L. (2011). Multimodal discourse analysis. In K. Hyland & B. Paltridge (Eds.), *Companion to discourse analysis* (2nd ed., pp. 249-266). Bloomsbury.
- O'Halloran, K. L. (2008). Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication*, 7(4), 443-475. <https://doi.org/10.1177/1470357208096210>
- Painter, C., Martin, J. R., & Unsworth, L. (2013). *Reading visual narratives: Image analysis of children's picture books*. Equinox.
- Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). Sage.
- Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction*, 3(3), 227-238. [https://doi.org/10.1016/0959-4752\(93\)90006-L](https://doi.org/10.1016/0959-4752(93)90006-L)
- Pham, B. N. (2022). *Towards a multimodal analysis of images in a medical textbook*. In *VietTESOL International Convention Proceedings 2021*, (pp. 271-279).

- Rowley-Jolivet, E. (2002). Visual discourse in scientific conference papers: A genre-based study. *English for Specific Purposes*, 21(1), 19-40.
- Royce, T. (1998). Synergy on the page: Exploring intersemiotic complementarity in page-based multimodal text. *JASFL Occasional Papers*, 1(1), 25-49.
- Royce, T. D. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quarterly*, 36(2), 191-205.
- Royce, T. (2007). *Intersemiotic complementarity: A framework for multimodal discourse analysis*. In T. D. Royce & W. L. Bowcher (Eds.), *New directions in the analysis of multimodal discourse* (pp. 63-109). Lawrence Erlbaum Associates.
- Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health*, 18(2), 179-183.
<https://doi.org/10.1002/nur.4770180211>
- Salager-Meyer, F. (2014). Origin and development of English for medical purposes. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 31-50). Wiley-Blackwell.
- Schwandt, T. A. (1998). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues* (pp. 221-259). Sage.
- Schwandt, T. A. (2007). *The SAGE dictionary of qualitative inquiry* (3rd ed.). Sage.
- Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. K. (2010). Science textbooks' use of graphical representations: A descriptive analysis of four sixth grade science texts. *Reading Psychology*, 31(3), 301-325.
<https://doi.org/10.1080/02702710902733550>
- Sunderland, J., & McGlashan, M. (2012). The linguistic, visual, and multimodal representation of two-mum and two-dad families in children's picture books. *Language and Literature*, 21(2), 189-210.
<https://doi.org/10.1177/0963947011435863>
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

- Swales, J. M. (2002). Integrated and fragmented worlds: EAP materials and corpus linguistics. In J. Flowerdew (Ed.), *Academic discourse* (pp. 150-164). Longman.
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4(4), 295-312.
- Tang, K.-S. (2013). Instantiation of multimodal semiotic systems in secondary school science textbooks. *Science Education*, 97(3), 442-478.
- Tang, K.-S., Delgado, C., & Moje, E. B. (2014). An integrative framework for the analysis of multiple and multimodal representations for meaning-making in science education. *Science Education*, 98(2), 305-326. <https://doi.org/10.1002/sce.21099>
- Tomlinson, B. (Ed.). (2011). *Materials development in language teaching* (2nd ed.). Cambridge University Press.
- Ton, N. M. N. (2017). A multimodal analysis of mathematical discourse in English for young learners. *VNU Journal of Foreign Studies*, 33(6), 93-101, <https://doi.org/10.25073/2525-2445/vnufs.4211>
- Ton, N. M. N. Nguyen, T. D. M. & Nguyen, T. H. (2020). A study on modality in English-medium research articles. *VNU Journal of Foreign Studies*, 36 (6), 74-92. <https://doi.org/10.25073/2525-2445/vnufs.4630>
- Tracy, S. J. (2010). Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. *Qualitative Inquiry*, 16(10), 837-851.
- Tran, H. L., & Lam, Q. D. (2025). English-Vietnamese aeronautical engineering textbook cover designs: A contrastive study from the theory of multimodal discourse analysis, *VNU Journal of Foreign Studies*, 41(1), 18-37. <https://doi.org/10.63023/2525-2445/jfs.ulis.5315>
- Tran, T. V. D. (2024), Ứng dụng phân tích diễn ngôn đa phương thức trong dạy học ngoại ngữ, *Journal of Educational Equipment: Applied Research*, 2 (317), 34-36.
- Trevors, G., Muis, K. R., Pekrun, R., Sinatra, G. M., & Winne, P. H. (2016). Identity and epistemic emotions during knowledge revision. *Contemporary Educational Psychology*, 47, 1-15. <https://doi.org/10.1016/j.cedpsych.2016.06.001>

- Trinh, C. D. (2010). English today and tomorrow from a critical perspective. *VNU Journal of Foreign Studies*, 26(3).
<https://doi.org/jfs.ulis.vnu.edu.vn/index.php/fs/article/view/2558>
- Troncoso, C. R. (2010). The effect of language materials on the development of intercultural competence. In B. Tomlinson & H. Masuhara (Eds.), *Research for material development in language learning: Evidence for best practice* (pp.83-102). Continuum.
- Unsworth, L. (2001). *Teaching multiliteracies across the curriculum: Changing contexts of text and image in classroom practice*. Open University Press.
- Unsworth, L. (2006). Towards a metalanguage for multiliteracies education: Describing the meaning-making resources of language-image interaction. *English Teaching: Practice and Critique*, 5(1), 55-76.
- Unsworth, L. (2006). *E-literature for children: Enhancing digital literacy learning*. Routledge.
- Unsworth, L. (2006). Multiliteracies and a metalanguage of image/text relations: Implications for teaching English as a first or additional language in the 21st century. *TESOL in Context*, 16 (Special edition), 147-162.
- Unsworth, L. (2008). *Multiliteracies and metalanguage: Describing image/text relations as a resource for negotiating multimodal texts*. *English Teaching*, 7(2), 55-76.
- Unsworth, L. (2008). Multimodal semiotics: Functional analyses in contexts of education. In L. Unsworth (Ed.), *Multimodal semiotics: Functional analysis in contexts of education* (pp. 1-30). Continuum.
- Unsworth, L. (2014). Multimodal reading comprehension: Curriculum expectations and large-scale literacy testing practices. *Pedagogies: An International Journal*, 9(1), 26-44.
<https://doi.org/10.1080/1554480X.2014.878968>
- Unsworth, L., & Cléirigh, C. (2009). Multimodality and reading: The construction of meaning through image-text interaction. In C. Jewitt

- (Ed.), *The Routledge handbook of multimodal analysis* (pp. 151-163). Routledge.
- Unsworth, L., & Ngo, T. (2014). Vai trò của hình ảnh trong sách giáo khoa dạy tiếng Anh ở Việt Nam [The role of visual images in English textbooks in Vietnam]. *The Journal of Language and Life*, 11 (299), 93-99.
- van Leeuwen, T. (2005). *Introducing social semiotics*. Routledge.
- van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis*. Oxford University Press.
- van Leeuwen, T. (2015). Multimodality in education: Some directions and some questions. *TESOL Quarterly*, 49 (3), 582-589. <https://doi.org/10.1002/tesq.242>
- Vietnam National University. (2012). *Regulations on textbook management of Vietnam National University*. https://vnu.edu.vn/upload/vanban/2012/08/28/2012_05_18-1600_QD-DHQGHN_Quy-dinh-vecong-tac-giao-trinh-cua-DHQGHN---Danh-sach.pdf

APPENDICES

APPENDIX A

PAGE-ENSEMBLE INVENTORY FOR MEDICAL TERMINOLOGY AND CAREER PATHS: MEDICAL

Textbook Source	Book / Edition	Page	Page-Ensemble Description (Visual-Verbal Content)	Primary Visual-Pedagogic Function
MT	7 th Ed. Pearson (2019)	59	Root/suffix formation tables with structured terminology examples	Visual scaffolding of morphological segmentation to support systematic word-building
MT	7 th Ed. Pearson (2019)	93	Dermatological terminology clusters arranged by condition/type	Categorization-based consolidation of semantic fields in organ-system learning
MT	7 th Ed. Pearson (2019)	122	Digestive system terminology listing with grouped functional labels	Anchoring conceptual domains through semantic grouping across medical fields
MT	7 th Ed. Pearson (2019)	148	Urinary system vocabulary patterns organized by clinical relevance	Linking terminology to functional anatomy through structural classification
MT	7 th Ed. Pearson	154	Medical term groupings with brief clinical	Reinforcement of clinical conceptual

	(2019)		reference glosses	meaning via lexicon clustering
MT	7 th Ed. Pearson (2019)	170	Reproductive system terminology sets and semantic associations	Conceptual coherence between biological function and terminology learning
MT	7 th Ed. Pearson (2019)	205	Musculoskeletal/orthopedic clustered terminology lists	Systematic naming for structural organization and knowledge recall
MT	7 th Ed. Pearson (2019)	231	Neurological terminology sets arranged morphologically	Cognitive mapping of neural-system lexicons through visual linguistic segmentation
MT	7 th Ed. Pearson (2019)	273	Respiratory terminology classification tables	Structured semantic fields enabling conceptual stability in system-based learning
MT	7 th Ed. Pearson (2019)	279	Respiratory pathological/diagnostic procedural naming patterns	Representing disease states as patterned variations of medical lexicon
MT	7 th Ed. Pearson (2019)	290	Cardiovascular systemic terminology clusters	Hierarchical semantic organization for conceptual clarity and retention

MT	7 th Ed. Pearson (2019)	316	Hematology/Immunology vocabulary mappings	Visual-semantic integration of biochemical and immunological lexicon
MT	7 th Ed. Pearson (2019)	330	Endocrine-related terminology groupings	Structuring hormone-related discourse through controlled vocabulary systems
MT	7 th Ed. Pearson (2019)	349	Musculoskeletal diagnostic and procedural terms	Clinical classification logic expressed through lexical naming conventions
MT	7 th Ed. Pearson (2019)	355	Disease–procedure pairing terminology	Reinforcing relational semantic meaning in diagnostic workflow
MT	7 th Ed. Pearson (2019)	401	Clinical descriptive terminology for symptoms/signs	Preparing learners for clinical documentation discourse
MT	7 th Ed. Pearson (2019)	404	Procedural vocabulary code	Mapping language to standardized clinical intervention practices
MT	7 th Ed. Pearson (2019)	431	Oncology and diagnostic imaging terminology	Specialized lexicon acquisition for advanced diagnostic interpretation

MT	7 th Ed. Pearson (2019)	476	Pharmacological naming patterns	Systematic professional encoding of medication classification
MT	7 th Ed. Pearson (2019)	496	Glossary-style cumulative term consolidation	Summative reinforcement of complete lexical framework
CPM	Book 1	9	Hospital role diagram with staff/patient positioning	Introducing institutional medical hierarchy and professional identities
CPM	Book 1	22	Patient admission / intake dialogue scene	Modeling interpersonal clinical communication routines
CPM	Book 1	27	Medical equipment labeling chart	Vocabulary acquisition through object-function visual matching
CPM	Book 2	9	Reception interaction scene	Social courtesy language in care- giving encounters
CPM	Book 2	18	Symptom communication conversation	Verbal expression of bodily experience in a patient-centered frame
CPM	Book 2	21	Physical examination	Directives and

			instruction dialogue	procedural communication in clinical tasks
CPM	Book 2	22	Treatment recommendation interaction	Negotiated clinical decision-making discourse
CPM	Book 2	26	Appointment scheduling scenario	Administrative language embedded in medical process continuity
CPM	Book 2	28	Diagnostic explanation exchange	Translating medical knowledge into patient-understandable form
CPM	Book 2	32	Lab result explanation dialogue	Verbalization of biomedical data in consultative discourse
CPM	Book 2	33	Preventive health counseling scene	Instructional communication for public health and lifestyle guidance
CPM	Book 3	7	Operating theatre teamwork illustration	Role-coordination and collective workflow representation
CPM	Book 3	8	Emergency response communication	High-stakes clinical clarity and rapid directive discourse

CPM	Book 3	9	ICU monitoring discussion	Precision reporting and closed-loop communication
CPM	Book 3	10	Ward round professional interaction	Hierarchical yet collaborative clinical reasoning exchange
CPM	Book 3	22	Chronic care follow-up	Long-term clinician– patient relational management
CPM	Book 3	24	Palliative/end-of-life communication	Empathic stance- taking and emotional sensitivity
CPM	Book 3	25	Cross-cultural care dialogue	Sociocultural competence and intercultural communication in healthcare
CPM	Book 3	27	Informing diagnosis conversation	Ethical patient communication during critical events
CPM	Book 3	28	Discharge planning instructions	Coordination language for patient transition- of-care

APPENDIX B

FULL RUBRICIZED MULTIMODAL CODEBOOK

APPENDIX B1. VISUAL CODEBOOK

Code	Visual Category	Criteria for Identification	Indicators in Page Ensembles	Example Pages
V-REP-CON	Conceptual Representational Image	Shows structures, body systems, components, or taxonomies	Anatomical diagrams, labeled organs, structural charts	MT- P205, MT- P231, MT- P316
V-REP-PROC	Procedural Representational Image	Depicts clinical process or sequence of actions	Arrows, step sequences, illustrated movement	MT- P476, CPM- B3 P24
V-INT-OFF	Offer Image (no gaze)	Human figure does not look at viewer → encourages observation	Patient diagrams, anatomical posters	MT- P273, MT- P349
V-INT-DEMAND	Demand Image (direct gaze)	Human figure looks at viewer → invites interpersonal alignment	Doctor–patient scenes, communication photos	CPM- B1P9, CPM- B3P8

V-COMP- IDEAL- REAL	Vertical Information Hierarchy	Top = conceptual / bottom = application	Headers → diagrams → examples	MT- P59, MT- P93
V-COMP- GIVEN- NEW	Left → Right Knowledge Flow	Left = established info, Right = new info or outcome	Clinical dialogues arranged horizontally	CPM- B2P21, CPM- B3P27
V-COMP- SAL	Visual Salience Marking	Color, contrast, zoom, bold labels highlight focal item	Red color- coded anatomy, arrows pointing	MT- P154, MT- P290

APPENDIX B2. VERBAL CODEBOOK

Code	Linguistic Function	Criteria	Common Signals in Text	Example Pages
T-TERM	Terminological Definition	Names and classifies medical entities	Bold terms, suffix/root breakdowns	MT- P148, MT- P401
T-PROC	Procedural / Instructional Language	Describes sequence or clinical action	Imperatives: assess, measure, record	CPM- B2P33, CPM- B3P10
T-EXPL	Explanatory Clarification	Explains function, mechanism, or purpose	“This structure functions to...”	MT- P316, MT- P476
T-INT	Interpersonal Patient-Provider Dialogue	Models tone, empathy, communication	Dialogue scripts, patient questions	CPM- B1P22, CPM- B3P25

APPENDIX B3. INTERSEMIOSIS CODEBOOK

Code	Intersemiotic Relation Type	Definition	Methods of Detection	Example Pages
I-ELAB	Elaboration (Image = Text)	Text and image convey the same conceptual content	Labels match diagram elements	MT- P205, MT- P231
I-EXT	Extension (Text Adds Information)	Text introduces new concepts beyond what is shown	Clinical explanation extends static diagram	MT- P316, CPM- B2P26
I- ENACT	Enactment (Image Performs the Text)	Image demonstrates the action described in text	Procedural sequences, communication role-play	CPM- B1 p.9, CPM- B3P24
I-DISJ	Disjunction (Weak Link)	Image and text co-present but not integrated	Decorative image unrelated to learning goal	Rare (but appears lightly in CPM covers)

APPENDIX B4. AUTO-CODED PAGE ENSEMBLES (MEDICAL TERMINOLOGY)

Page Ensemble	Visual Representational Code	Visual Interactive Code	Visual Compositional Code	Verbal Function Code	Intersemiosis Code	Meaning Summary
MT- P59	V-REP-CON (Conceptual: classificatory schema)	V-INT-OFF (Observer role)	V-COMP-HIER (Hierarchical layout)	T-TERM (Terminology building)	I-ELAB (Elaborative coupling)	The page constructs a taxonomic system of terms, visually anchoring linguistic elements to structured medical categories
MT- P93	V-REP-CON	V-INT-OFF	V-COMP-LABELED (Label sequencing)	T-MORPH (Morphological breakdown)	I-ALIGN (Form-to-meaning mapping)	Visual hierarchy reinforces morphological decoding of medical terms

MT- P122	V-REP-PROC (Processual depiction)	V-INT-OFF	V-COMP- FLOW (Information sequencing flow)	T-EXPL (Explanatory text)	I-EXT (Extension from image to function)	Shows function/process; text clarifies physiological mechanism
MT- P148	V-REP-CON	V-INT-OFF	V-COMP-HIER	T-TERM	I-ELAB	Visual system maps terminology to organ-level anatomy
MT- P154	V-REP-CON	V-INT-OFF	V-COMP- LABELED	T-TERM	I-ALIGN	Emphasizes naming accuracy through structural labeling
MT- P170	V-REP-PROC	V-INT-OFF	V-COMP- FLOW	T-EXPL	I-EXT	Visualizes procedural function and system interaction

MT- P205	V-REP-CON	V-INT-OFF	V-COMP-HIER	T-TERM	I-ELAB	Terminology learning is anchored in sectional anatomy
MT- P231	V-REP-CON	V-INT-OFF	V-COMP-LABELED	T-MORPH	I-ALIGN	Breaks lexical items along anatomical boundaries to support memory
MT- P273	V-REP-PROC	V-INT-OFF	V-COMP-FLOW	T-EXPL	I-EXT	Depicts operational system interaction to reinforce meaning construction
MT- P279	V-REP-CON	V-INT-OFF	V-COMP-HIER	T-TERM	I-ELAB	Conceptual map of structures guides semantic field formation

MT- P290	V-REP-CON	V-INT-OFF	V-COMP-LABELED	T-EXPL	I-ELAB	Labeling supports unitization of system components
MT- P316	V-REP-PROC	V-INT-OFF	V-COMP-FLOW	T-EXPL	I-EXT	Physiological pathway visually integrated with explanatory text
MT- P330	V-REP-CON	V-INT-OFF	V-COMP-HIER	T-TERM	I-ALIGN	Structured visual-territorial segmentation aids lexical stability
MT- P349	V-REP-CON	V-INT-OFF	V-COMP-LABELED	T-TERM	I-ELAB	Reinforces systemic relations in anatomy terminology learning

MT- P355	V-REP-PROC	V-INT-OFF	V-COMP- FLOW	T-EXPL	I-EXT	Dynamizes physiological sequence understanding
MT- P401	V-REP-CON	V-INT-OFF	V-COMP-HIER	T-TERM	I-ELAB	Conceptual chart supports memory of categorical terms
MT- P404	V-REP-CON	V-INT-OFF	V-COMP- LABELED	T-MORPH	I-ALIGN	Morphological clarity linked to structural visualization
MT- P431	V-REP-PROC	V-INT-OFF	V-COMP- FLOW	T-EXPL	I-EXT	Functional process depiction helps narrative reasoning

MT- P476	V-REP-CON	V-INT-OFF	V-COMP-HIER	T-TERM	I-ELAB	Terminology anchored in spatialized organization
MT- P496	V-REP-CON + V-REP-PROC (Hybrid conceptual-process)	V-INT-OFF	V-COMP-INTEGR (Integrated layout)	T-EXPL + T-TERM	I-SYN (Metafunctional synthesis)	Integrates naming + function, preparing transition to applied medical reasoning

APPENDIX B5. AUTO-CODED PAGE ENSEMBLES (CAREER PATHS: MEDICAL)

Page Ensemble	Visual Representational Code	Visual Interactive Code	Visual Compositional Code	Verbal Function Code	Intersemiosis Code	Meaning Summary
CPM- B1P9	V-REP-PERS (Personalized scenario)	V-INT-ENG (Engagement, direct gaze)	V-COMP-SPLIT (Scenario + glossary split layout)	T-DIAL (Dialogue practice)	I-ANCH (Image anchors communicative context)	Images situate vocabulary in real medical interaction
CPM- B1P22	V-REP-PERS	V-INT-SOC (Social interaction, equal status)	V-COMP-BAL (Symmetrical layout)	T-ROLE (Role-play instructions)	I-COOR (Coordination of speech + action cues)	Visuals guide interpersonal dynamics in clinical settings

CPM- B1P27	V-REP-PROC (Procedural: clinical operations)	V-INT-OBS (Learner as onlooker)	V-COMP- FLOW (Ordered instructional panels)	T-EXPL (Step explanations)	I-EXT (Text extends visual procedure)	Depicts task sequences to support performance-based learning
CPM- B2P9	V-REP-PERS	V-INT-ENG	V-COMP- LABELED (Label + definition)	T-VOC (Vocabulary introduction)	I-ALIGN (Meaning alignment via caption)	Connects terms to familiar medical equipment
CPM- B2P18	V-REP-PERS	V-INT-SOC	V-COMP- SPLIT	T-DIAL	I-ANCH	Dialogue visualized in patient-interview scenario
CPM- B2P21	V-REP-PROC	V-INT-OBS	V-COMP- FLOW	T-EXPL	I-EXT	Clinical procedure depicted in step- by-step manner

CPM- B2P22	V-REP-PERS	V-INT-ENG	V-COMP-BAL	T-ROLE	I-COOR	Visuals scaffold communicative stance and professional tone
CPM- B2P26	V-REP-PERS	V-INT-SOC	V-COMP-SPLIT	T-VOC	I-ANCH	Vocabulary grounded in interpersonal clinical task
CPM- B2P28	V-REP-PROC	V-INT-OBS	V-COMP-FLOW	T-EXPL	I-EXT	Task-oriented learning through action sequencing
CPM- B2P32	V-REP-PERS	V-INT-SOC	V-COMP-BAL	T-DIAL	I-COOR	Visual framing guides tone, empathy, and communication

CPM- B2P33	V-REP-PROC	V-INT-OBS	V-COMP- FLOW	T-EXPL	I-EXT	Integration of visual steps with procedural instructions
CPM- B3P7	V-REP-PERS	V-INT-SOC	V-COMP- SPLIT	T-DIAL	I-ANCH	Establishes patient-care relational dynamics
CPM- B3P8	V-REP-PERS	V-INT-SOC	V-COMP-BAL	T-ROLE	I-COOR	Images help learners adopt professional identities in talk
CPM- B3P9	V-REP-PROC	V-INT-OBS	V-COMP- FLOW	T-EXPL	I-EXT	Shows procedural communication embedded in care delivery

CPM- B3P10	V-REP-PERS	V-INT-ENG	V-COMP-SPLIT	T-VOC	I-ALIGN	Vocabulary linked to embodied clinical setting
CPM- B3P22	V-REP-PROC	V-INT-OBS	V-COMP-FLOW	T-EXPL	I-EXT	Depiction of hospital workflow supports action-based reasoning.
CPM- B3P24	V-REP-PERS	V-INT-SOC	V-COMP-BAL	T-DIAL	I-COOR	Visual signals model empathy-oriented communication
CPM- B3P25	V-REP-PROC	V-INT-OBS	V-COMP-FLOW	T-EXPL	I-EXT	Breaks down protocol execution into teachable stages.

CPM- B3P27	V-REP-PERS + V-REP-PROC (Hybrid)	V-INT-SOC	V-COMP- INTEGR (Integrated scenography)	T-DIAL + T- ROLE	I-SYN (Intersemiotic synthesis)	Integrates talk + action to create situated clinical competence
CPM- B3P28	V-REP-PERS	V-INT-ENG	V-COMP-BAL	T-DIAL	I-COOR	Reinforces professional bedside manner and rapport building

APPENDIX C. PAGE-ENSEMBLE ANALYTICAL MATRIX

This appendix presents the systematic coding of page-ensembles drawn from *Medical Terminology: A Living Language* and *Career Paths: Medical*, based on the Rubricized Multimodal Codebook (Appendix B). Each page-ensemble is analyzed across three semiotic modules:

- **Module A (A1-A3):** Visual Meaning based on Kress and van Leeuwen's (2021) *Grammar of Visual Design*, addressing representational structure (A1), interactive viewer relations (A2), and compositional organization (A3).
- **Module B (B1-B3):** Linguistic Meaning, assessing how accompanying text encodes ideational meaning (B1), interpersonal meaning (B2), and textual organization (B3), following Halliday & Matthiessen (2014) and Unsworth (2006).
- **Module C (C1-C3):** Intersemiotic Coupling, examining how visual and linguistic modes work together to produce meaning: elaboration (C1), extension (C2), and enhancement (C3), informed by Royce (2007) and Martinec & Salway (2005).

- Scores are assigned on the 0–3 scale (Appendix B):

Symbol	Score	Interpretation
✓	3	Meaning feature is strongly and explicitly realized
•	2	Meaning feature is moderately realized and context-dependent
()	0-1	Meaning feature is weak or absent

- To provide a summary indication of meaning density:
 $A\Sigma = A1 + A2 + A3 \rightarrow$ Strength of visual meaning construction
 $B\Sigma = B1 + B2 + B3 \rightarrow$ Strength of linguistic meaning construction
 $C\Sigma = C1 + C2 + C3 \rightarrow$ Strength of intersemiotic integration
- The column Visual Meaning Function identifies the pedagogic semiotic role of each image (e.g., anatomical labeling, process demonstration, taxonomic differentiation, systemic mapping).
- The final column, Concise Analytic Note (linked to RQ1), provides a short research-oriented interpretation indicating how the image contributes to disciplinary knowledge-building and multimodal meaning construction, directly addressing RQ1.1, RQ1.2, and/or RQ1.3.

Ord.	Page Ensemble ID	Representational Focus of Page Ensembles	A1	A2	A3	AΣ	B1	B2	B3	BΣ	C1	C2	C3	CΣ	Concise Analytic Note (linked to RQ1)
1	MT-P59 (Figures 3-1 & 3-2: Skin structure cross-section + photomicrograph)	Anatomical cross-sectional visualization showing hierarchical tissue organization	✓	•	✓	8/9	✓	•	✓	8/9	✓	•	()	5/9	The detailed cross-sectional layering visually elaborates technical terminology (e.g., epidermis/dermis/subcutaneous), enabling intersemiotic elaboration of micro-macro correspondence. Supports RQ1.1 (representational organization) and RQ1.2 (image-text coupling for concept stabilization).
2	MT-P93 (Figures 4-1 & 4-2: Bone)	Taxonomic classification + structural decomposition	✓	()	✓	7/9	✓	•	✓	8/9	✓	()	()	4/9	The combination of whole-to-part visual structuring with morphological terminology reinforces systematic

	classification + long bone components)														categorization, exemplifying compositional hierarchy relevant to RQ1.1. Text supports semantic labeling but intersemiotic coupling remains mostly elaborative, informing RQ1.2.
3	MT-P122 (Figure 4-23: Skeletal, smooth, cardiac muscle comparison)	Comparative morphological differentiation for functional contrast	✓	✓	✓	9/9	✓	()	✓	7/9	✓	•	()	5/9	Morphological comparison maps muscle type to functional role, forming a visually-driven conceptual taxonomy. Strong interactive positioning draws learner into embodied interpretation, linking to RQ1.1 & RQ1.3 (pedagogical reasoning).

4	MT-P148 (Figure 5-1: Pulmonary vs systemic circulation diagram)	System process mapping and directional flow visualization	✓	()	✓	7/9	✓	✓	✓	9/9	✓	•	•	7/9	The diagram externalizes a dynamic physiological process using color-coded semiotic contrast (oxygenated vs deoxygenated), enabling enhancement intersemiosis (C3) and supporting computational reasoning. Strong relevance to RQ1.2.
5	MT-P154 (Figures 5-6 & 5-7: Cardiac conduction + ECG waveform)	Process-sequence mapping with temporal progression	✓	✓	•	7/9	✓	✓	✓	9/9	✓	✓	•	7/9	The conduction pathway schematic maps directly onto the ECG temporal trace, forming a cross-modal causal chain. This is a clear case of intersemiotic enhancement, supporting RQ1.2 and conceptual reasoning in clinical training (RQ1.3).

6	MT-P170 (Holter / Stress Test / CPR / Defibrillation)	Clinical monitoring + intervention task visualization showing device-patient interaction	✓	•	✓	8/9	✓	✓	•	8/9	✓	()	()	5/9	Visuals clarify procedural purpose and human-device interaction, supporting RQ1.2 (applied meaning + functional interpretation).
7	MT-P205 (Tonsils / Spleen / Thymus / Immunity)	Anatomical immune tissue localization + functional role mapping	✓	✓	•	8/9	✓	•	✓	8/9	✓	•	()	6/9	Shows immune organ locations with labeled function cues → strengthens micro–macro immune system understanding (RQ1.1 & RQ1.2).
8	MT-P231 (Respiratory Nasopharynx–Laryngopharynx)	Cross-sectional airway spatial pathway mapping	✓	✓	✓	9/9	✓	✓	•	8/9	✓	•	()	6/9	Anatomical sectional labeling stabilizes terminology through spatial correspondence → RQ1.1 & RQ1.2.

9	MT-P273 (Teeth: Upper/Lower + X-ray + Internal Structure)	Multi-view surface / radiographic tissues representation integration	✓	✓	✓	9/9	✓	✓	✓	9/9	✓	•	()	6/9	Three-mode visualization strengthens intersemiotic alignment → concept stabilization (RQ1.1 & RQ1.2).
10	MT-P279 (Salivary / Liver / Gallbladder ducts)	Digestive accessory secretion + duct pathway visualization	✓	•	✓	8/9	✓	✓	•	8/9	✓	()	()	5/9	Sequential organ-duct structure visually supports functional continuity → RQ1.1.
11	MT-P290 (Polyposis / UC / Volvulus)	Pathological morphology illustrating abnormal tissue change	✓	✓	•	8/9	✓	✓	•	8/9	✓	•	()	7/9	Visible disease pattern differences support conceptual contrast → RQ1.2 (image-text stabilization of pathology terms).

12	MT-P316 (Nephron + Bladder)	Hierarchical micro–macro system scaling (filtration → storage)	✓	✓	✓	9/9	✓	•	✓	8/9	✓	()	()	5/9	Reinforces systemic function explanations via scale-linked visuals → RQ1.1.
13	MT-P330 (Lithotripsy + Hemodialysis)	Therapeutic task sequence & medical apparatus visualization	•	✓	✓	7/9	✓	✓	✓	9/9	✓	•	()	6/9	Visuals anchor procedural meaning and patient interaction → RQ1.2.
14	MT-P349 (Ovarian Cycle + Ovum release)	Temporal biological progression (follicular development sequence)	✓	✓	✓	9/9	✓	✓	•	8/9	✓	()	()	6/9	Sequential staging builds conceptual continuity → RQ1.1.

15	MT-P355 (Stages of Labor & Delivery)	Chronological event staging visual narrative	✓	✓	✓	9/9	✓	✓	•	8/9	✓	•	()	7/9	Step-wise birth process visualization reinforces procedural understanding → RQ1.2.
16	MT-P401 (Endocrine System: Parathyroid & Pineal + anatomical positioning)	Endocrine gland localization and structural relations	✓	✓	•	7/9	✓	✓	•	7/9	✓	•	•	5/9	Visual spatial positioning clarifies gland location relative to adjacent organs, supporting RQ1.1 (representational mapping) and RQ1.2 (terminology-image coupling).
17	MT-P404 (Thymus & Thyroid Glands; thoracic spatial view)	Thoracic endocrine organ spatial integration	✓	✓	✓	9/9	✓	✓	•	7/9	✓	✓	•	7/9	The juxtaposition of anterior thoracic anatomy with endocrine labeling facilitates functional-locational reasoning, supporting RQ1.1.

18	MT-P431 (Brain lobes color-coded diagram)	Macro-level neuroanatomical regional differentiation	✓	✓	✓	9/9	✓	✓	✓	9/9	✓	✓	✓	9/9	High color-contrast partitioning and anchored terminology enable stable conceptual chunking of brain regions, strongly supporting RQ1.1-RQ1.2.
19	MT-P476 (Muscles of the Eye + Strabismus visual examples)	Functional-anatomical mapping of extraocular muscles, showing directional pull and clinical deviation outcomes	✓	✓	✓	9/9	✓	✓	✓	9/9	✓	✓	✓	9/9	The Page Ensemble integrates structural anatomy (muscle arrangement) with functional biomechanics (movement direction) and clinical manifestations (strabismus). The visual contrast between normal pull vs. misalignment stabilizes concept formation through intersemiotic coupling of form-function-outcome. Strongly supports RQ1.1 (representational

															organization) and RQ1.2 (image-text meaning integration).
20	MT-P496 (Hearing pathway, outer → middle → inner ear)	Sequential auditory conduction pathway representation	✓	✓	✓	9/9	✓	✓	✓	9/9	✓	✓	✓	9/9	The stepwise depiction of sound conduction supports causal chain reasoning and conceptual integration of auditory physiology (RQ1.2).
21	CPM-B1P9 (Supply Room Procedures)	Procedural communication visualization supporting task-oriented clinical workflow	✓	•	•	1/9	✓	✓	•	2/9	✓	•	•	1/9	Conversational and visual integration reinforces procedural reasoning. The combination of the listening text, poster, and supply form exemplifies intersemiotic scaffolding of workplace communication. Supports RQ1.2 (image-text coupling for procedural stabilization).

22	CPM-B1P21 (First Aid Techniques and Injuries)	Step-wise clinical intervention sequencing through labeled pathological representations	✓	✓	•	2/9	✓	✓	✓	3/9	✓	•	✓	2/9	Labeled images of burns, lacerations, and CPR illustrate treatment progression. Visual-text coordination enhances learners' comprehension of procedural order and clinical terminology. Supports RQ1.1 (representational organization) and RQ1.2 (image-text alignment).
23	CPM-B1P27 (Medication Dosage and Charting)	Quantitative—procedural mapping of drug dosage and administration routines	✓	•	•	1/9	✓	✓	•	2/9	✓	✓	•	2/9	Dialogue-driven dosage explanation aligns with chart-recorded data, modeling authentic nurse communication. The integrated page ensemble strengthens numeric accuracy and supports RQ1.2 (intersemiotic correspondence for dosage precision).

24	CPM-B2P9 (Skin condition consultation dialogue)	Exchange of diagnosis + treatment recommendations between doctor and patient.	✓	•	•	1/3	✓	✓	•	2/3	✓	•	•	1/3	Image + dialogue jointly construct clinical context → strengthens lexico-semiotic grounding of symptoms & advice → supports RQ1.2.
25	CPM-B2P18 (Taking a medical history)	Structuring patient history; elicitation of symptom onset & background.	•	✓	✓	2/3	✓	•	•	1/3	✓	•	•	1/3	Text-heavy; images are orienting, not explanatory → supports concept organization and procedural sequencing → RQ1.1.
26	CPM-B2P21 (Flu symptoms dialogue)	Symptom reporting + physician questioning + care recommendation.	✓	•	✓	2/3	✓	✓	•	2/3	✓	•	•	1/3	Visual cues reinforce symptom vocabulary & bodily states → enhances contextualization of clinical expressions → RQ1.2.

27	CPM-B2P22 (Physical examination steps)	Direct procedural demonstration of physical assessment techniques.	✓	✓	✓	3/3	✓	✓	•	2/3	✓	✓	•	2/3	Strong intersemiotic elaboration: images show technique while text names + explains → supports RQ1.1 + RQ1.2 robustly.
28	CPM-B2P26 (Diagnostic imaging equipment)	Functional and comparative identification of medical imaging devices.	✓	•	✓	2/3	✓	•	✓	2/3	✓	•	✓	2/3	Images support recognition but require inferential linking to text for meaning → supports conceptual differentiation → RQ1.1.
29	CPM-B2P28 (Family medicine overview)	Role, scope, and responsibilities in primary care.	•	✓	•	1/3	✓	•	•	1/3	✓	•	•	1/3	Images set context and role identity, but conceptual meaning is mainly text-driven → supports network-building → RQ1.1.
30	CPM-B2P32 (Geriatrics)	Common conditions + care challenges in	✓	✓	•	2/3	✓	✓	•	2/3	✓	•	✓	2/3	Emotional/affective imagery strengthens social-cognitive framing of care → supports

	+ elderly patient care)	aging populations.													RQ1.2 (compassion + understanding).
31	CPM-B2P33 (Nurse-to-nurse medication interaction dialogue)	Communication of medication changes + drug interaction risks.	✓	•	✓	2/3	✓	✓	✓	3/3	✓	•	•	1/3	Realistic scenario → images anchor clinical reasoning + safety communication → supports RQ1.2 (practice readiness).
32	CPM-B3P7 (Medication Record Problem – Nurses reviewing dosage notes)	Visual-linguistic alignment for clinical record verification and procedural checking	✓	•	✓	8/9	✓	•	✓	8/9	•	✓	•	5/9	Image cues reinforce attention to accuracy in medical documentation, supporting intersemiotic reasoning between dosage terms and patient charts → Supports RQ1.1 & RQ1.2.

33	CPM-B3P8 (Communicating with Patients - Memo + images of interaction)	Representation of interpersonal communicative competence in clinical settings	✓	•	✓	8/9	✓	✓	•	8/9	•	✓	•	5/9	Visuals illustrate rapport, empathy, anxiety, enabling learners to interpret interpersonal stance → Supports RQ1.2 (interpersonal metafunction).
34	CPM-B3P9 (Patient Communication Strategies - Doctor-Nurse conversation + assessment images)	Diagnostic talk + patient engagement visualized through staged dialogue	✓	✓	•	8/9	✓	•	✓	8/9	•	✓	•	5/9	Image-text pairing stabilizes meanings of <i>symptoms</i> , <i>reassurance</i> , <i>rapport-building</i> → Enables image-supported pragmatic competence (RQ1.2).

35	CPM-B3P10 (Challenges in Medical Profession -economic strain, caseload, paperwork)	Conceptual-expository visualization of systemic healthcare challenges	✓	•	✓	8/9	✓	•	✓	8/9	✓	•	•	5/9	Images anchor abstract institutional concepts (workload, burnout), grounding disciplinary literacy (RQ1.1).
36	CPM-B3P22 (ER Triage Guidelines + labeled clinical procedural images)	Procedural visualization of emergency intake and priority protocols	✓	•	✓	8/9	✓	•	✓	8/9	•	✓	•	5/9	Intersemiotic sequencing stabilizes triage logic: injury severity → resource allocation → treatment priority (RQ1.1 & RQ1.2).

37	CPM-B3P24 (Hospital Infections Overview - pathogens + sanitation images)	Cause-effect representation of infection spread + prevention	✓	✓	•	8/9	✓	•	✓	8/9	•	✓	•	5/9	Visual exemplars of vectors and sanitization concretize microbiological abstraction, enhancing conceptual clarity (RQ1.1).
38	CPM-B3P25 (Preventing MRSA-Nurse-Doctor clinical decision interaction)	Clinical reasoning visualized through treatment evaluation dialogue	✓	•	✓	8/9	✓	✓	•	8/9	•	✓	•	5/9	Coupling infection terminology with treatment negotiation strengthens procedural communication literacy (RQ1.2).

39	CPM-P27 (Job Interview + Role Identity Construction)	Professional identity & role expectation visualization	✓	•	✓	8/9	✓	•	✓	8/9	•	✓	•	5/9	Visual framing positions the nurse as credentialed professional, supporting social-interpersonal disciplinary positioning (RQ1.3).
40	CPM-P28 (OB/GYN Clinic Services - pregnancy, labor, menopause visuals)	Lifecycle medical care representation with patient condition transitions	✓	✓	•	8/9	✓	•	✓	8/9	•	✓	•	5/9	Visuals scaffold terminology progression (<i>trimester</i> → <i>contractions</i> → <i>labor</i> → <i>recovery</i>), enabling semantic stability across stages (RQ1.1 & RQ1.2).

APPENDIX D - TRACEABILITY OF MULTIMODAL ANALYSIS ACROSS PAGE ENSEMBLES

This appendix documents the traceability of analytical claims presented in Chapters IV and V to specific page ensembles and coding categories. It demonstrates how multimodal patterns were systematically derived from the rubricised codebook and analytical matrix.

A. MEDICAL TERMINOLOGY: A LIVING LANGUAGE

Table D.1 Medical Terminology: A Living Language (Fremgen & Frucht, 2019)

Page Ensemble ID	Textbook Location	Visual Profile (IMAF-S1)	Intersemiotic Relation (IMAF-S2)	Pedagogic Meaning (IMAF-S3)	Chapter IV	Chapter V	Linked Appendices
MT-P59	Pearson (2019), p.59	Conceptual-classificatory Offer; long shot Centre-Margin	Elaboration	Terminological decoding; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P93	Pearson (2019), p.93	Conceptual-analytical Offer; long shot Ideal-Real	Elaboration	Conceptual classification; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P122	Pearson (2019), p.122	Conceptual-analytical Offer; long shot Centre-Margin	Elaboration	Terminological decoding; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C

MT-P148	Pearson (2019), p.148	Narrative-actional Offer; long shot Given-New	Enhancement	Procedural understanding; Learner as observer	4.3.2	5.4	A; B1; B3; B4; C
MT-P154	Pearson (2019), p.154	Narrative-actional Offer; long shot Given-New	Enhancement	Procedural understanding; Learner as observer	4.3.2	5.4	A; B1; B3; B4; C
MT-P170	Pearson (2019), p.170	Conceptual-analytical Offer; long shot Centre-Margin	Extension	Canonical biomedical perspective; Learner as observer	4.4.2	5.6	A; B1; B3; B4; C
MT-P205	Pearson (2019), p.205	Conceptual- classificatory Offer; long shot Centre- Margin	Elaboration	Terminological decoding; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P231	Pearson (2019), p.231	Conceptual-analytical Offer; long shot Centre-Margin	Elaboration	Conceptual classification; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P273	Pearson (2019), p.273	Conceptual-analytical Offer; long shot Ideal-Real	Elaboration	Terminological decoding; Learner as observer	4.5.2	5.6	A; B1; B3; B4; C

MT-P279	Pearson (2019), p.279	Conceptual-analytical Offer; long shot Centre-Margin	Elaboration	Conceptual classification; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P290	Pearson (2019), p.290	Narrative-actional Offer; long shot Given-New	Enhancement	Procedural understanding; Learner as observer	4.3.2	5.4	A; B1; B3; B4; C
MT-P316	Pearson (2019), p.316	Conceptual- classificatory Offer; long shot Centre- Margin	Elaboration	Terminological decoding; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P330	Pearson (2019), p.330	Conceptual-analytical Offer; long shot Ideal-Real	Elaboration	Conceptual classification; Learner as observer	4.5.2	5.6	A; B1; B3; B4; C
MT-P349	Pearson (2019), p.349	Conceptual-analytical Offer; long shot Centre-Margin	Extension	Canonical biomedical perspective; Learner as observer	4.4.1	5.6	A; B1; B3; B4; C
MT-P355	Pearson (2019), p.355	Narrative-actional Offer; long shot Given-New	Enhancement	Procedural understanding; Learner as observer	4.3.2	5.4	A; B1; B3; B4; C

MT-P401	Pearson (2019), p.401	Conceptual- classificatory Offer; long shot Centre- Margin	Elaboration	Terminological decoding; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P404	Pearson (2019), p.404	Conceptual-analytical Offer; long shot Centre-Margin	Elaboration	Conceptual classification; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P431	Pearson (2019), p.431	Conceptual- classificatory Offer; long shot Centre- Margin	Elaboration	Terminological decoding; Learner as observer	4.3.1	5.3	A; B1; B3; B4; C
MT-P476	Pearson (2019), p.476	Narrative-actional Offer; long shot Given-New	Enhancement	Procedural understanding; Learner as observer	4.3.2	5.4	A; B1; B3; B4; C
MT-P496	Pearson (2019), p.496	Narrative-actional Offer; long shot Given-New	Enhancement	Procedural understanding; Learner as observer	4.3.2	5.4	A; B1; B3; B4; C

B. CAREER PATHS: MEDICAL

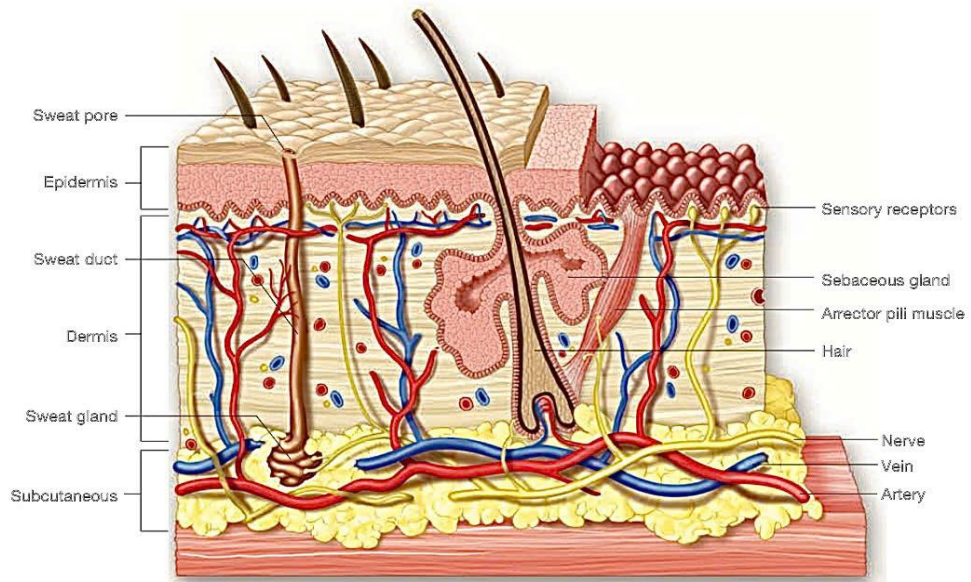
Table D.2 Career Paths: Medical (Evans, Dooley & Tran, 2012)

Page Ensemble ID	Textbook Location	Visual Profile (IMAF-S1)	Intersemiotic Relation (IMAF-S2)	Pedagogic Meaning (IMAF-S3)	Chapter IV	Chapter V	Linked Appendices
CPM-B1P9	Book 1, p.9	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B1P22	Book 1, p.22	Narrative-actional Offer; medium shot Given-New	Enhancement	Procedural understanding; Learner as novice participant	4.3.2	5.4	A; B1; B3; B5; C
CPM-B1P27	Book 1, p.27	Conceptual-analytical Offer; medium shot Ideal-Real	Elaboration	Terminological decoding; Learner as observer	4.5.1	5.3	A; B1; B3; B5; C
CPM-B2P9	Book 2, p.9	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C

CPM-B2P18	Book 2, p.18	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B2P21	Book 2, p.21	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B2P22	Book 2, p.22	Narrative-actional Offer; medium shot Given-New	Enhancement	Procedural understanding; Learner as novice participant	4.3.2	5.4	A; B1; B3; B5; C
CPM-B2P26	Book 2, p.26	Conceptual-analytical Offer; medium shot Centre-Margin	Elaboration	Terminological decoding; Learner as observer	4.3.1	5.3	A; B1; B3; B5; C
CPM-B2P28	Book 2, p.28	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B2P32	Book 2, p.32	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C

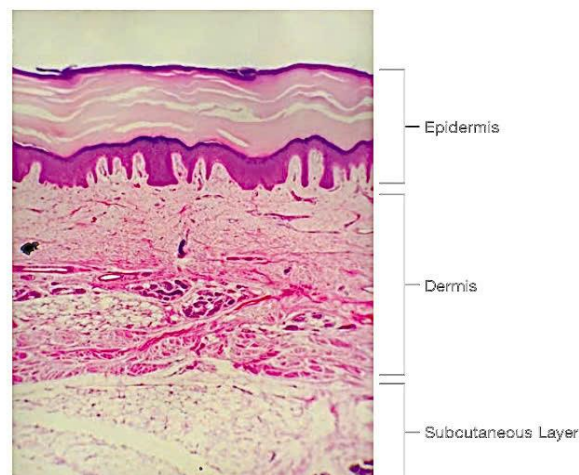
CPM-B2P33	Book 2, p.33	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B3P7	Book 3, p.7	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B3P8	Book 3, p.8	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B3P9	Book 3, p.9	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B3P10	Book 3, p.10	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B3P22	Book 3, p.22	Narrative-actional Offer; medium shot Given-New	Enhancement	Procedural understanding; Learner as novice participant	4.3.2	5.4	A; B1; B3; B5; C

CPM-B3P24	Book 3, p.24	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B3P25	Book 3, p.25	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B3P27	Book 3, p.27	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C
CPM-B3P28	Book 3, p.28	Narrative-reactional Offer; medium shot Given-New	Projection	Professional role positioning; Learner as novice participant	4.3.3	5.5	A; B1; B3; B5; C

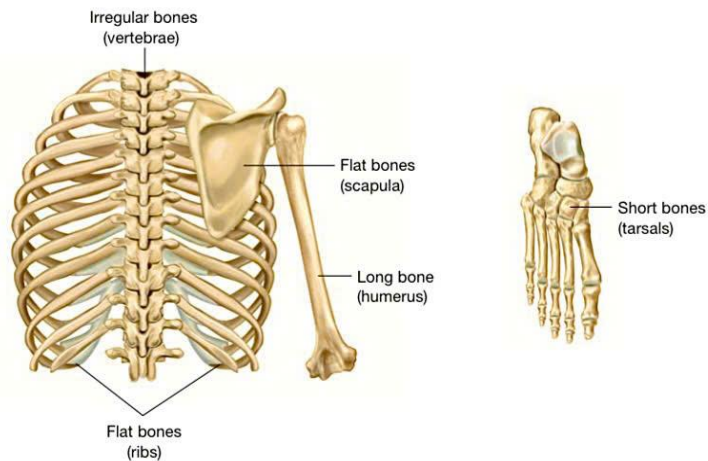


■ **Figure 3-1** Skin structure, including the layers of the skin, the subcutaneous layer, and the accessory organs: sweat gland, sebaceous gland, and hair.

The epidermis is composed of **stratified squamous epithelium** (see Figure 3-2 ■). This type of epithelial tissue consists of flat, scale-like cells arranged in overlapping layers or strata. The epidermis does not have a blood supply or any connective tissue, so it is dependent for nourishment on the deeper layers of skin.



■ **Figure 3-2** Photomicrograph showing the two layers of the skin and the subcutaneous layer. (Jubal Harshaw/Shutterstock)



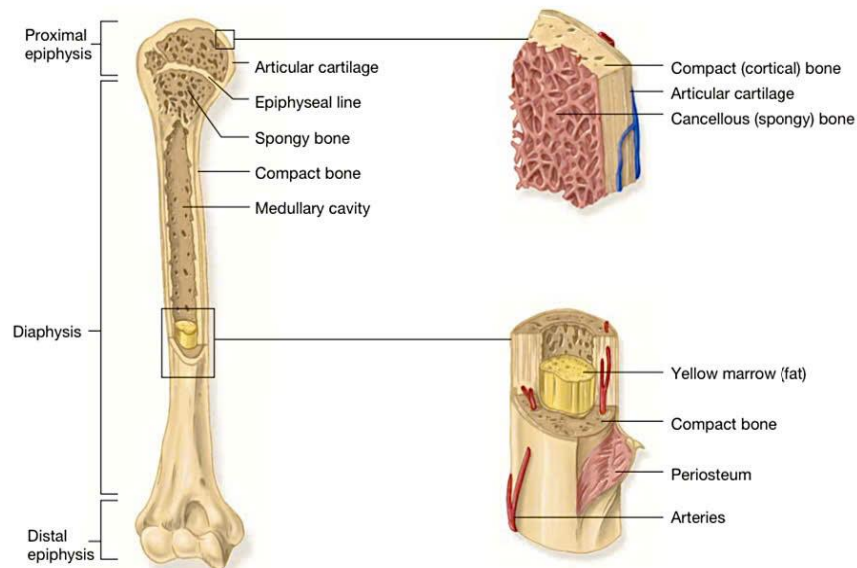
■ **Figure 4-1** Classification of bones by shape.

nerves, and lymphatic vessels. The dense and hard exterior surface bone is called **cortical** or **compact bone**. **Cancellous** or **spongy bone** is found inside the bone. As its name indicates, spongy bone has spaces in it, giving it a spongelike appearance. These spaces contain **red bone marrow**, which manufactures most of the blood cells and is found in some parts of all bones.

The center of the diaphysis contains an open canal called the **medullary cavity**. Early in life, this cavity also contains red bone marrow, but as a person ages, the red bone marrow of the medullary cavity gradually converts to **yellow bone marrow**, which consists primarily of fat cells. Figure 4-2 ■ contains an illustration of the structure of long bones.

Med Term Tip

The term *diaphysis* comes from the Greek term meaning *to grow between*.



■ **Figure 4-2** Components of a long bone. The entire long bone is on the left side, accompanied by a blow-up of the proximal epiphysis and a section of the diaphysis.

Anatomy and Physiology of the Muscular System

Med Term Tip

The term *muscle* is the diminutive form of the Latin word *mus* or "little mouse." This is thought to describe how the skin ripples when a muscle contracts, like a little mouse running.

What's In A Name?

Look for these word parts:
cardi/o = heart
-ac = pertaining to
in- = not

muscle fibers

muscles

Muscles are bundles of parallel **muscle fibers**. As these fibers contract (shorten in length) they produce movement of or within the body. The movement may take the form of bringing two bones closer together, pushing food through the digestive system, or pumping blood through blood vessels. In addition to producing movement, muscles also hold the body erect and generate heat.

Types of Muscles

cardiac muscle

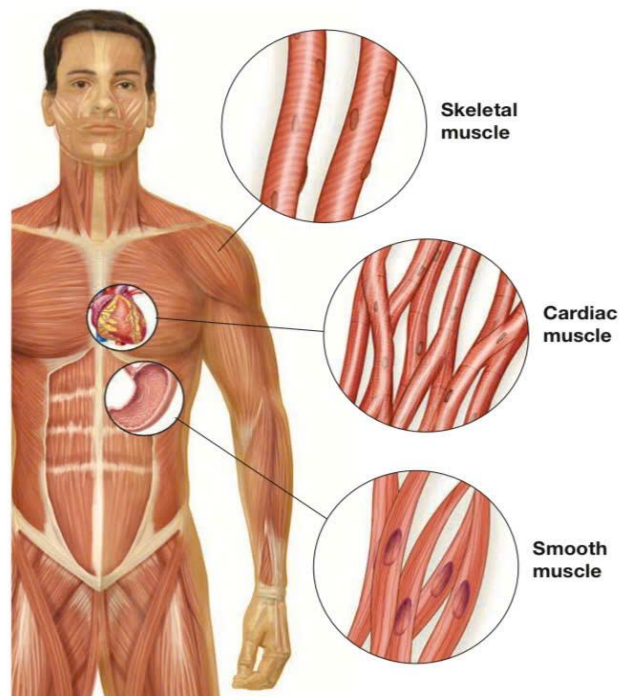
smooth muscle

involuntary muscles

voluntary muscles

skeletal muscle

The three types of muscle tissue are **skeletal muscle**, **smooth muscle**, and **cardiac muscle** (see Figure 4-23 ■). Muscle tissue may be either voluntary or involuntary. **Voluntary muscles** are those muscles for which a person consciously chooses to contract and for how long and how hard to contract them. The skeletal muscles of the arm and leg are examples of this type of muscle. **Involuntary muscles** are the muscles under the control of the subconscious regions of the brain. The smooth muscles found in internal organs and cardiac muscles are examples of involuntary muscle tissue.



■ **Figure 4-23** The three types of muscles: skeletal, smooth, and cardiac.

Anatomy and Physiology of the Cardiovascular System

arteries
blood vessels
capillaries
carbon dioxide
circulatory system
deoxygenated (dee-OK-sih-jen-ay-ted)
heart

oxygen
oxygenated (OK-sih-jen-ay-ted)
pulmonary circulation (PULL-mon-air-ee / ser-kyoo-LAY-shun)
systemic circulation (sis-TEM-ik / ser-kyoo-LAY-shun)
veins

What's In A Name?

Look for these word parts:

ox/o = oxygen

pulmon/o = lung

system/o = system

-ary = pertaining to

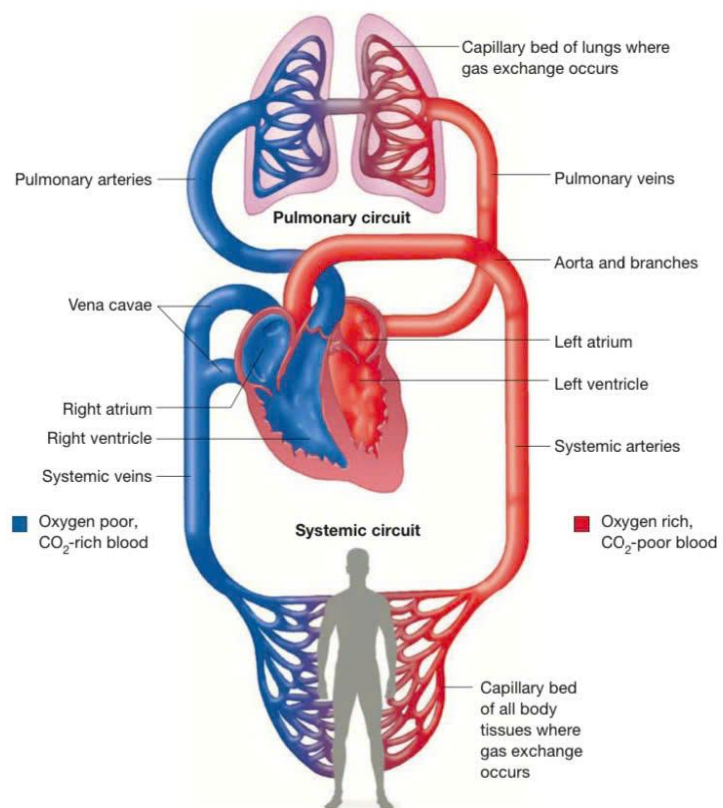
-ic = pertaining to

de- = without

di- = two

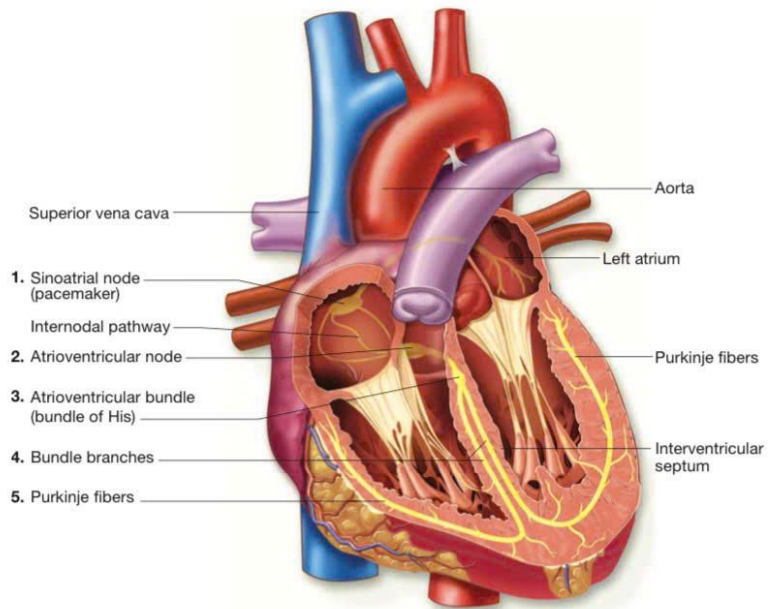
The cardiovascular (CV) system, also called the **circulatory system**, maintains the distribution of blood throughout the body and is composed of the **heart** and the **blood vessels**—**arteries**, **capillaries**, and **veins**.

The circulatory system is composed of two parts: the **pulmonary circulation** and the **systemic circulation**. The pulmonary circulation, between the heart and lungs, transports **deoxygenated** blood to the lungs to get oxygen, and then back to the heart. The systemic circulation carries **oxygenated** blood away from the heart to the tissues and cells, and then back to the heart (see Figure 5-1 ■). In this way, all the body's cells receive blood and oxygen.



■ **Figure 5-1** A schematic of the circulatory system illustrating the pulmonary circulation picking up oxygen from the lungs and the systemic circulation delivering oxygen to the body.

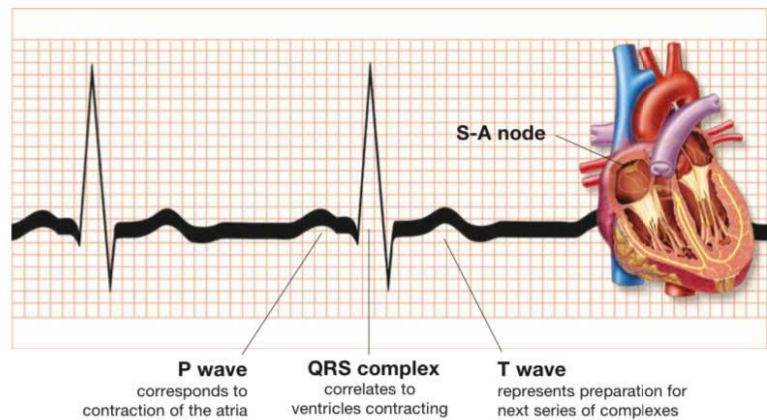
■ **Figure 5-6** The conduction system of the heart; traces the path of the electrical impulse that stimulates the heart chambers to contract in the correct sequence.



Med Term Tip

The electrocardiogram, referred to as an EKG or ECG, is a measurement of the electrical activity of the heart (see Figure 5-7 ■). This can give the physician information about the health of the heart, especially the myocardium.

■ **Figure 5-7** An electrocardiogram (EKG or ECG) wave record of the electrical signal as it moves through the conduction system of the heart. This signal stimulates the chambers of the heart to contract and relax in the proper sequence.



PRACTICE AS YOU GO

A. Complete the Statement

1. The study of the heart is called _____.
2. The three layers of the heart are _____, _____, and _____.
3. The impulse for the heartbeat (the pacemaker) originates in the _____.

Diagnostic Procedures (continued)

Term	Word Parts	Definition
Holter monitor		Portable ECG monitor worn by patient for a period of a few hours to a few days to assess heart and pulse activity as person goes through activities of daily living; used to assess patient who experiences chest pain and unusual heart activity during exercise and normal activities
stress testing		Method for evaluating cardiovascular fitness; patient is placed on treadmill or bicycle and then subjected to steadily increasing levels of work; EKG and oxygen levels are taken while patient exercises; test is stopped if abnormalities occur on EKG; also called <i>exercise test</i> or <i>treadmill test</i>



■ **Figure 5-18** Man undergoing a stress test on a treadmill while physician monitors his condition. (Serafino Mozzo/Shutterstock)

Therapeutic Procedures

Term	Word Parts	Definition
Medical Procedures		
cardiopulmonary resuscitation (CPR) (kar-dee-oh-PULL-mon-air-ee / ree-suss-ih-TAY-shun)	cardi/o = heart pulmon/o = lung -ary = pertaining to	Procedure to restore cardiac output and oxygenated air to lungs for person in cardiac arrest; combination of chest compressions (to push blood out of heart) and artificial respiration (to blow air into lungs) is performed by one or two CPR-trained rescuers
defibrillation (dee-fib-rih-LAY-shun)	de- = without	Procedure that converts serious irregular heartbeats, such as fibrillation, by giving electric shocks to heart using instrument called defibrillator; also called <i>cardioversion</i> ; automated external defibrillators (AEDs) are portable devices that automatically detect life-threatening arrhythmias and deliver appropriate electrical shock; designed to be used by nonmedical personnel and found in public places such as shopping malls and schools



■ **Figure 5-19** An emergency medical technician positions defibrillator paddles on the chest of a supine male patient. (Floyd Jackson/Pearson Education, Inc.)

Tonsils

adenoids (AD-eh-noydz)

lingual tonsils (LING-gwal)

palatine tonsils (PAL-ah-tyne)

pharyngeal tonsils (fair-IN-jee-al)

pharynx (FAIR-inks)

The tonsils are collections of lymphatic tissue located on each side of the throat or **pharynx** (see Figure 6-11 ■). There are three sets of tonsils: **palatine tonsils**, **pharyngeal tonsils** (commonly referred to as the **adenoids**), and **lingual tonsils**. All tonsils contain a large number of leukocytes and act as filters to protect the body from the invasion of pathogens through the digestive or respiratory systems. Tonsils are not vital organs and can safely be removed if they become a continuous site of infection.



■ **Figure 6-11** The shape of a tonsil.

What's In A Name?

Look for these word parts:

lingu/o = tongue
palat/o = palate
pharyng/o = pharynx
-al = pertaining to
-eal = pertaining to
-ine = pertaining to

Spleen

blood sinuses

macrophages (MAK-roh-fay-jez)

The spleen, located in the upper left quadrant of the abdomen, consists of lymphatic tissue that is highly infiltrated with blood vessels (see Figure 6-12 ■). These vessels spread out into slow-moving **blood sinuses**. The spleen filters out and destroys old red blood cells, recycles the iron, and also stores some of the blood supply for the body. Phagocytic **macrophages** line the blood sinuses in the spleen to engulf and remove pathogens. Because the blood is moving through the organ slowly, the macrophages have time to carefully identify pathogens and worn-out red blood cells. The spleen is also not a vital organ and can be removed due to injury or disease. However, without the spleen, a person's susceptibility to a bloodstream infection may be increased.



■ **Figure 6-12** The shape of the spleen.

What's In A Name?

Look for these word parts:

macro- = large
-phage = to eat

Thymus Gland

T cells

thymosin (THIGH-moh-sin)

T lymphocytes

The thymus gland, located in the upper portion of the mediastinum, is essential for the proper development of the immune system (see Figure 6-13 ■). It assists the body with the immune function and the development of antibodies. This organ's hormone, **thymosin**, changes lymphocytes to **T lymphocytes** (simply called **T cells**), which play an important role in the immune response. The thymus is active in the unborn child and throughout childhood until adolescence, when it begins to shrink in size.



■ **Figure 6-13** The shape of the thymus gland.

What's In A Name?

Look for this word part:

-ous = pertaining to

Immunity

acquired immunity

active acquired immunity

bacteria (bak-TEE-ree-ah)

cancerous tumors

fungi (FUN-jeye)

immune response

immunity (im-YOO-nih-tee)

immunizations (im-yoo-nih-ZAY-shuns)

natural immunity

passive acquired immunity

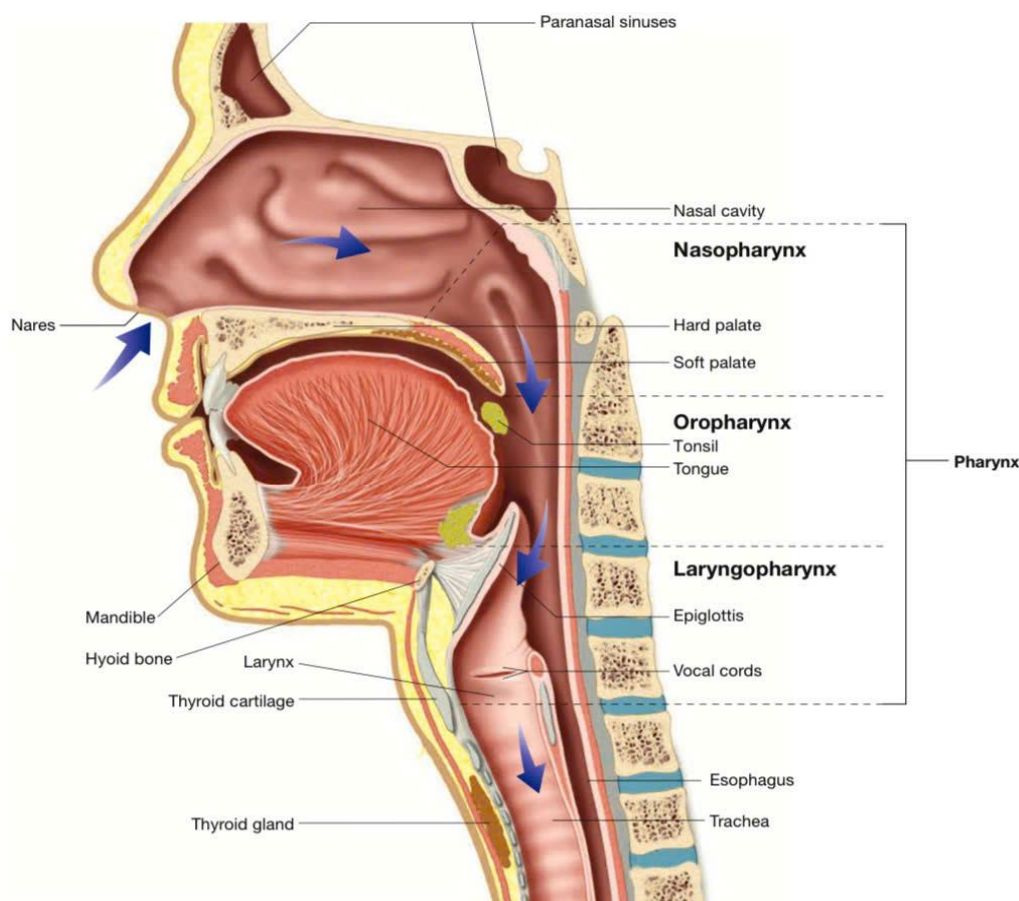
protozoans (proh-toh-ZOH-anz)

toxins

vaccinations (vak-sih-NAY-shuns)

viruses

Immunity is the body's ability to defend itself against pathogens, such as **bacteria**, **viruses**, **fungi**, **protozoans**, **toxins**, and **cancerous tumors**. Immunity comes in two forms: **natural immunity** and **acquired immunity**. Natural immunity, also called **innate immunity**, is not specific to a particular disease and does not require prior exposure to the pathogenic agent. A good example of natural immunity is the



■ **Figure 7-1** Sagittal section of upper respiratory system illustrating the internal anatomy of the nasal cavity, pharynx, larynx, and trachea.

and filter out large dirt particles before they can enter the lungs. Capillaries in the mucous membranes warm inhaled air as it passes through the airways. Additionally, several **paranasal sinuses**, or air-filled cavities, are located within the facial bones. The sinuses act as an echo chamber during sound production and give resonance to the voice.

Pharynx

adenoids (AD-eh-noydz)

auditory tube

eustachian tube (yoo-STAY-shee-en)

laryngopharynx (lah-ring-goh-FAIR-inks)

lingual tonsils (LING-gwal)

nasopharynx (nay-zoh-FAIR-inks)

oropharynx (or-oh-FAIR-inks)

palatine tonsils (PAL-ah-tyne)

pharyngeal tonsils (fair-IN-jee-al)

Air next enters the pharynx, also referred to as the *throat*, which is used by both the respiratory and digestive systems. At the end of the pharynx, air enters the trachea while food and liquids are shunted into the esophagus.

Word Watch

The term *cilia* means *hair*, and there are other body systems that have cilia or cilia-like processes. For example, when discussing the eye, *cilia* means *eyelashes*.

What's In A Name?

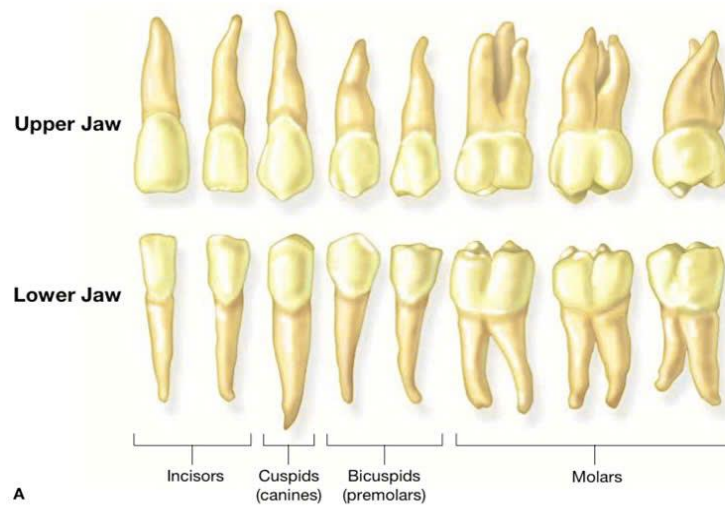
Look for these word parts:

audit/o = hearing

lingu/o = tongue

-al = pertaining to

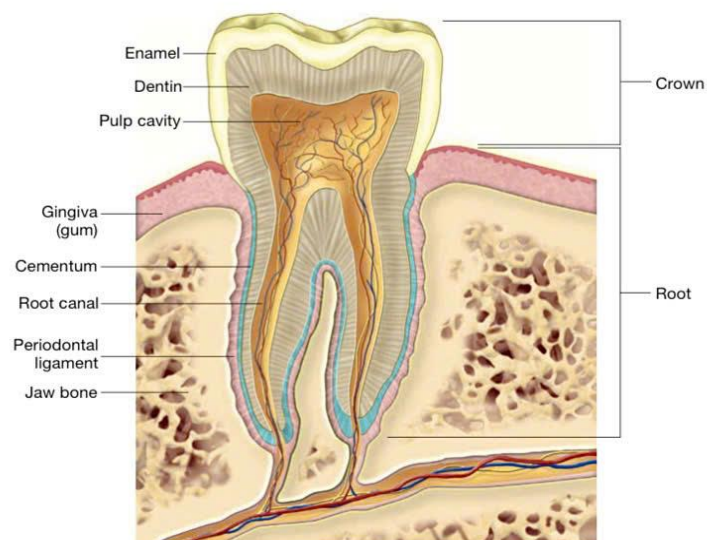
-ory = pertaining to



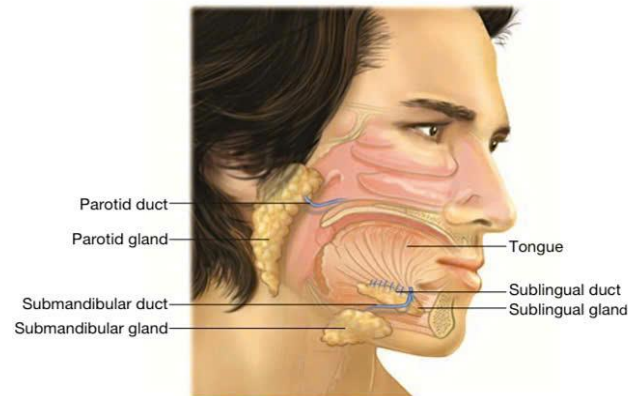
A



B



■ **Figure 8-4** An adult tooth, longitudinal view showing internal structures of the crown and root.



■ **Figure 8-9** The salivary glands: parotid, sublingual, and submandibular. This image shows the position of each gland and its duct emptying into the oral cavity.

Liver

bile (BYE-al)

emulsification (ee-mull-sih-fih-KAY-shun)

The liver, a large organ located in the right upper quadrant of the abdomen, has several functions including processing the nutrients absorbed by the intestines, detoxifying harmful substances in the body, and producing **bile** (see Figure 8-10 ■). Bile is important for the digestion of fats and lipids because it breaks up large fat globules into much smaller droplets, making them easier to digest in the watery environment inside the intestines. This process is called **emulsification**.

Med Term Tip

The liver weighs about four pounds and has so many important functions that people cannot live without it. It has become a major transplant organ. The liver is also able to regenerate itself. You can lose more than half of your liver, and it will regrow.

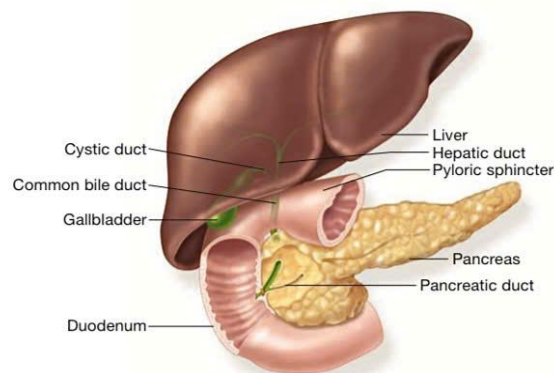
Gallbladder

common bile duct

cystic duct (SIS-tik)

hepatic duct (heh-PAT-ik)

Bile produced by the liver is stored in the gallbladder (GB). As the liver produces bile, it travels down the **hepatic duct** and up the **cystic duct** into the gallbladder (see again Figure 8-10). In response to the presence of fat in the chyme, the

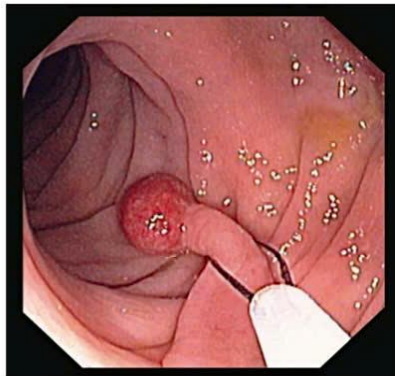


■ **Figure 8-10** The liver, gallbladder, and pancreas. Image shows the relationship of these three organs and their ducts to the duodenum.

Pathology (continued)

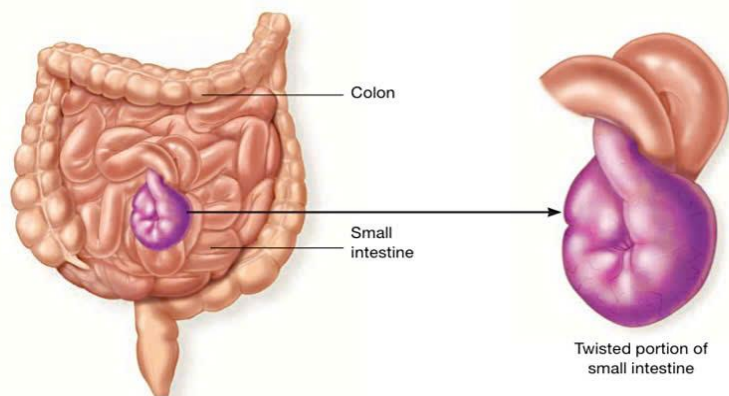
Term	Word Parts	Definition
polyposis (pol-ee-POH-sis)	polyp/o = polyp -osis = abnormal condition	Presence of small tumors, called polyps , containing a pedicle or stemlike attachment in mucous membranes of the large intestine (colon); may be precancerous

■ **Figure 8-16** Endoscopic view of a polyp in the colon. Note the mushroom-like shape, an enlarged top growing at the end of a stem. It is being removed by means of a wire loop slipped over the polyp and then tightened to cut it off. (David M. Martin, M.D./Science Source)

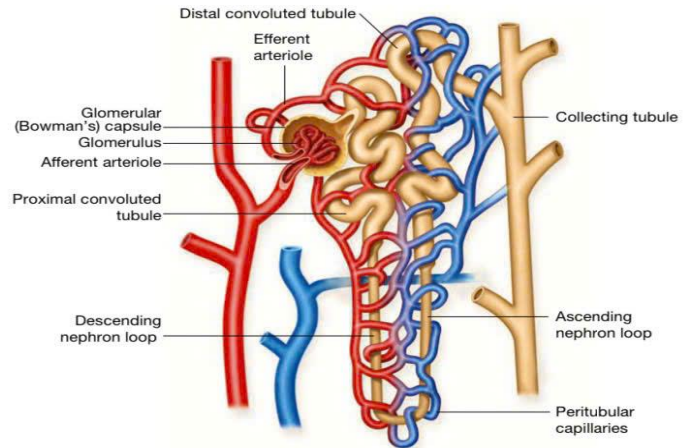


proctoptosis (prok-top-TOH-sis)	proct/o = rectum and anus -ptosis = drooping	Prolapsed or drooping rectum and anus
ulcerative colitis (UL-ser-ah-tiv / koh-LYE-tis)	col/o = colon -itis = inflammation	Chronic inflammatory condition resulting in numerous ulcers formed on mucous membrane lining of the colon; cause is unknown; also known as <i>inflammatory bowel disease</i> (IBD)
volvulus (VOL-vyoo-lus)		Condition in which the bowel twists upon itself, causing an obstruction; painful and requires immediate surgery

■ **Figure 8-17** Volvulus. A length of small intestine has twisted around itself, cutting off blood circulation to the twisted loop.



■ **Figure 9-3** The structure of a nephron, illustrating the nephron structure in relation to the circulatory system.



What's In A Name?

Look for these word parts:

ex- = outward

in- = inward

-al = pertaining to

Word Watch

The terms *ureter* and *urethra* are frequently confused. Remember that there are two ureters carrying urine from the kidneys into the bladder. There is only one urethra, and it carries urine from the bladder to the outside of the body.

Med Term Tip

Terms such as *micturition*, *voiding*, and *urination* all mean basically the same thing—the process of releasing urine from the body.

Urinary Bladder

external sphincter (SFINGK-ter)

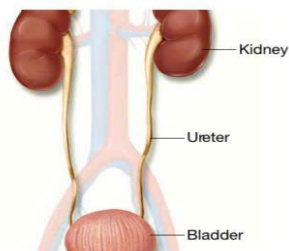
internal sphincter

rugae (ROO-jee)

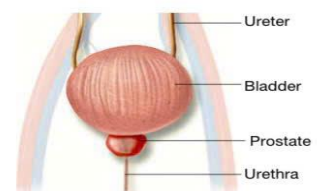
urination

The urinary bladder is an elastic muscular sac that lies in the base of the pelvis just behind the pubic symphysis (see Figure 9-5 ■). It is composed of three layers of smooth muscle tissue lined with mucous membrane containing **rugae**, or folds, that allow it to stretch. The bladder receives the urine directly from the ureters, stores it, and excretes it by **urination** through the urethra.

Generally, an adult bladder signals the urge to void (or empty the bladder) when it contains 300–400 mL of urine. Involuntary muscle action causes the bladder to contract and the **internal sphincter** to relax. The internal sphincter prevents the bladder from emptying at the wrong time. Voluntary action controls

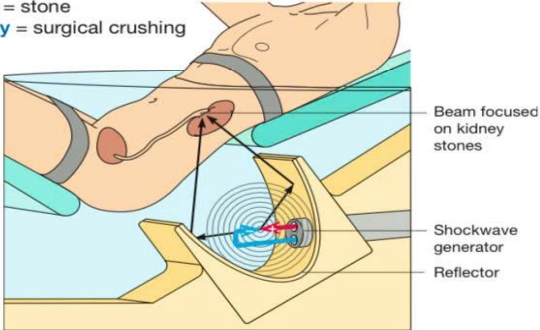


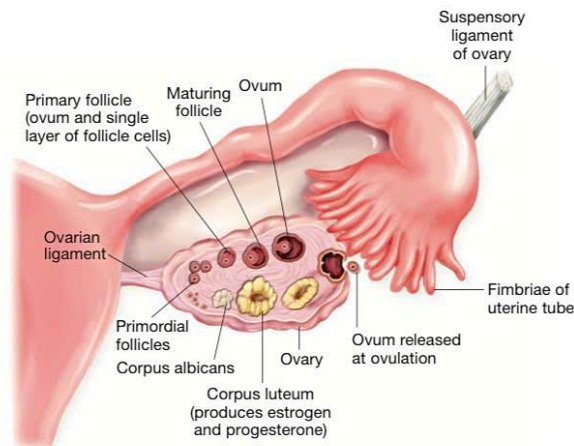
■ **Figure 9-4** The ureters extend from the kidneys to the urinary bladder.



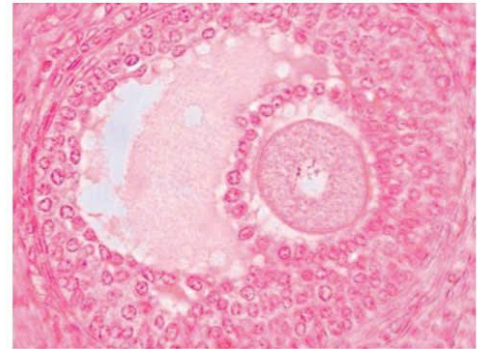
■ **Figure 9-5** The structure of the urinary bladder. (Note the prostate gland.)

Therapeutic Procedures

Term	Word Parts	Definition
Medical Treatments		
catheter (KATH-eh-ter)		Flexible tube inserted into body for purpose of moving fluids into or out of body; most commonly refers to tube threaded through the urethra into the bladder to withdraw urine (see again Figure 9-9)
catheterization (cath) (kath-eh-ter-ih-ZAY-shun)		Insertion of tube through the urethra and into the urinary bladder for purpose of withdrawing urine or inserting dye
extracorporeal shockwave lithotripsy (ESWL) (eks-trah-kor-POR-ee-al / shockwave / LITH-oh-trip-see)	extra- = outside of corpor/o = body -eal = pertaining to lith/o = stone -tripsy = surgical crushing	Use of ultrasound waves from outside the body to break up stones; process does not require invasive surgery
■ Figure 9-12 Extracorporeal shockwave lithotripsy, a noninvasive procedure using high-frequency sound waves to shatter kidney stones.		
hemodialysis (HD) (hee-moh-dye-AL-ih-sis)	hem/o = blood	Use of artificial kidney machine that filters the blood of a person to remove waste products; use of this technique in patients who have defective kidneys is lifesaving
■ Figure 9-13 Patient undergoing hemodialysis. Patient's blood passes through hemodialysis machine for cleansing and is then returned to the body. (Gopixa/Shutterstock)		



■ **Figure 10-2** Structure of the ovary and uterine (fallopian) tube. Figure illustrates stages of ovum development and the relationship of the ovary to the uterine tube.



■ **Figure 10-3** Photomicrograph of human ovary showing ovum in its follicle prior to ovulation. (Anna Jurkowska/Shutterstock)

follicle-stimulating hormone (FSH) and **luteinizing hormone (LH)**, stimulate maturation of an ovum and trigger **ovulation**, the process by which one ovary releases an ovum (or **oocyte**) (see Figure 10-3 ■). The principal female sex hormones produced by the ovaries, **estrogen** and **progesterone**, stimulate the lining of the uterus to be prepared to receive a fertilized ovum. These hormones are also responsible for the female secondary sexual characteristics.

Uterine Tubes

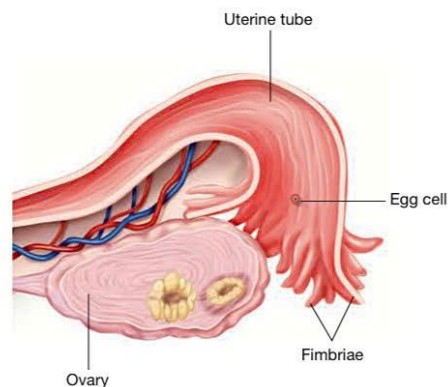
conception (kon-SEP-shun)

fallopian tubes (fah-LOH-pee-an)

fimbriae (FIM-bree-ee)

oviducts (OH-vih-dukts)

The uterine tubes, also called the **fallopian tubes** or **oviducts**, are approximately 4 inches (10 cm) long and run from the area around each ovary to either side of the upper portion of the uterus (see Figure 10-4 ■ and Figure 10-5 ■). As they near the ovaries, the unattached ends of these two tubes expand into finger-like projections called **fimbriae**. The fimbriae catch an ovum after ovulation and direct it into



What's In A Name?

Look for these word parts:

estr/o = female

o/o = egg

ov/o = ovum

-cyte = cell

-gen = that which produces

pro- = before

What's In A Name?

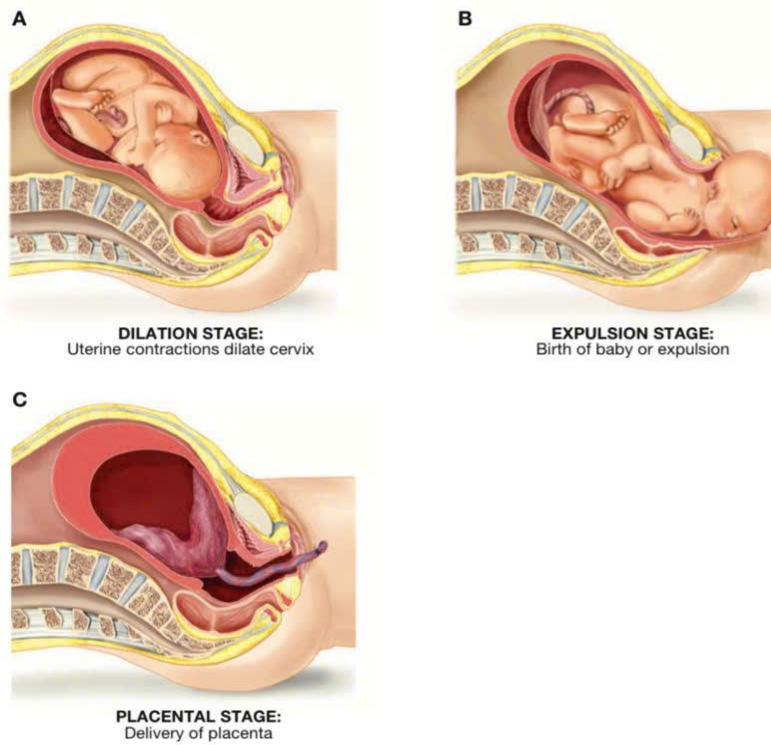
Look for this word part:

ov/i = ovum

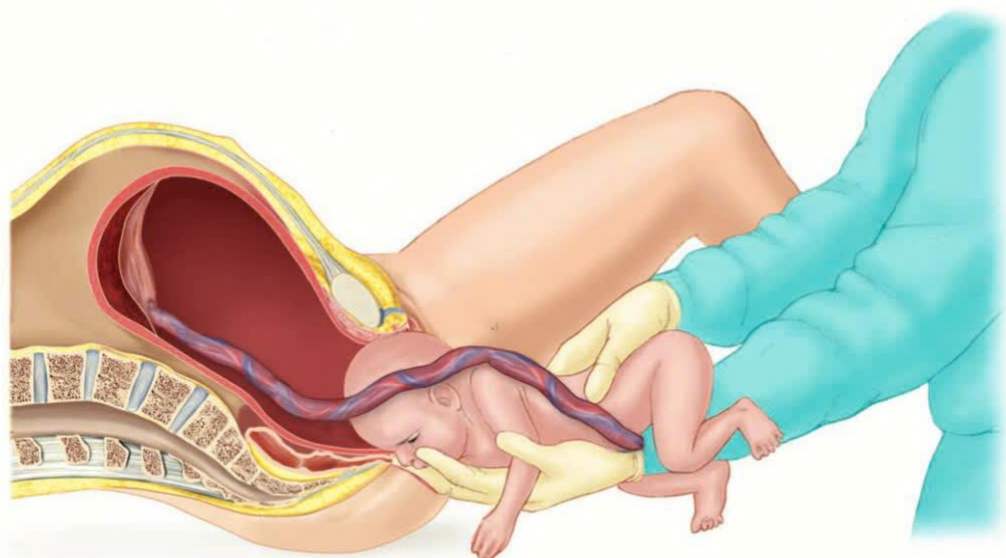
Med Term Tip

When the fertilized egg adheres or implants to the uterine tube instead of moving into the uterus, a condition called **tubal pregnancy** exists. There is not enough room in the uterine tube for the fetus to grow normally. Implantation of the fertilized egg in any location other than the uterus is called an **ectopic pregnancy**. *Ectopic* is a general term meaning *in the wrong place*.

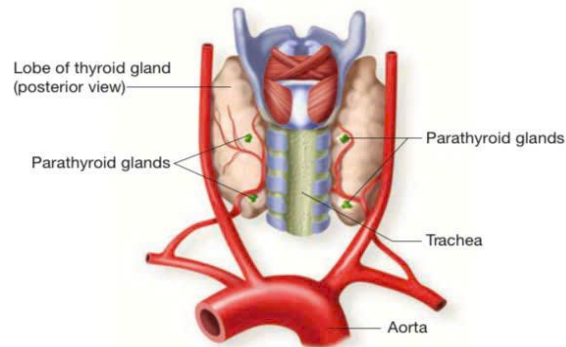
■ **Figure 10-4** Uterine (fallopian) tube, showing released ovum within the uterine tube.



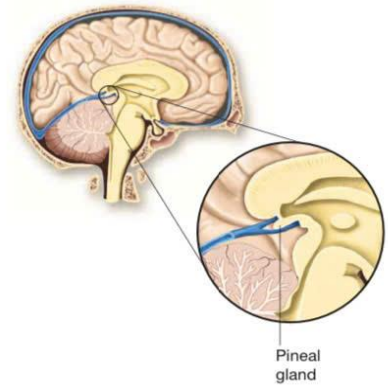
■ **Figure 10-12** The stages of labor and delivery. A) During the dilation stage the cervix thins and dilates to 10 cm. B) During the expulsion stage the infant is delivered. C) During the placental stage the placenta is delivered.



■ **Figure 10-13** A breech birth. This image illustrates a newborn that has been delivered buttocks first.



■ **Figure 11-4** The parathyroid glands. These four glands are located on the posterior side of the thyroid gland. They secrete parathyroid hormone.



■ **Figure 11-5** The pineal gland is a part of the thalamus region of the brain. It secretes melatonin.

Pituitary Gland

adrenocorticotrophic hormone

(ah-dree-noh-kor-tih-koh-TROH-pik)

anterior lobe

antidiuretic hormone (an-tye-dye-yoo-RET-ik)

follicle-stimulating hormone (FALL-ih-kl / STIM-yoo-lay-ting)

gonadotropins (goh-nad-oh-TROH-pins)

growth hormone

hypothalamus (high-poh-THAL-ah-mus)

luteinizing hormone (LOO-teh-nigh-zing)

melanocyte-stimulating hormone

oxytocin (ok-see-TOH-sin)

posterior lobe

prolactin (proh-LAK-tin)

somatotropin (soh-mat-oh-TROH-pin)

thyroid-stimulating hormone

The pituitary gland is located underneath the brain (see Figure 11-6 ■). The small marble-shaped organ is divided into an **anterior lobe** and a **posterior lobe**.

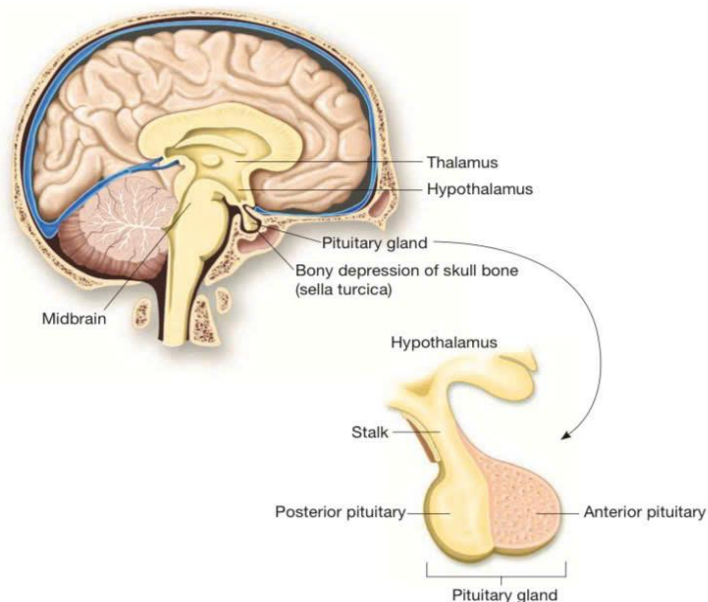
What's In A Name?

Look for these word parts:

somat/o = body

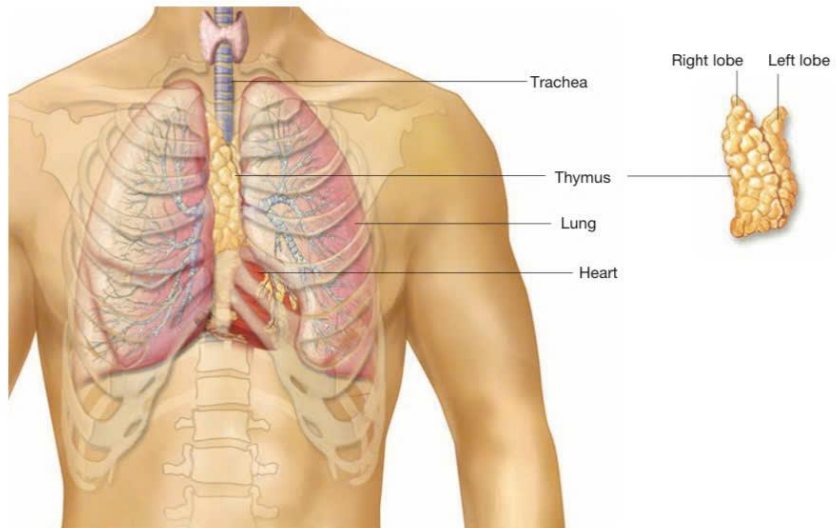
-tropin = to stimulate

hypo- = below



■ **Figure 11-6** The pituitary gland lies just underneath the brain. It is subdivided into anterior and posterior lobes. Each lobe secretes different hormones.

■ **Figure 11-9** The thymus gland. This gland lies in the mediastinum of the thoracic cavity, just above the heart. It secretes thymosin.

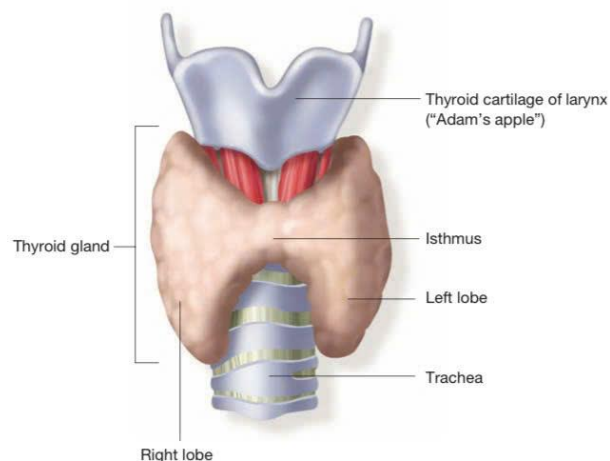


Med Term Tip

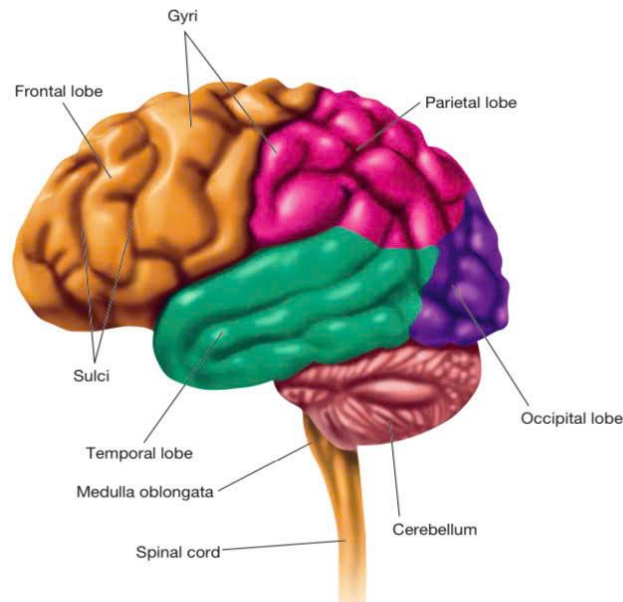
Iodine is found in many foods, including vegetables and sea-food. It is also present in iodized salt, which is one of the best sources of iodine for people living in the Goiter Belt, composed of states located away from saltwater. A lack of iodine in the diet can lead to thyroid disorders, including *goiter*.

thyroid cartilage, or Adam's apple, is located just above the thyroid gland. This gland produces the hormones **thyroxine** (T_4) and **triiodothyronine** (T_3) (see again Table 11-1). These hormones are produced in the thyroid gland using the mineral **iodine**. Thyroxine and triiodothyronine help to regulate the production of energy and heat in the body to adjust the body's metabolic rate. The minimum rate of metabolism necessary to support the function of the body at rest is called the **basal metabolic rate** (BMR).

The thyroid gland also secretes **calcitonin** (CT) in response to hypercalcemia (too high blood calcium level). Its action is the opposite of the parathyroid hormone and stimulates the increased deposition of calcium into bone, thereby lowering blood levels of calcium.



■ **Figure 11-10** The thyroid gland is subdivided into two lobes, one on each side of the trachea.



■ **Figure 12-3** The brain; colors indicate the different cerebral lobes. (Lightspring/Shutterstock)

The final portion of the brain is the brainstem, which has three components: **midbrain**, **pons**, and **medulla oblongata**. The midbrain acts as a pathway for impulses to be conducted between the brain and the spinal cord. The pons—a term meaning *bridge*—connects the cerebellum to the rest of the brain. The medulla oblongata is the most inferior positioned portion of the brain; it connects the brain to the spinal cord. However, this vital area contains the centers that control respiration, heart rate, temperature, and blood pressure. Additionally, this is the site where nerve tracts cross from one side of the brain to control functions and movement on the other side of the body. In other words, with few exceptions, the left side of the brain controls the right side of the body and vice versa.

The brain has four interconnected cavities called **ventricles**: one in each cerebral hemisphere, one in the thalamus, and one in front of the cerebellum. These contain **cerebrospinal fluid** (CSF), which is the watery, clear fluid that provides protection from shock or sudden motion to the brain and spinal cord.

PRACTICE AS YOU GO

A. Complete the Statement

1. The organs of the central nervous system are the _____ and _____.
2. The nerves of the peripheral nervous system are either _____ nerves or _____ nerves.
3. The three basic parts of a neuron are _____, _____, and _____.
4. _____ is a fatty substance that insulates some axons.

Muscles of the Eye

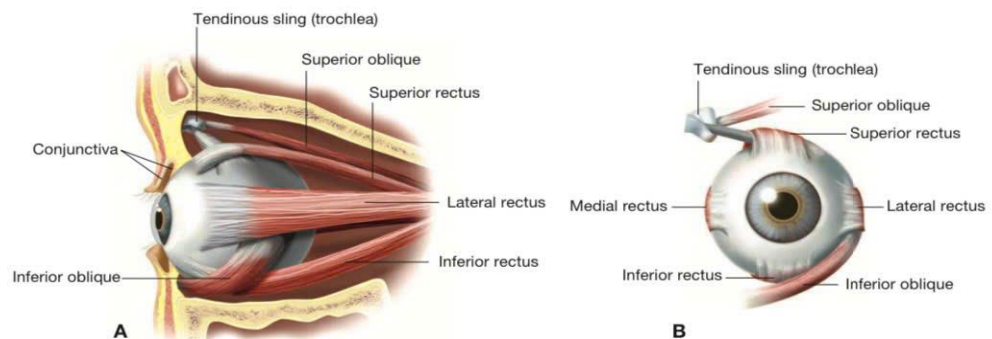
Med Term Tip

Like many other muscles, the names *rectus* and *oblique* provide clues regarding the direction of their fibers, or their *line of pull*. Rectus means *straight* and oblique means *slanted*. Rectus muscles have a straight line of pull. Because the fibers of an oblique muscle are slanted on an angle, they produce rotation.

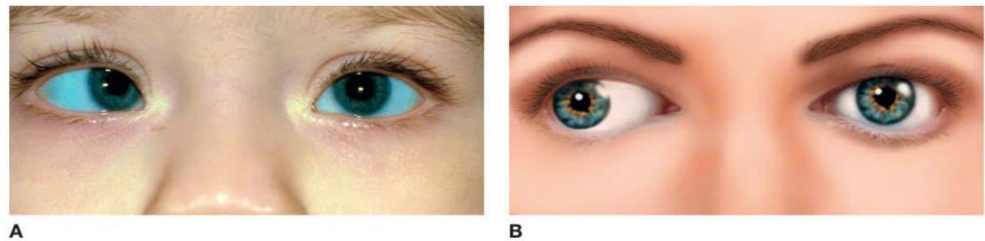
oblique muscles (oh-BLEEK)

rectus muscles (REK-tus)

Six muscles connect the actual eyeball to the skull (see Figure 13-3 ■). These muscles allow for change in the direction of each eye's sightline. In addition, they provide support for the eyeball in the eye socket. Children may be born with a weakness in some of these muscles and may require treatments such as eye exercises or even surgery to correct this problem, commonly referred to as crossed eyes or *strabismus* (see Figure 13-4 ■). The muscles involved are the four **rectus** and two **oblique muscles**. Rectus (meaning *straight*) muscles pull the eye up, down, left, or right in a straight line. Oblique muscles are on an angle and produce diagonal eye movement.



■ **Figure 13-3** The arrangement of the external eye muscles, A) lateral and B) anterior views.



■ **Figure 13-4** Examples of common forms of strabismus. A) Esotropia with the right eye turning inward. (Biophoto Associates/Science Source) B) Exotropia with the right eye turning outward. (Gwen Shockey/Science Source)

The Eyelids

cilia (SIL-ee-ah)

eyelashes

sebaceous glands (sih-BAY-shus)

A pair of eyelids over each eyeball provides protection from foreign particles, injury from the sun and intense light, and trauma (see again Figure 13-1). Both the upper and lower edges of the eyelids have **eyelashes**, or **cilia**, that protect the eye from foreign particles. In addition, **sebaceous glands** located in the eyelids secrete lubricating oil onto the eyeball.

What's In A Name?

Look for these word parts:

seb/o = oil

-ous = pertaining to

What's In A Name?

Look for these word parts:

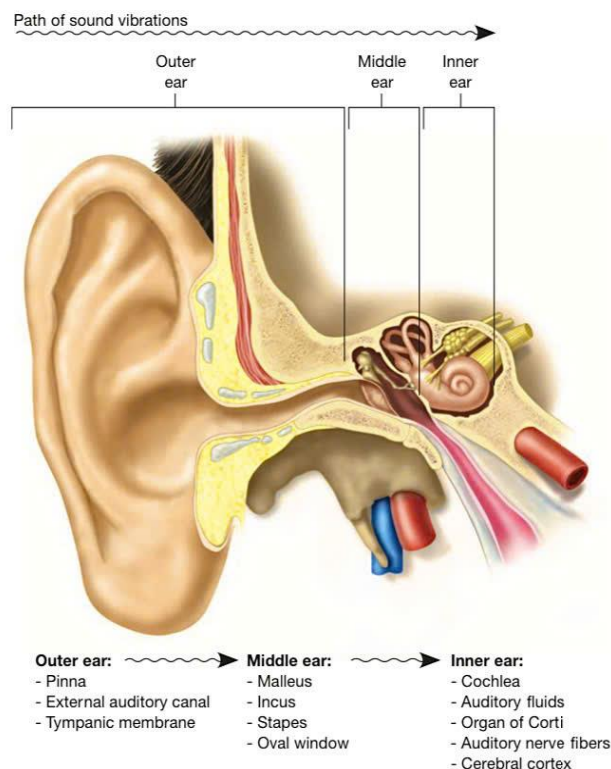
neur/o = nerve**-al** = pertaining to**Med Term Tip**

Hearing impairment is becoming a greater problem for the general population for several reasons. First, people are living longer. Hearing loss can accompany old age, and there are a greater number of people over age 50 requiring hearing assistance. In addition, sound technology has produced music quality that was never available before. However, listening to loud music either naturally or through earphones can cause gradual damage to the hearing mechanism.

How Hearing Works**conductive hearing loss** (kon-DUK-tiv)**sensorineural hearing loss** (sen-soh-ree-NOO-ral)

Figure 13-16 ■ outlines the path of sound through the outer ear and middle ear and into the cochlea of the inner ear. Sound waves traveling down the external auditory canal strike the eardrum, causing it to vibrate. The ossicles conduct these vibrations across the middle ear from the eardrum to the oval window. Oval window movements initiate vibrations in the fluid that fills the cochlea. As the fluid vibrations strike a hair cell, they bend the small hairs and stimulate the nerve ending. The nerve ending then sends an electrical impulse to the brain on the cochlear portion of the vestibulocochlear nerve.

Hearing loss can be divided into two main categories: **conductive hearing loss** and **sensorineural hearing loss**. Conductive refers to disease or malformation of the outer or middle ear. All sound is weaker and muffled in conductive hearing loss since it is not conducted correctly to the inner ear. Sensorineural hearing loss is the result of damage or malformation of the inner ear (cochlea) or the cochlear nerve. In this hearing loss, some sounds are distorted and heard incorrectly. There can also be a combination of both conductive and sensorineural hearing loss.



■ **Figure 13-16** The path of sound waves through the outer, middle, and inner ear.

- 5 Listen and read the supply room poster again. How should syringes and medical waste be disposed?

Listening

- 6 Listen to a conversation between two nurses. Choose the correct answers.

- 1 What are the speakers mainly talking about?
 - A ways to use less supplies
 - B how to organize the storage closet
 - C supplies that need to be replaced
 - D how to properly dispose of syringes
- 2 What will the man likely do next?
 - A empty the sharps containers
 - B check the exam rooms
 - C find new latex gloves
 - D order more syringes

- 7 Listen again and complete the conversation.

Nurse 1: Gary, have you done your supply checks for the day?
 Nurse 2: Yes, I just finished them up.
 Nurse 1: Are we 1 _____ ?
 Nurse 2: Yes, 2 _____ more latex gloves in exam rooms 1 and 3.
 Nurse 1: Okay. 3 _____ ?
 Nurse 2: There are 4 _____ syringes left in the storage closet.
 Nurse 1: Okay, I'll have to order some of those. That reminds me, 5 _____
 _____ empty the sharps containers before you leave.
 Nurse 2: Will do. 6 _____
 _____ Carol.



Speaking

- 8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

We need some ...
Don't forget to ...
That reminds me ...

Student A: You are a nurse. Talk to Student B about:

- what items need to be replaced
- something that should be ordered
- what not to forget to do

Student B: You are a nurse. Talk to Student A about what items need to be replaced.

Writing

- 9 Use the conversation from Task 8 to fill out the supply order form.

Family Clinic

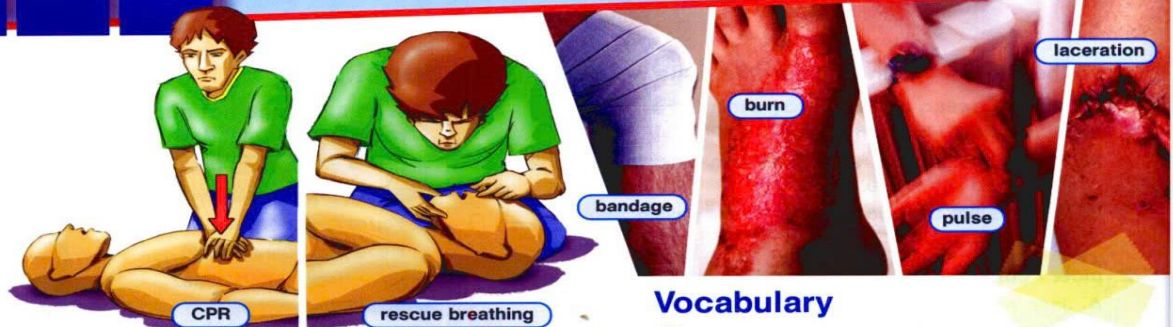
Supply

ORDER FORM

Exam Rooms: _____

Storage Closet: _____

10 First aid



Basic FIRST AID Techniques

In a medical emergency, always call for assistance first. The following are basic steps to take to help an injury victim before medical help arrives.

Assess the victim's condition. Give him or her plenty of space. Prevent **shock** by elevating the legs and covering the victim.

Apply **pressure** to **lacerations** or other open **wounds** to stop bleeding. If bleeding has stopped, wounds should be cleaned and wrapped in a **bandage**. Cool minor **burns** with running water.

If the victim's **pulse** or breathing has stopped, **CPR** or **rescue breathing** can be performed. Be sure to clear the victim's airway before rescue breathing is carried out.

Get ready!

- ① Before you read the passage, talk about these questions.

- 1 What are some injuries that first aid can help?
- 2 What are some common first aid techniques?

Reading

- ② Read the first aid guide. Then, mark the following statements as true (T) or false (F).

- 1 ___ Bleeding wounds should be wrapped with a bandage.
- 2 ___ Use running water to cool burns.
- 3 ___ Rescue breathing can clear a victim's airway.

Vocabulary

- ③ Match the words (1-6) with the definitions (A-F).

- | | |
|-------------|------------------------|
| 1 ___ shock | 4 ___ rescue breathing |
| 2 ___ burn | 5 ___ laceration |
| 3 ___ pulse | 6 ___ CPR |

- A a normal, regular heartbeat that can be felt when an artery is pressed
 B a first aid procedure involving chest compression and artificial breathing
 C an open injury caused by a hard impact to soft body tissue
 D a medical emergency in which the circulatory system cannot provide oxygen to the body
 E a first aid procedure involving providing air for someone who has stopped breathing, but still has a heartbeat
 F a physical injury that can be caused by heat

Vocabulary

- ④ Fill in the blanks with the correct words and phrases from the word bank.

Word BANK

wound first aid bandage
pressure assess

- 1 The doctor will need to _____ the victim's condition.
- 2 I am trained in basic _____ techniques.
- 3 Apply _____ to the injury to stop the bleeding.
- 4 Wash the injury and wrap a(n) _____ around it.
- 5 The nurse used stitches to close the _____.

- 4 Read the sentence pair. Choose where the words best fit the blanks.

1 milliliters / drops

- A Please put 2 _____ of cleaning solution in each ear. Use a dropper.
B Half a liter is equal to 500 _____.

2 grams / ounces

- A There are sixteen _____ in a pound.
B One thousand _____ make up one kilogram.

- 5 Listen and read the medication list for patients again. How are different dosages measured out?

Listening

- 6 Listen to a conversation between a nurse and a head nurse. Mark the following statements as true (T) or false (F).

- 1 ___ The woman thinks a patient's dose is too high.
2 ___ The patient must now take antibiotics.
3 ___ The man thanks the woman for catching an incorrect measurement.

- 7 Listen again and complete the conversation.

Nurse: Hey, Mark. I 1 _____ about Mr. Paul's medication.
Head Nurse: Yes, Nancy? 2 _____?
Nurse: This note calls for three drops in each eye twice a day. That seems 3 _____.
Head Nurse: That's right. They're a little higher than usual. Mr. Paul has a 4 _____.
Nurse: Oh. So that's why there's this new antibiotic here. He'll take fifty 5 _____ once a day by mouth?
Head Nurse: Right. He'll take it until the eye infection 6 _____.
Nurse: All right. That's all I wanted to know.
Head Nurse: Okay. Thanks for checking with me.



Speaking

- 8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

*This note calls for ...
They're a little higher ...
That's why there's ...*

Student A: You are a nurse. Ask Student B about:

- a patient's medication
- why a dose has changed
- new medications

Student B: You are a head nurse. Talk to Student A about a patient's medication.

Writing

- 9 Use the conversation from Task 8 to fill out a patient medication chart.

Name: _____

Medication 1: _____

Dose: _____

Number of Doses per day: _____

Medication 2: _____

Dose: _____

Number of Doses per day: _____

- 4 Read the sentence pair. Choose where the words best fit the blanks.

1 acne / subcutaneous fat

A The hypodermis is composed of _____.

B _____ is a common disease of the skin.

2 epidermis / dermis

A The _____ is the outermost layer of the skin.

B The hypodermis is a layer of skin directly beneath the _____.

- 5 Listen and read the dermatologist's web page on skin conditions again. Which part of the skin does an abrasion affect?

Listening

- 6 Listen to a conversation between a doctor and a patient. Mark the following statements as true (T) or false (F).

- 1 ___ The man's condition is common in teenagers.
- 2 ___ The woman believes the man washes his face too often.
- 3 ___ The woman writes the man a prescription.

- 7 Listen again and complete the conversation.

Doctor: So James, it looks like you have developed
1 _____. Have you heard of it before?

Patient: Yeah, but I don't know much about it. Is it bad?

Doctor: It's not too bad. It's actually
2 _____ in teenagers.

Patient: What causes it?

Doctor: Oils that become trapped in the skin's pores cause it. This is why the 3 _____ form.

Patient: Can we make it go away?

Doctor: The best thing 4 _____
_____ is to wash your face twice a day. Can you do that?

Patient: Yeah. Should I use special
5 _____?

Doctor: Yes. 6 _____
_____ some to take home.

Patient: Is that all?



Speaking

- 8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

What you have is ...

It's quite common ...

The best thing for you to do is ...

Student A: You are a doctor. Talk to Student B about:

- a skin condition
- the severity of the condition
- treatment options/measures

Student B: You are a patient. Talk to Student A about a skin condition.

Writing

- 9 Use the conversation from Task 8 to complete the doctor's notes about a patient's skin condition.

John Saxton
DERMATOLOGIST

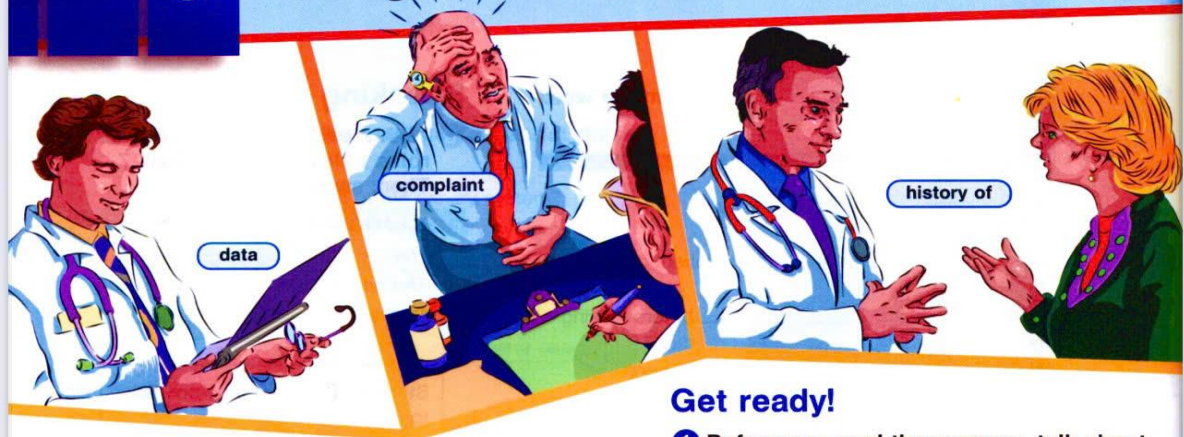
Patient Name: _____

Skin Condition: _____

Treatment: _____

8

Taking a history



family medical history



Assist-A-Med

Guide to Taking Medical Histories

As a medical professional, you should be prepared to take thorough medical histories. The **data** you collect reveals important details about a patient. Doctors use these details to uncover likely causes of medical conditions and establish diagnoses.

First, get as many details as possible about a patient's **complaint**. Find out if the patient has a **history of** the condition. Ask when the patient experienced the **onset** of symptoms. Check what makes it worse and what makes it better.

Then find out about the patient's **past medical history**. Record any previous medical problems, especially those that could relate to the **current** condition.

Other details can also contribute to a useful medical history. A good **social history** includes information about the patient's personal habits and lifestyle. A **family medical history** reveals conditions that are more likely to occur in a particular family.

Before any medical **procedure**, make sure you have an updated record of the patient's **allergies**. This way, you avoid exposing patients to medications or substances that might harm them.

Get ready!

1 Before you read the passage, talk about these questions.

- 1 What kind of information is included in a patient's medical history?
- 2 How do doctors use medical histories?

Reading

2 Read the guide to taking medical histories. Then, complete the table using information from the passage.

Type of History	Information Included
Past Medical History	_____
Social History	_____
Family Medical History	_____

Vocabulary

3 Match the words (1-6) with the definitions (A-F).

- | | |
|-------------|----------------------------|
| 1 — data | 4 — complaint |
| 2 — onset | 5 — procedure |
| 3 — allergy | 6 — family medical history |

- A the first instance of something
 B a record of a patient's family's conditions
 C a collection of information
 D a condition that causes a reaction or illness
 E a medical treatment
 F a pain or illness reported by a patient

4 Read the sentence and choose the correct word.

- 1 The pain in my back **wheezes** / throbs.
- 2 Helen was coughing because of the **muscle ache** / flu.
- 3 According to the thermometer, the patient has a **headache** / fever.
- 4 Fatigue is a **symptom** / cold of the flu.

5 Listen and read the doctor's notes on a patient's symptoms again. What indicates that a patient has the flu?

Listening

6 Listen to a conversation between a doctor and a patient. Mark the following statements as true (T) or false (F).

- 1 — The woman believes she has a cold.
- 2 — The woman has trouble breathing.
- 3 — The woman's fever is severe.

7 Listen again and complete the conversation.

Doctor: What's the trouble today?

Patient: Well, Doctor, I don't 1 _____ at all. I'm sure I have the flu.

Doctor: I'm sorry to hear that. 2 _____ coughing?

Patient: A little. But mostly I'm wheezing. It's 3 _____.

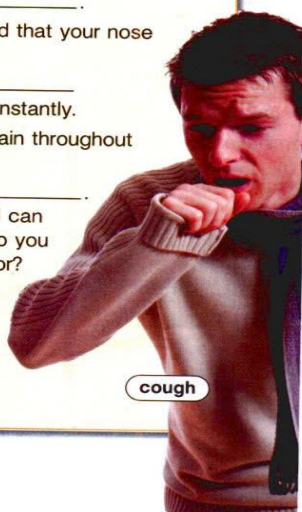
Doctor: And have you noticed that your nose is runny?

Patient: Oh, yes. I've been 4 _____ constantly.

Doctor: Okay. Do you have pain throughout your body?

Patient: I do have 5 _____ . And I'm so tired that I can hardly do anything. Do you think it's the flu, Doctor?

Doctor: Well, you have a pretty 6 _____ .



cough

Speaking

8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

I think I have ...

Have you noticed ...

It's hard to ...

Student A: You are a doctor. Talk to Student B about:

- his or her symptoms
- likely cause of symptoms
- your recommendations

Student B: You are a patient. Talk to Student A about your symptoms.

Writing

9 Use the conversation from Task 8 to complete the flu symptoms checklist.

Does your patient have the flu?

Symptoms:

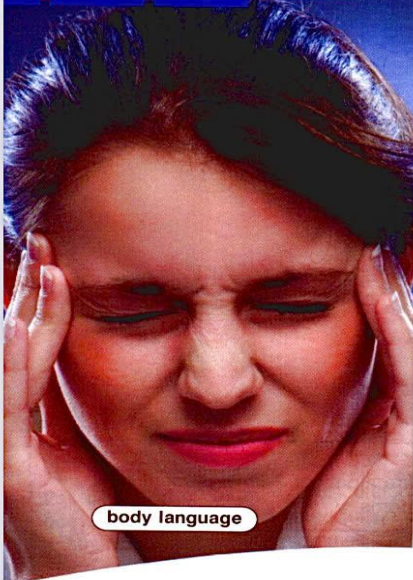
— fever

— muscle aches

— other, if yes, please describe:

Recommendations: _____

10 Physical examinations



body language

Giving a Physical Examination - For Providers of Medical Care

A physical examination is one of the most basic of medical care procedures. Follow these simple steps to be sure that your exams are thorough and effective every time.

- 1** Make sure the patient is comfortable and that the examination environment is optimal.
- 2** Check **vital signs**. With the patient sitting, take his/her temperature using an oral thermometer. Then measure the patient's **pulse**. Use a stethoscope to conduct **auscultation** of the patient's breathing.
- 3** Have the patient lie down. Use **percussion** techniques to check the condition of organs beneath the abdomen.
- 4** Conduct a general **inspection** of physical body parts. Check the arms, legs, back, and neck with **palpation**. **Evaluate** the patient's posture. Make **observations** of the patient's **body language**. It can reveal symptoms that **affect** the diagnosis.



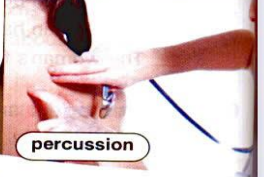
vital signs



pulse



auscultation



percussion

Get ready!

- 1** Before you read the passage, talk about these questions.

- 1 What is involved in giving a physical examination?
- 2 What does it mean to check a patient's vital signs?

Reading

- 2** Read the instructions on giving a physical exam. Then, choose the correct answers.

- 1 What is the purpose of the instructions?
 - A to explain how to give a physical exam
 - B to demonstrate how body language affects vital signs
 - C to compare percussion and auscultation techniques
 - D to describe the importance of physical examinations
- 2 Which of the following is NOT an aspect of checking vital signs?

A checking pulse	C taking temperature
B auscultation	D evaluating posture
- 3 What is one use for percussion techniques?
 - A to measure a patient's pulse
 - B to determine the quality of a patient's breathing
 - C to check the condition of internal organs
 - D to evaluate a patient's mood

Vocabulary

- 3** Place the words and phrases from the word bank under the correct heading.

Word BANK

palpation	auscultation
pulse	vital signs
percussion	body language

Examination techniques

Things to check in an exam

12 Diagnostic equipment



ultrasound



MRI



CT scan

Community Health Clinic Medical Imaging Services

At the Community Health Clinic, we have a wide range of **medical imaging** techniques available. Which technique to use depends on the situation and needs of the patient.

Radiography is the use of **X-rays** to take pictures of the internal body. It is a common imaging technique used in many different situations. A **CT scan** is an advanced form of radiography. It uses many different two-dimensional X-rays to create a three-dimensional image.

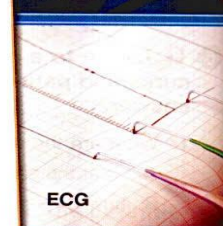
An **ultrasound** uses the reflection of sound waves to create images. A special **gel** placed on the skin aids this process. Unborn children are often examined through ultrasounds.

Recently the use of **endoscopes** during examinations and surgery has become quite common. A doctor **feeds** a tube with a tiny camera on its end directly into the organ or cavity to be examined.

Advanced imaging techniques include **MRIs** and **ECGs**. An MRI utilizes the magnetic properties of atoms to create very detailed pictures. An ECG measures the electrical activity of the heart.



endoscope



ECG



X-ray

Get ready!

- 1 Before you read the passage, talk about these questions.

- What are some different kinds of medical imaging?
- Why should a health professional be familiar with all kinds of diagnostic equipment?

Reading

- 2 Read the clinic's website about medical imaging. Then, choose the correct answers.

- What is the website mainly about?
 - when to use advanced imaging
 - the cost of medical imaging techniques
 - different kinds of medical imaging
 - how X-rays are being replaced by other imaging techniques
- Which of the following is NOT an advanced medical imaging technique?

A MRI	C CT scan
B X-ray	D ECG
- Who is most likely to be examined by an ultrasound?
 - a patient in surgery
 - a patient with a broken bone
 - a patient who is pregnant
 - a patient with abnormal heart activity

Vocabulary

- 3 Match the words (1-5) with the definitions (A-E).

1 — endoscope

2 — CT scan

3 — radiography

4 — ultrasound

5 — MRI

- a form of medical imaging that uses many two-dimensional X-rays to create a three-dimensional image of an object
- a thin, tubular instrument used to examine the inside of an organ or body cavity
- a medical imaging technique that uses magnetic forces on atoms to produce an image of the body
- the use of x-rays to view images of the internal human body
- a medical imaging technique that uses the reflection of sound waves to produce an image of the body

13 Family medicine



Sunday Herald

Family Medicine

by Dr. John Olson, MD

The Heart of Primary Care

For decades the family medical clinic has been the first point of **primary care** for millions of patients. However, today patient demand threatens to overcome the supply of family medicine **practitioners**. Thus it is important to emphasize the need for these vital health care professionals.

Family practitioners provide a variety of health care services to all patients, regardless of age. Not only do they treat injuries and illnesses, they also provide **preventative** care such as routine checkups and immunizations. They **counsel** patients about their health care options and **advise** a recommended course of action. They are a vital source of **health education**. If needed, they **refer** patients to specialists for specific diseases and injuries. Even though they do not work in an emergency room, many family practitioners provide **urgent care** to patients requiring immediate attention.

Family medicine will always be at the core of the health care system. Medical professionals should be encouraged to follow this rewarding career.

Get ready!

① Before you read the passage, talk about these questions.

- 1 How is family medicine structured in your country?
- 2 What are some of the duties of a family practitioner?

Reading

② Read the newspaper article on family medicine. Then, choose the correct answers.

- 1 What is the main idea of the article?
 - A Family medicine is not as important as in the past.
 - B More medical professionals should pursue a career in family medicine.
 - C Family practitioners should not provide urgent care.
 - D Family practitioners are the fastest growing group of medical professionals.
- 2 Which of the following do family practitioners NOT provide?

A urgent care	C specialized care
B checkups and immunizations	D health education
- 3 Which of the following is part of preventative care?

A regular checkup	B treating an injury
C urgent care	D prescribing medicine

Vocabulary

③ Match the words (1-5) with the definitions (A-E).

- | | |
|----------------------|------------------|
| 1 — primary care | 4 — urgent care |
| 2 — health education | 5 — preventative |
| 3 — family medicine | |

- A medicine that provides long-term, general health care for all individuals
- focusing on preventing diseases from occurring rather than curing them
- the act of teaching people about their bodies and good health habits
- the first and main point of medical assistance for patients in a health care system
- health care provided to immediately respond to an injury or illness, but not in an emergency room

15 Geriatrics

Get ready!

1 Before you read the passage, talk about these questions.

- 1 What are some common diseases in elderly patients?
- 2 Why might senior citizens need special care?



Reading

2 Read the website on geriatrics. Then, mark the following statements as true (T) or false (F).

- 1 ___ Bedsores are a result of incontinence.
- 2 ___ Dizziness could be a sign of a more serious disease.
- 3 ___ Nursing homes provide the most freedom for seniors needing assistance.

Vocabulary

3 Match the words (1-6) with the definitions (A-F).

- | | |
|--------------------|--------------------|
| 1 ___ stroke | 4 ___ geriatrics |
| 2 ___ incontinence | 5 ___ polypharmacy |
| 3 ___ nursing home | 6 ___ bed sore |

- A the use of too many medications at once
 B a field of medicine that focuses on preventing and treating diseases in older people
 C a rapid loss of brain functions due to a loss of blood to the brain
 D the inability to control one's bladder or bowel movements
 E a skin wound caused by too much pressure cutting off blood circulation to a part of the body
 F a facility that provides 24-hour care to elderly patients

Medical Specialties: Geriatrics

Geriatrics has developed as a field of medicine focusing on the prevention and treatment of diseases in **elderly** people.

Common Ailments

Certain diseases become more common with advanced age. For example, the risk of a **stroke** increases significantly. **Hearing loss** can require implants to maintain functional levels. Inability to be physically active can cause **bed sores**. Additionally, **Alzheimer's Disease** is a terminal illness that robs people of their memories and their ability to think clearly. Other ailments, such as spells of **dizziness** or **incontinence**, could indicate more serious problems.

Caring for Seniors

Senior citizens must often take several regular medications. This puts them at risk for **polypharmacy**, which can cause dangerous **drug interactions**. It is important to assist seniors who may have trouble understanding when and how much medicine to take. Should their health issues make them incapable of caring for themselves, several care options exist. **Assisted living** centers help seniors with their medical needs while still granting them independence. Seriously ill seniors may need to be placed in a **nursing home**.

4 Fill in the blanks with the correct words and phrases from the word bank.

Word BANK

assisted living
 hearing loss elderly
 Alzheimer's Disease
 drug interaction dizziness

- 1 You must manage your prescriptions to prevent a serious _____.
- 2 My grandfather suffers from _____ and often forgets my name.
- 3 My mother really enjoys the _____ arrangement in her apartment.
- 4 It is important to recognize the specific illnesses that affect _____ people.
- 5 This morning I felt such _____ I could barely stand up.
- 6 My grandmother's _____ is worse so she is getting an ear implant.

- 5 Listen and read the website on geriatrics again. What are some guidelines for taking care of elderly patients?

Listening

- 6 Listen to a conversation between two nurses. Choose the correct answers.

- What are the speakers mainly talking about?
 - how to help seniors move around safely
 - the effects of some drug interactions
 - special needs of some patients
 - a new patient's medication
- Why should the woman help Ms. Clark?
 - She is not able to move around by herself.
 - She may not know that she has a new medication.
 - She may have developed bed sores.
 - She must not take a certain medication right now.

- 7 Listen again and complete the conversation.

- Nurse 1:** Ms. Clark has recently started a new medication. You might have to 1 _____ and make sure she understands it.
- Nurse 2:** 2 _____ is she taking?
- Nurse 1:** She's on Cordarone for an irregular 3 _____. While she's taking that, she can't have any Zocor for her cholesterol. 4 _____ a dangerous interaction.
- Nurse 2:** Okay, 5 _____ she knows to stay off the Zocor.
- Nurse 1:** Good. The rest of the patients require 6 _____.
- Nurse 2:** That sounds fine. Let's get started.



Speaking

- 8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

You might have to ...
That could cause ...
I'll make sure ...

Student A: You are a nurse. Talk to Student B about:

- what care two patients need
- a medication a patient is taking
- a possible drug interaction

Student B: You are a nurse. Talk to Student A about what care two patients need.

Writing

- 9 Use the conversation from Task 8 to complete the nurse's notes.

Assisted Living Center

Nurse: _____
 Wing: 2 East

Patient: _____

Special Care Required: _____

4 Read the sentence and choose the correct word.

- 1 Please organize these files in **chronological order** / **status**.
- 2 Nurse Wilson takes over for me at the **plan of care** / **shift change**.
- 3 Since the doctor's writing was not **comprehensive** / **legible**, the pharmacist could not read it.

5 Listen and read the email about staff communications again. What is a danger of poor communication?

Listening

6 Listen to a conversation between two nurses. Choose the correct answers.

- 1 What is wrong with the medical record?
 - A The wrong medication is listed.
 - B The notes are not in chronological order.
 - C The patient's name is not clearly marked.
 - D The correct dosage amount is missing.
- 2 How will the nurses handle the problem?
 - A They will ask the patient about his medication.
 - B They will review the medical record again.
 - C They will find out who wrote the patient's notes.
 - D They will contact a nurse from the previous shift.

7 Listen again and complete the conversation.

- Nurse 1:** Hey Wendy, is Mr. Williams 1 _____ medication?
- Nurse 2:** I don't know. What does his 2 _____ say?
- Nurse 1:** Well, I'm not sure. I 3 _____ these notes.
- Nurse 2:** Let me see ... Wow, this doesn't make any sense. I think he's supposed to get his meds every day at nine AM, but it doesn't look like 4 _____ dosages for the last three days.
- Nurse 1:** Really? That 5 _____. How could that happen?
- Nurse 2:** Wait, I think I see what happened. These notes aren't in 6 _____. Someone entered last week's dosages first, after this week's. It's all mixed up.
- Nurse 1:** So did he get his medication today or not?
- Nurse 2:** I'm not sure. Who was on duty at nine AM?
- Nurse 1:** I'll check the shift board ... Tom was here this morning.
- Nurse 2:** These notes are 7 _____. We'd better give him a call.
- Nurse 1:** That's a good idea. We certainly don't want to give Mr. Williams the wrong dosage.

Speaking

8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

I can't read ...
It doesn't look like ...
These notes are ...

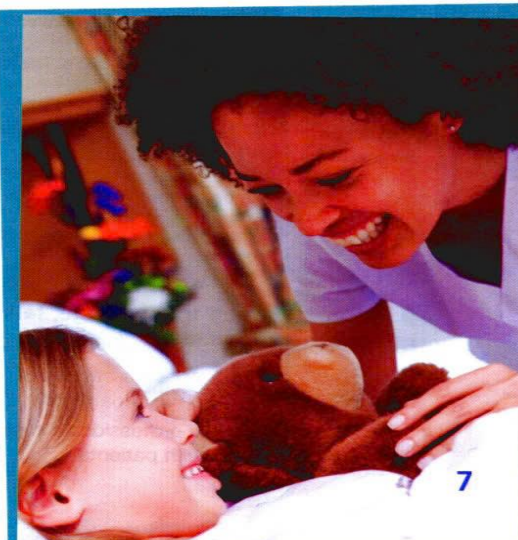
Student A: You are a nurse. Talk to Student B about:

- a patient's medication
- a problem with a medical record
- what to do next

Student B: You are a nurse. Talk to Student A about a problem with a medical record.

Writing

9 Use the conversation from Task 8 and the email to write a memo to the hospital director. Include a description of a communication problem and your steps to fix it.



3

Communicating with patients



To: All Staff

Subject: Communicating with Patients

Clear communication is vital to providing the best care possible. However, recently we have not been getting all the information we need.

If you believe you are not getting adequate information from a patient, remember some of the possible causes of **miscommunication**:

- **Poor Health Literacy** - Uninformed patients may not understand important terms.
- **Negative Preconceptions** - Patients may not trust medical professionals. They may experience **anxiety** when forced to interact with them.
- **Cultural Differences** - Different cultures may have different approaches to dealing with illness and health care.

If you recognize any of these issues, a few simple steps can be all that is required to improve communication:

- Establish a good **rapport** with the patient. Patients heal better in a **therapeutic** environment.
- Show **empathy** and understanding at all times.
- Be aware of how your **verbal** and **nonverbal** cues are affecting the interaction.
- Avoid using **jargon** as much as possible. This may confuse or intimidate the patient.
- Ask **open-ended questions** to signal to patients that you care about their input.

Remember, everyone shares the responsibility of effective communication with patients. The more information we can give and receive, the better care we can provide.

Get ready!

1 Before you read the passage, talk about these questions.

- 1 What are some possible causes of miscommunication with patients?
- 2 How can medical professionals improve communication with patients?

8

rapport

Reading

2 Read the memo on communicating with patients. Then, choose the correct answers.

- 1 What is the main idea of the article?
 - A Medical professionals must learn to trust patients.
 - B There are easy ways to improve communication with patients.
 - C Verbal and nonverbal communication are equally important.
 - D Poor communication with patients can lead to life-threatening risks.
- 2 Which of the following is NOT a cause of miscommunication?
 - A asking too many open-ended questions
 - B low levels of health education
 - C diverse cultural beliefs
 - D mistrust of medical personnel
- 3 What is an appropriate way to improve communication with patients?
 - A Use a lot of technical words.
 - B Avoid becoming too friendly.
 - C Limit nonverbal communication.
 - D Show proper empathy to patients.

Vocabulary

3 Match the words (1-6) with the definitions (A-F).

- | | |
|-------------------------|---------------|
| 1 — verbal | 4 — rapport |
| 2 — therapeutic | 5 — anxiety |
| 3 — open-ended question | 6 — nonverbal |

- A a feeling of understanding and a connection between two people
- B a feeling of nervousness and worry
- C an inquiry that requires an answer longer than a one-word response
- D communication without the use of spoken language
- E spoken
- F helping someone to heal or feel relaxed

4 Read the sentence pair. Choose where the words best fit the blanks.

1 jargon / empathy

- A Using too much _____ can confuse patients.
B Show _____ to your patients without sounding fake.

2 health literacy / miscommunication

- A You must speak clearly to avoid a _____ with a patient.
B The _____ level of many patients is too low.

3 cultural differences / preconceptions

- A You must be aware of _____ that can lead to misunderstandings.
B Never let _____ shape how you treat a patient.

5 Listen and read the memo on communicating with patients again. Why is it important for medical professionals to help patients relax?

Listening

6 Listen to a conversation between a doctor and a nurse. Mark the following statements as true (T) or false (F).

- 1 ___ The patient lied about his medical history.
2 ___ The patient may not understand certain jargon.
3 ___ The man tried to engage the patient in casual conversation.

7 Listen again and complete the conversation.

Nurse: It's possible. But I can't tell because he's barely speaking to me.
Doctor: He might 1 _____. It can be scary to feel terrible but not know why.
Nurse: Yeah, he does seem 2 _____. In fact, I noted that as an additional symptom.
Doctor: Did you use 3 _____? In my experience, that makes things worse.
Nurse: I might have. Since he has 4 _____, I asked him if he has a history of hypoglycemia.
Doctor: See, that term may 5 _____ to him. Try "low blood sugar" instead.
Nurse: Good point.
Doctor: Also, 6 _____ establish a rapport. Even just by asking about his day.
Nurse: Oh. I didn't really do that. I just jumped in with my questions.

Speaking

8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

*She/He's barely speaking ...
Did you use lots of ...
Make sure to ...*

Student A: You are a doctor. Talk to Student B about:

- a communication problem with a patient
- how to improve communication
- an example of unnecessary jargon

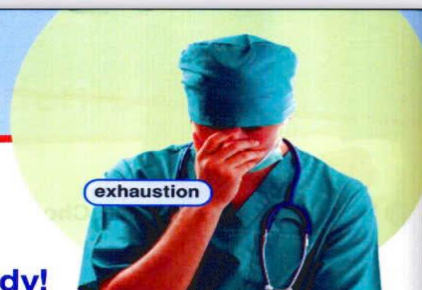
Student B: You are a nurse. Talk to Student A about a communication problem with a patient.

Writing

9 Use the conversation from Task 8 and the memo to write some patient communication guidelines. Include sections on reducing anxiety, avoiding jargon, and establishing a rapport.



4 Challenges



Get ready!

1 Before you read the passage, talk about these questions.

- 1 How can economic hardships affect hospitals?
- 2 How can economic hardships affect individual medical professionals?

Reading

2 Read the article on challenges in the medical profession. Then, choose the correct answers.

- 1 What is the main idea of the article?
 - A how to improve medical practices
 - B poor hospital management strategies
 - C shortfalls in medical education
 - D difficulties faced by medical workers
- 2 Which is NOT a challenge mentioned in the article?
 - A shortages of medical supplies
 - B rising medical school costs
 - C longer shifts for doctors
 - D increased paperwork
- 3 According to the article, what might prevent a student from pursuing medical school?
 - A substantial loans
 - B increased caseloads
 - C too much paperwork
 - D exhaustion

The Ailing Medical Profession

Hospitals today are facing economic hardships. While most doctors and nurses are competent, hardworking individuals, they encounter various challenges that can interfere with patient care.

The **demand** for medical professionals has increased in recent years. With rapidly rising populations, communities need good health care now more than ever. At the same time, however, the economic downturn has led to shortfalls in the **budget** of health care providers, so hospitals cannot afford to hire new staff. Consequently, doctors and nurses are paid less to work longer **shifts** and take on larger **caseloads**. The added **paperwork** forces many to work unpaid overtime just to maintain the necessary records. More and more of these workers are suffering from **exhaustion**.

The **shortage** of medical professionals is partly due to individual financial hardships as well. Increases in medical school costs have forced students to take out **loans** and incur substantial **debt** before they can start practicing. And today's doctors must also pay hefty insurance premiums each month in case they are sued for medical **malpractice**. For many would-be doctors, these expenses are too high, and so students are pursuing more affordable career paths.

10

Vocabulary

3 Write a word that is similar in meaning to the underlined part.

- 1 Dr. Bainbridge was sued for neglectful or illegal action after the failed surgery.
_ _ l p _ _ c t _ _ e
- 2 I have to file this official documentation in the patient's medical record. p _ _ e r _ _ k
- 3 Our available amount of money decreased this year by 15%. b _ _ g e _
- 4 Henry cannot afford to pay off his amount of money owed. _ e b _

10 Traumatic injuries



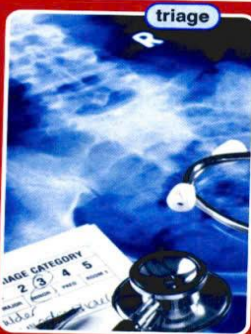
Heartland Hospital

ER TRIAGE GUIDELINES

Heartland Hospital strives to provide the best care it can with limited **resources**. To ensure that each patient is seen at the appropriate time and receives quality services, please follow these guidelines for the **triage** process.

Trauma patients who arrive with obviously **critical** and life-threatening **injuries** receive the highest **priority**. They should be taken to an available trauma room immediately. Any necessary staff and equipment should be devoted to them until their vital signs are stabilized.

All other patients should report to the **triage station** to undergo an assessment by a nurse. The nurse will utilize the **Emergency Severity Index** to determine the severity of a patient's injuries and when they should be seen. This is a common-sense system that ranks patients according to the danger level of their situation. For example, an open **wound** will rank higher on the Index than a **minor** illness such as a low fever. The nurse will assign an **ESI level** between 1 and 5, with 1 being the most serious, to each patient. In this way, patients will be seen according to the **acuity** of their injuries.



trauma



resources



wound



triage station

Get ready!

1 Before you read the passage, talk about these questions.

- 1 What kind of patients are seen in an emergency room?
- 2 How does the triage process generally work?

Reading

2 Read the poster on treating patients in an emergency room. Then, mark the following statements as true (T) or false (F).

- 1 — Fevers are not ranked on the Emergency Severity Index.
- 2 — A patient with a level one ESI rating has the highest priority.
- 3 — Only an emergency room doctor should assign a patient an ESI rating.

Vocabulary

3 Match the words (1-7) with the definitions (A-G).

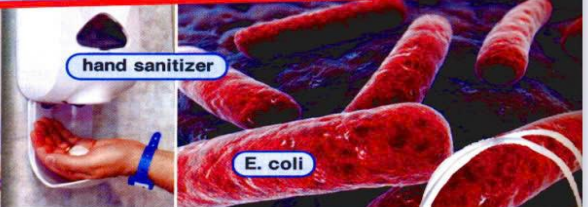
- | | |
|---------------|--------------|
| 1 — priority | 5 — triage |
| 2 — resource | 6 — minor |
| 3 — wound | 7 — critical |
| 4 — ESI level | |

- A a unit of measurement of a trauma patient's injuries
- B the process of determining the order to treat trauma patients based on the severity of their injuries
- C something that is more important than other things
- D requiring immediate medical attention
- E an injury in which a part of the body is damaged
- F not very serious
- G anything a person uses to achieve a goal

11 Infections



nosocomial



hand sanitizer

E. coli

urinary catheter

MAY ISSUE

Journal of Hospital Issues

Stopping Infections in Your Hospital

Hospitals are full of **vectors** of infectious diseases. According to estimates, about 1.7 million American hospital patients get a **nosocomial** infection each year. Health care facilities have a responsibility to be **vigilant** about protecting their patients from infections. There are many simple ways this can be achieved.

The majority of illnesses are spread through touch, so it is essential that hospitals have many hand-washing stations available to patients, visitors, and employees. Problems like **E. coli** and **staph infections** can be prevented by using **alcohol-based hand sanitizers**. However, these sanitizers do not kill all bacteria. Washing frequently with soap and water will help guarantee reduced infection outbreaks.

Hospitals should try to prevent illnesses before they spread. Diagnostic tests can detect infections, such as those caused by **drug-resistant MRSA** bacteria. Catching an infection early is a key step to stopping it in its tracks.

Finally, hospitals must have standard processes and tasks for procedures. For example, in intensive care units, nurses should brush their patient's teeth regularly. This kills the bacteria that causes **ventilator-associated pneumonia**. Similarly, LPNs should have checklists of the steps for fitting patients with **urinary catheters**, since an improper fit can lead to infection. In addition, cleaning procedures for medical tools must be followed carefully. It is equally important not to **contaminate** sterilized tools.

24

Get ready!

1 Before you read the passage, talk about these questions.

- How do infections spread?
- What are some ways to prevent infections from spreading?

Reading

2 Read the article on preventing hospital infections. Then, complete the table using information from the article.

Problems	Prevention
E. coli and staph infections	1 _____
Drug-resistant MRSA bacteria	2 _____
Ventilator-associated pneumonia	3 _____
Improper fit of urinary catheter	4 _____

Vocabulary

3 Fill in the blanks with the correct words and phrases from the word bank.

Word BANK

alcohol-based drug resistant
vigilant hand sanitizers

- The _____ nurse washed her hands after working with each patient.
- This substance is _____. It's not made with water.
- Peter is not getting better because his virus is _____.
- _____ kill most bacteria, but not all of them.

4 Check (✓) the sentence that uses the underlined part correctly.

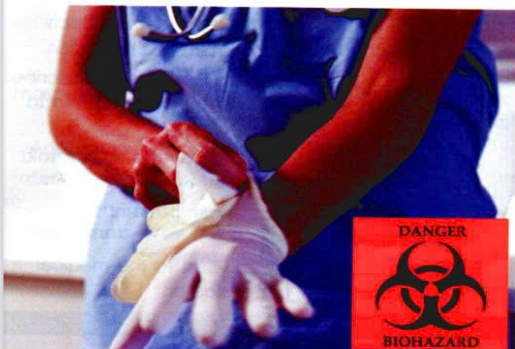
- 1 — A nosocomial infection is one that cannot be cured with most drugs.
— B The hands are a vector of illness.
- 2 — A The man contaminated the wound when he touched it with his bare hands.
— B Ventilator-associated pneumonia is common in pediatric units.
- 3 — A MRSA can be treated using over-the-counter drugs.
— B Lewis contracted E. coli after eating raw spinach.
- 4 — A Staph infections are caused when a patient is on a ventilator for too long.
— B The woman has a urinary catheter because she can't leave her bed.

5 Listen and read the article on preventing hospital infections again. What are some ways to prevent MRSA from spreading?

Listening

6 Listen to a conversation between a nurse and a doctor. Mark the following statements as true (T) or false (F).

- 1 — The man believes that antibiotics will not help the patient's condition.
- 2 — The diagnostic tests showed that the other man in the patient's room also has MRSA.
- 3 — The patient's room is being disinfected to prevent MRSA from spreading.



7 Listen again and complete the conversation.

Nurse: Well, the red, swollen wounds on his body are now 1 _____ pus.

Doctor: 2 _____?

Nurse: Yes, he said that he was short of breath, and his temperature is 3 _____.

Doctor: I'm afraid it may be MRSA. We need 4 _____ him on an aggressive round of antibiotics. And he needs to be in isolation.

Nurse: Right. I'll get started on that 5 _____.

Doctor: Has anyone else in this unit been infected?

Nurse: We ran some diagnostic tests on the other patient in Mr. Harris' room, but his test results 6 _____ negative. None of the other patients are exhibiting symptoms.

Speaking

8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

Anything else?

Has anyone else ...

The test results came back ...

Student A: You are a doctor. Talk to Student B about:

- a patient's current condition
- what could happen to the patient

Student B: You are a nurse. Talk to Student A about a sick patient.

Writing

9 Use the conversation from Task 8 and the article to write notes on a patient. Include the patient's name, condition and symptoms, preventative measures, and treatment directions and concerns.

- 5 🎧 Listen and read the job postings again. What are some common tasks for RNs?

Listening

- 6 🎧 Listen to a conversation between a job interviewer and a nurse. Mark the following statements as true (T) or false (F).

- 1 ___ The woman is a nurse practitioner.
- 2 ___ One of the job duties is to give vaccines.
- 3 ___ The woman prefers the fast patient flow of hospitals.

- 7 🎧 Listen again and complete the conversation.

Interviewer: Jane, I'm Richard Jenson. I'll be interviewing you for the 1 _____.

Applicant: Nice to meet you, Mr. Jenson.

Interviewer: So what kind of experience do you have?

Applicant: Well, I started volunteering in hospitals as a teenager. I've worked as an RN for the past 2 _____.

Interviewer: And why did you go into nursing?

Applicant: I love taking care of people. It brings me a lot of joy.

Interviewer: Excellent. Now, this is a general practice. You'll be giving lots of 3 _____ and assisting with annual physicals.

Applicant: Everyday tasks. I understand.

Interviewer: Is that appealing to you?

Applicant: Yes, because there's more interaction with patients. Hospitals are 4 _____.

Interviewer: Right, it's definitely slower here. But we still have to maintain patient flow.

Applicant: Of course.

Interviewer: Now, have you worked with 5 _____?

Applicant: Yes, we 6 _____ them several years ago.

Speaking

- 8 With a partner, act out the roles below based on Task 7. Then, switch roles.

USE LANGUAGE SUCH AS:

I'll be interviewing you for ...

I have worked ...

Student A: You are a job interviewer. Talk to Student B about:

- his or her experience
- what kind of work he or she will be doing
- familiarity with medical systems

Student B: You are a nurse. Talk to Student A about your qualifications.

Writing

- 9 Use the conversation from Task 8 and the job posting to write notes on the interview. Include: the applicant's experience, the reason he or she entered nursing, and what tasks he or she will be responsible for.



13 OB/GYN



MIDWEST OB/GYN CLINIC

Services Offered

At Midwest **OB/GYN** Clinic we offer complete medical care for women at every stage of their life cycle. Many specific services are tailored to pregnant women and those going through **menopause**.

Pregnancy

Midwest OB/GYN Clinic provides care during all three **trimesters** of **pregnancy**. Regular appointments and ultrasounds safeguard the health of both the mother and child. Our clinic also conducts **fertility** treatments for women finding it difficult to become pregnant. For example, some women undergo **progesterone** therapy before and during pregnancy.

Midwest also offers classes for expecting parents. Topics include:

- How to prepare for **labor**
- What to do when your **water breaks**
- The possibility of a **c-section** delivery when necessary
- What to do when you arrive home with your newborn

Please remember that our clinic does not provide emergency services. You should contact the hospital's **OB Triage** station if you are having rapid **contractions** or experience a medical emergency during pregnancy.

Menopause

Menopause is a natural stage in every woman's life. Still, the **hot flashes**, emotional impact, and other effects can be difficult to handle. Women going through menopause can schedule regular appointments at our clinic. We can advise patients on lifestyle changes to make menopause more comfortable. We also offer hormone therapies, such as **estrogen** treatments, should they be necessary.

Get ready!

1 Before you read the passage, talk about these questions.

- What medical services do pregnant women need?
- What special medical services do older women need?

Reading

2 Read the OB/GYN clinic's website. Then, mark the following statements as true (T) or false (F).

- ☐ The clinic offers classes on the changes that occur in menopause.
- ☐ Pregnant women should contact the clinic in an emergency.
- ☐ The clinic gives progesterone treatments to women in menopause.

Vocabulary

3 Match the words (1-7) with the definitions (A-G).

1 ☐ estrogen	5 ☐ labor
2 ☐ menopause	6 ☐ progesterone
3 ☐ hot flash	7 ☐ break water
4 ☐ trimester	

- a sudden sensation of heat in the face and neck that is triggered by hormone changes
- a natural childbirth process in which a woman's uterus contracts in preparation to deliver the baby
- a hormone that is produced in the ovaries which prepares the uterus lining for pregnancy
- the end of a woman's menstruation, after which she is no longer capable of becoming pregnant
- a period of three months that pregnancy is commonly divided into
- a hormone that is produced in the ovaries which regulates the menstrual cycle
- to release amniotic fluid after the placenta ruptures

LIST OF THE AUTHOR'S SCIENTIFIC PUBLICATIONS RELATED TO THE THESIS

- Pham, B. N. (2022). Interactive meaning of visual images in medical textbooks: A multimodal discourse analysis. In *Proceedings of the National Scientific Conference UNC 2022* (Vol.1, pp. 334-340). Vietnam National University Press.
- Pham, B. N. (2022). Towards a multimodal analysis of images in a medical textbook. In *VietTESOL International Convention 2021: Rethinking English language education in the Covid Era*, (Vol.2, pp. 271-279). Vinh University Press. <https://proceedings.viettesol.org.vn/index.php/vic/issue/view/1>
- Pham, B. N. (2023). Interactive mode of visual images in a medical textbook: A multimodal discourse analysis. In *International Graduate Research Symposium*, (Vol.1, pp. 1113-1122). Vietnam National University Press.
- Lam, Q. D., & Pham, B. N. (2024). A multimodal discourse analysis of a medical textbook in Vietnam: Preliminary findings. *VNU Journal of Foreign Studies*, 40(3), 21-34. <https://doi.org/10.63023/2525-2445/jfs.ulis.5220>
- Pham, B. N. (2025). A multimodal discourse analysis of visual images in a medical textbook. *Journal of Language and Life: Linguistics Society of Vietnam*, 3 (365), 51-58. National Political Publishing House.
- Pham, B. N. (2025). An analysis of communication modes of visual images in a medical textbook. *Journal of Language and Life: Linguistics Society of Vietnam*, 6B (369), 190-203. National Political Publishing House.
- Pham, B. N. & Lam, Q. D. (2025). What do the visual images say in a medical textbook? In *International Graduate Research Symposium*, (Vol.1, pp. 908-916). Vietnam National University Press.
- Pham, B. N. & Lam, Q. D. (2025). Deconstructing visual discourse: A multimodal perspective on images in a medical textbook. In *Proceedings of the International Scientific Conference: Functional Linguistics and Multimodality* (pp. 472-482). Vietnam National University Press.

Pham, B. N. (2025). A multimodal discourse analysis of images in a medical textbook: Representational modes in focus. In *Proceedings of the International Conference on Language, Media and Culture (UniC2025.2)* (pp. 357-370). Vietnam National University Press.