

EXPLORING PRACTITIONER COLLABORATIVE ASSESSMENT IDENTITY TO DEVELOP A PRINCIPLED MULTI-LENS APPROACH TO ASSESSING COLLABORATIVE ASSIGNMENTS IN ENGLISH FOR ACADEMIC PURPOSES

Peter Levrai

Supervised by:
Dr. María del Pilar García Mayo

University of the Basque Country
2024



Director: Dr. Maria del Pilar Garcia Mayo
Universidad del Pais Vasco – Euskal Herriko Unibertsitatea (UPV/EHU)
Department of English, German and Translation and Interpretation

Acknowledgements

A PhD is a long, laborious process, particularly one which was conducted during the COVID pandemic, which disrupted both the world, and this work, resetting it to zero more than once. Having said that, the main thanks for this dissertation has to go to Averil Bolster, researcher of the companion dissertation, and very best friend and partner anyone could wish for. It's been a long journey through innumerable challenges but what has been of great help is knowing we're engaged in doing something important and necessary. I would never have started this PhD without her, and I certainly wouldn't have been able to navigate it without her support and counsel.

Entering a new discourse community requires guidance and Dr Maria del Pilar Garcia Mayo, who played both the roles of guide and critical friend, more than fulfilled that role. Her constant support and encouragement provided a space to adapt, and adapt again, in light of changing circumstances. Her feedback was always prompt, thoughtful, and necessary. Any errors which remain in the work are very much mine.

Ultimately, however, this PhD is built on the knowledge and experience of fellow EAP practitioners, who selflessly gave their time, thought and consideration to this study. Regardless of how this PhD is received, their thoughtful contributions assure me EAP is in good hands. They cannot be named for reasons of confidentiality, but each and every one of you has my thanks. Gregory Hadley was also an important inspiration during the necessary pandemic pivot, turning this dissertation to a Grounded Theory project, which gave those participants their voice.

Thanks must also go to my colleagues at University of Turku. Undertaking a part-time PhD is a time-consuming endeavour but I never lacked for support from my manager, Michael Nelson and of considerable help was the support of colleagues also undertaking PhD studies. Good luck to all of us.

It would also be remiss to ignore the contributions of people well before I thought of undertaking a PhD. Dr Naeema Hann was a great supervisor for my MA in ELT and Materials Development, pushing me toward my best work, ably supported by Heather Buchanan, and that study

introduced me to the vital work of Brian Tomlinson. My first step from practitioner to practitioner-researcher took place during my International Diploma in Language Teaching Management and thanks must also go to Ron White, Andy Hockley, and George Pickering.

It is also important to acknowledge my colleagues in University of Nottingham Ningbo China (UNNC) and University of Macau, who were with me through the experience of realising collaborative assignments exist and are a fact of life we need to deal with in EAP. UNNC, in particular, provided a great grounding in EAP, providing my first experience in the academic English field.

Thanks must also go to my family, for their support every step of the way, with particular consideration for my father's constant backing. Ultimately, however, the acknowledgements for this dissertation must return to Averil Bolster, without whom it certainly wouldn't exist. You've always guided me to the things I need to do and supported me in doing them. Thank you.

“In a programme of assessment, individual assessments are purposefully chosen in such a way that the whole is more than the sum of its parts. Not every individual assessment, therefore, needs to be perfect.” (Van de Vleuten et al., 2015, p. 641)

Table of Contents

List of Acronyms	xvii
List of Tables.....	xix
List of Figures.....	xxi
Abstract.....	xxiv
Resumen.....	xxviii
INTRODUCTION	1
An opening vignette.....	2
Researcher motivation.....	5
Origins of this dissertation.....	7
Research focus	8
Understanding the issue	9
Overview of this dissertation	10
1 LITERATURE REVIEW	17
1.1 Chapter overview.....	17
1.2 Collaboration.....	19
1.2.1 Collaboration as a valued global skill.....	19
1.2.2 Collaborative learning in education.....	21
1.2.3 Sociocultural Theory	22
1.2.3.1 Zone of Proximal Development and the More Knowledgeable Other	23

1.2.4	Collaboration in language education	25
1.2.4.1	Groupwork	25
1.2.5	Collaborative writing	28
1.3	English for Academic Purposes in Higher Education	29
1.3.1	Background study: Collaboration in EAP	33
1.4	Assessment	35
1.4.1	Bridging assessment and learning outcomes	37
1.4.2	Assessment of/for/as Learning	38
1.4.3	Evaluative judgement and the role of self- and peer- assessment	39
1.4.4	Assessment and Sociocultural Theory	40
1.4.5	Locating collaborative assessment	41
1.4.6	Assessing collaboration	46
1.4.6.1	Summative grading of collaborative assignments	47
1.4.6.2	How collaborative assessment can impact student behaviour	48
1.4.6.3	Freeloaders, social loafing and suckers	50
1.4.6.4	Looking inside a collaboration	51
1.4.6.5	What to assess when assessing a collaborative assignment	52
1.4.7	Assessment literacy	54
1.5	Grounded Theory	58
1.5.1	The Grounded Theory family	59
1.5.2	Core concepts in Grounded Theory	62
1.5.2.1	Theory	62

1.5.2.2	Emergence	63
1.5.2.3	Theoretical sampling and theoretical saturation.....	63
1.5.2.4	The process of grounded theory.....	64
1.5.2.5	Theory evaluation	66
1.5.3	Critiques of Grounded Theory	68
1.6	Chapter conclusion	69
2	METHODOLOGY	73
2.1	Chapter overview.....	73
2.2	Philosophical grounding.....	74
2.2.1	Philosophical eclecticism	74
2.2.2	Philosophical framework of the study	76
2.3	The evolution of this dissertation	78
2.3.1	The rise and fall of the original research plan.....	81
2.3.2	The approach of the original research plan	81
2.3.3	Original research questions and hypotheses.....	82
2.3.4	Original research plan data collection tools	83
2.3.4.1	Teacher and student surveys	84
2.3.4.2	Teacher Interviews.....	84
2.3.4.3	Student focus groups	84
2.3.4.4	Gradebook analysis	85
2.3.5	The pandemic interruption	85
2.4	The current research plan.....	86

2.4.1	Application of Grounded Theory in this dissertation.....	87
2.4.1.1	The process of this dissertation	88
2.4.1.2	Approach to literature	90
2.4.1.3	Role of the insider researcher	91
2.4.2	Research questions and hypotheses.....	93
2.4.3	Ethics	94
2.4.4	Consent	94
2.4.5	Anonymity and confidentiality.....	94
2.5	Technological aids.....	95
2.5.1	Microsoft Office Suite	95
2.5.2	Webropol	95
2.5.3	Zoom	95
2.5.4	NVivo12 & NVivo20.....	96
2.5.5	WhatsApp.....	96
2.5.6	Otter.ai	96
2.5.7	Samsung Notes.....	97
2.6	Data collection	97
2.6.1	Sampling.....	97
2.6.1.1	Initial sampling.....	97
2.6.1.2	Theoretical sampling.....	98
2.6.2	Research tools.....	99
2.6.2.1	EAP Practitioner survey.....	99

2.6.2.2	Pre-interview task	101
2.6.2.3	Pre-interview task coverage	101
2.6.2.4	Models of collaborative assessment.....	102
2.6.2.5	EAP Practitioner Interview	104
2.7	Participants	107
2.7.1	Andrea (pilot round)	109
2.7.2	Ava.....	109
2.7.3	Barry (pilot round).....	110
2.7.4	Bertha.....	110
2.7.5	Cristobal	110
2.7.6	Dolly	111
2.7.7	Dorian (pilot round)	111
2.7.8	Fay	111
2.7.9	Fernand	112
2.7.10	Gonzalo	112
2.7.11	Humberto.....	113
2.7.12	Jerry.....	113
2.7.13	Karen	114
2.7.14	Kyle.....	114
2.7.15	Laura.....	115
2.7.16	Lorenzo.....	115
2.7.17	Olga	116

2.7.18	Tanya (pilot round).....	116
2.7.19	Van	117
2.7.20	Wendy.....	117
2.7.21	Anna, Claudette and Elsa	118
2.8	Data analysis	118
2.8.1	Types of reasoning	119
2.8.2	Memoing.....	120
2.8.2.1	Illustration of memoing	121
2.8.3	Open exploration	122
2.8.3.1	Approach to coding.....	123
2.8.4	Focussed exploration	127
2.8.5	Theoretical coding and diagramming	133
2.8.6	A return to immersion	139
2.9	Going beyond sufficiency.....	139
2.9.1	Actualizing the collaborative assessment algorithm: Starter guide design	140
2.9.2	Evaluating the starter guide.....	142
2.10	Chapter conclusion	144
3	FINDINGS	147
3.1	Chapter overview	147
3.2	EAP and collaboration.....	148
3.2.1	Participants' understanding of EAP	149
3.2.2	Participants' understanding of collaboration	154

3.2.3	Collaboration in EAP	159
3.3	Participant evaluation of models of assessment	164
3.3.1	Model A: Group grade	165
3.3.1.1	Model A: Positive reactions	167
3.3.1.2	Model A: Negative reactions.....	168
3.3.1.3	Suggestions for implementing Model A.....	171
3.3.1.4	General summary of Model A evaluation.....	172
3.3.2	Model B: Group grade + complementary individual assignments.....	173
3.3.2.1	Model B: Positive reactions	174
3.3.2.2	Model B: Negative reactions.....	176
3.3.2.3	Suggestions for implementing Model B.....	177
3.3.2.4	General summary of Model B evaluation	178
3.3.3	Mode C: Group grade + individual collaboration grade.....	179
3.3.3.1	Model C: Positive reactions	181
3.3.3.2	Model C: Negative reactions.....	182
3.3.3.3	Suggestions for implementing Model C.....	188
3.3.3.4	General summary of Model C evaluation	189
3.3.4	Model D: Group grade x weighted collaboration grade	190
3.3.4.1	Model D: Positive reactions	192
3.3.4.2	Model D: Negative reactions	194
3.3.4.3	Suggestions for implementing Model D	195
3.3.4.4	General summary of Model D evaluation.....	196

3.4	Participants' proposals of a model of assessment.....	197
3.5	Key principles and guiding orientations.....	198
3.5.1	Simplicity.....	200
3.5.1.1	Product glare.....	201
3.5.2	Transparency.....	203
3.5.3	Fairness	207
3.5.3.1	Differentiation and surface level differentiation	210
3.5.4	Reward orientation	212
3.5.5	Practical Orientation	216
3.5.5.1	Real world relevance.....	216
3.5.5.2	Expertise gaps	217
3.5.5.3	Tutor and student training.....	219
3.5.5.4	Awareness of own limits	220
3.5.5.5	Workload.....	221
3.5.5.6	Assessment overload	223
3.5.6	Defensive orientation	224
3.5.6.1	Environment: Working in HE.....	225
3.5.6.2	The interinfluence of expectation and experience	229
3.5.6.3	Expectations of collaborative assignments.....	236
3.5.6.3.1	Personality and personal relationships.....	237
3.5.6.3.2	Educational maturity.....	240
3.5.6.3.3	Language level	243

3.5.6.3.4	Free-riders and high achievers.....	244
3.5.6.3.5	Circumventing collaboration.....	247
3.5.6.3.6	Grade focus and performative collaboration	249
3.5.6.3.7	Students as assessment agents	252
3.6	Defensive avoidance, acceptance and adaptation	258
3.7	Pertinent findings of the companion dissertation	260
3.8	Evaluations of the starter guide to assessing collaborative assignments.....	262
3.8.1	Evaluations of the first draft of the starter guide	262
3.8.2	Evaluations of the second draft of the starter guide	266
3.8.2.1	Bertha’s review (Reviews of the starter guide, p. 9-10)	267
3.8.2.2	Jerry’s review (Reviews of the user guide p. 12-13)	268
3.8.2.3	Lorenzo’s review (Reviews of the user guide p. 14-15)	269
3.9	Chapter conclusion	269
4	DISCUSSION AND THEORY GENERATION.....	273
4.1	Chapter overview.....	273
4.2	Collaborative assessment identity	275
4.2.1	Eight collaborative assessment archetypes	280
4.3	A collaborative assessment algorithm: Designing a principled assessment scheme for a collaborative assignment	285
4.3.1	Layer 1 – Motivation: Rationalising collaborative assignments in EAP	288
4.3.2	Layer 2 – Purpose: Conceptualising a valid collaborative assessment	290

4.3.3	Layer 3 – Readiness: Considerations before implementing a collaborative assignment	292
4.3.4	Layer 4 – Lens selection: Selecting assessment lenses	293
4.3.4.1	The final product as assessment lens	296
4.3.4.2	Individual complementary assignments as assessment lenses	301
4.3.4.3	Collaboration itself as an assessment lens	303
4.3.4.3.1	What is “contribution”?	305
4.3.4.3.2	Utilising self- and peer-assessment to evaluation collaboration.....	309
4.3.4.3.3	The challenges of assessing collaboration and the case for cooperation	310
4.3.4.3.4	Criteria driven assessment of collaboration	313
4.3.4.3.5	Assessing awareness, not performance	315
4.3.4.4	Learning outcome directed lens selection	316
4.3.5	Layer 5 – Evaluation: Using the dimensions of collaborative assessment identity to evaluate a collaborative assignment assessment scheme	319
4.4	Application of the algorithm to the opening vignette	321
4.5	From theory to practice: From algorithm to guide	322
4.5.1	Reactions to the starter guide	324
4.6	Theory evaluation	325
4.6.1	Credibility	328
4.6.2	Originality	328
4.6.3	Resonance	329
4.6.4	Usefulness	329

4.7	Chapter conclusion	330
5	CONCLUSION	335
5.1	Summary of findings	336
5.2	Research question and emergent areas of interest.....	342
5.2.1	Research questions and hypotheses from the pre-pandemic research plan	344
5.3	Limitations and future directions.....	348
5.4	Implications and significance	351
5.5	Closing thoughts.....	353
	REFERENCES.....	357
	LIST OF APPENDICES	383
	Appendix 1: Dissertation data collection tools (including ethic consent, surveys, pre-interview tasks & interview protocols)	383
	Appendix 2: Survey data.....	383
	Appendix 3: Pre-interview task data	383
	Appendix 4: Interview transcripts	383
	Appendix 5: Secondary Interview Data	383
	Appendix 6: Data processing codebooks.....	383
	Appendix 7: Data collection tools for original study (including ethic consent, student and teacher surveys, student and teacher interview protocols)	383
	Appendix 8: Motivation and background studies	383
	Appendix 9: First draft starter guide to assessing collaborative assessments.....	383
	Appendix 10: Information sheet for starter guide evaluation	383

Appendix 11: Reviews of the starter guide	383
Appendix 12: Figure 4-9 - Designing a collaborative assignment assessment scheme	383
Appendix 13: Second draft starter guide to assessing collaborative assignments	383

List of Acronyms

AA – Authentic assessment

AaL – Assessment as learning

AfL – Assessment for learning

AoL – Assessment of learning

AT – Activity Theory

ARG – Assessment Reform Group

BA – Bachelor of Arts

BALEAP – Formerly British Association of Lecturers of English for Academic Purposes

BALEAP CFTEAP - Competency Framework for Teachers of English for Academic Purposes

BALEAP TEAP – BALEAP Teaching of English for Academic Purposes

CBA – Classroom-based assessment

CEFR – Common European Framework of Reference

CSCC - Computer Supported Collaborative Communication

DA – Dynamic assessment

EAP – English for Academic Purposes

EATAW - European Association for the Teaching of Academic Writing

EGAP – English for General Academic Purposes

ELC – English Language Centre

EMI – English medium instruction

ESAP – English for Specific Academic Purposes

ESP – English for Specific Purposes

GLAID - Group Learning Activities Instructional Design

GT – Grounded Theory

GTA – Grounded Theory Approach

HE – Higher Education

LLMs – Large language models

L2 – Second language

MA – Masters of Arts

MKO – More knowledgeable other

NFEAP – Norwegian Forum for English for Academic Purposes

NNES – Non-native English speaking

PhD – Doctor of Philosophy

PISA - Programme for International Student Assessment

ROG – Rules of the Game

SCT – Sociocultural theory

TESOL - Teaching English as a Second or Other Language

UKVI - United Kingdom Visas and Immigration

UM – University of Macau

UTU – University of Turku

ZPD – Zone of proximal development

List of Tables

Table 0.1 Grade allocation for the collaborative essay assignment (Levrai & Bolster, 2019, p. 194)	6
Table 1.1 Eight features of groups (tabulated from Jaques, 2000, p. 1-2)	26
Table 1.2 Defining features of collaborative writing tasks (Storch & Knoch, 2023, p. 229)	28
Table 1.3 CFTEAP Competencies	33
Table 1.4 Key questions for assessment in EAP (Ding, 2013)	35
Table 2.1 Collaborative assignment complementary dissertations	79
Table 2.2 How data was anonymised	94
Table 2.3 Summary of invitations to participate and responses	98
Table 2.4 Survey Coverage	100
Table 2.5 Summary of the general coverage of the three iterations of the pre-interview task	101
Table 2.6 Models of collaborative assessment used in the pre-interview task	103
Table 2.7 Summary of the general coverage of the three iterations of the interview protocol	106
Table 2.8 Interview length	107
Table 2.9 Participant biodata from the survey	108
Table 2.10 Initial coding methods of pilot data in Saldaña's nomenclature (2016)	124
Table 2.11 Nascent concepts from early data analysis	127
Table 3.1 Participants' star ratings for the four models of assessment from the pre-interview task	165
Table 3.2 Star rating for Model A	165
Table 3.3 Suggestions to implement Model A	171
Table 3.4 Star rating for Model B	173
Table 3.5 Suggestions to implement Model B	178
Table 3.6 Star rating for Model C	179
Table 3.7 Suggestions to implement Model C	188
Table 3.8 Star rating for Model D	190

Table 3.9 Suggestions to implement Model D.....	195
Table 3.10 Key principles and guiding orientations for EAP practitioners assessing collaborative assignments	199
Table 4.1 Sample criteria for assessing a collaborative assignment	314
Table 4.2 Using collaborative assessment identity as criteria to evaluate an assessment scheme for a collaborative assignment.....	320
Table 4.3 Tentative assessment scheme for the vignette assignment.....	322

List of Figures

Figure 0-1 Student attitude to group essays (Levrai and Bolster, 2017)	7
Figure 1-1 Theoretical framework for examining three core concepts.....	17
Figure 1-2 Global Skills and their four foundational frameworks (adapted from Oxford University Press (n.d.)).....	20
Figure 1-3 Locating collaborative assessment within AaL	41
Figure 1-4 The six qualities of test usefulness (van Daalen, 1999, p. 3)	44
Figure 1-5 The three elements of collaborative writing (based on Storch & Knoch, 2023, p.234)	52
Figure 1-6 Suggested elements of collaborative writing in EAP	53
Figure 1-7 Charmaz's Four Main Strategies of Grounded Theory	60
Figure 1-8 Defining actions of a Grounded Theory Researcher (adapted from Charmaz, 2014, p. 15).....	61
Figure 1-9 Core Methodological Characteristics of GTM (Hadley, 2017, p. 59)	61
Figure 1-10 Summary of techniques used in GTM (Hadley, 2017, p.131)	65
Figure 1-11 A visual representation of Grounded Theory (Charmaz, 2014, p. 18).....	66
Figure 1-12 Matrix for Grounded Theory Evaluation (Hadley, 2017)	67
Figure 2-1 Aligning structure, pattern and process to philosophical paradigms.....	76
Figure 2-2 Autopoietic Movement within the History of Qualitative Research (Hadley, 2017, p. 27).77	
Figure 2-3 Research tools and data in the complementary studies	80
Figure 2-4 Fundamental Tenets of Grounded Theory (Charmaz, 2008, p. 155).....	88
Figure 2-5 The process of this study	89
Figure 2-6 Concurrent processes in Grounded Theory.....	90
Figure 2-7 Approach to Literature in the study	91
Figure 2-8 Illustration of the cycle of reasoning, from inductive to deductive to abductive	119
Figure 2-9 Extract from near the beginning of the project memo	121
Figure 2-10 Iterative immersion and coding analysis.....	122

Figure 2-11 Illustration of open coding	126
Figure 2-12 Grouping & categorising open codes	130
Figure 2-13 Illustration of focused coding	132
Figure 2-14 Illustration of multi-lens assessment of collaboration (theoretical diagramming).....	133
Figure 2-15 Key principles shaping attitude to collaborative assessment (theoretical diagramming)	134
Figure 2-16 Principles & Orientations informing a collaborative assessment identity (theoretical diagramming)	135
Figure 2-17 Defensive orientation: origins and outcomes (theoretical diagramming)	136
Figure 2-18 Guiding orientations (theoretical diagramming).....	137
Figure 2-19 Illustration of concept coding.....	138
Figure 2-20 Example of a Sketchplanation used in the starter guide illustrating the cobra effect (Hey, 2019)	142
Figure 3-1 Summary of evaluation of Model A: Group Grade.....	167
Figure 3-2 Summary of evaluation of Model B: Group grade + individual complementary assignments	176
Figure 3-3 Summary of evaluation of Model C: Group grade + individual collaboration grade.....	181
Figure 3-4 Summary of evaluation of Model D: Group grade x weighted collaboration grade	192
Figure 3-5 Key factors informing the practical orientation	216
Figure 3-6 Factors shaping the defensive orientation	225
Figure 3-7 Impactful expectations of student collaboration	237
Figure 3-8 The central role in assessment of the reluctant overlord teacher	257
Figure 4-1 Reconceptualising teacher assessment identity (Looney et al., 2018, p. 456)	276
Figure 4-2 The overlapping key principles and embedded guiding orientations on EAP practitioners for collaborative assessment.....	277
Figure 4-3 The spectrum of collaborative assessment archetypes	281

Figure 4-4 A collaborative assessment algorithm: 5 layers in designing an assessment scheme for a collaborative assignment	286
Figure 4-5 Impact of assessment purpose on assessment focus.....	291
Figure 4-6 Readiness considerations before implementing a collaborative assignment	292
Figure 4-7 Potential assessment lenses for collaborative assignments	294
Figure 4-8 The collaborative assessment identity prism: key principles encompassing guiding orientations.....	294
Figure 4-9 Learner training opportunity cost of assessment lenses.....	296
Figure 4-10 Product glare masking quality of collaboration.....	297
Figure 4-11 Interaction patterns based on Storch, 2002; 2013, Li & Zhu, 2013, 2017; Edstrom, 2015	304
Figure 4-12 Teacher-mediated assessment of a collaborative assignment	310
Figure 4-13 Designing a collaborative assignment assessment scheme	317
Figure 4-14 The materials developer-practitioner-researcher	323
Figure 4-15 Evaluation of the work using Glaser's 1978 criteria	327

Abstract

In a complex world, collaboration is increasingly valued as a global skill (Oxford University Press, n.d.) to help address the multitude of challenges we face. This has seen collaboration become a valued graduate attribute (World Economic Forum, 2016), and thus a common Higher Education (HE) learning outcome. Collaborative assignments are used across faculties and disciplines, to facilitate the development of collaborative skills.

English for Academic Purposes (EAP) serves a preparatory role in readying students for their academic studies and future careers (Ding and Bruce, 2017; Gillett, 2017; Hyland and Hamp-Lyons, 2002). A part of this is readying students for collaboration, so EAP programmes also utilise collaborative assessments, such as group essays and group presentations (Bolster & Levrai, 2022). As well as meeting desired learning outcomes, collaboration brings benefits in terms of learning gains, academic socialisation and increased learner autonomy (Dörnyei and Malderez, 1997; Storch, 2013; Storch, 2021, Thomas, 2014; Wigglesworth & Storch, 2009; Wilson et al., 2018). That is not to say collaborative assignments are not without tensions, with free-riding (where a member of a group makes no contribution but benefits from the work of others) being a prime example (Strijbos, 2016).

However, the major source of tension in collaborative assignments is assessment, which is an area fraught with difficulty (Berry, 2007; Meijer et al., 2020; Strauss & U, 2007; Strijbos, 2016). There is no agreed approach to assessing collaborative assignments, with basing assessment solely on the product, complementing the collaborative assignment with individual assignments, and assessing the collaboration itself all possibilities (de Hei, Strijbos et al., 2016; Meijer et al., 2020). Solely assessing the product has been critiqued as unfair as students may have put different amounts of effort or contribution toward the assignment yet receive the same grade (Meijer, et al., 2020). A product-only approach also fails to recognise learning about collaboration that could have taken place but is not reflected in the product (Boud and Bearman, 2022). Individual assessment in a collaborative assignment has also been critiqued for potentially misaligning student behaviour and compromising

collaboration (Meijer et al., 2020). Assessing collaboration itself probably carries the most concerns due to an under-conception of what is being assessed and a lack of clear means of doing so (Gammie and Matson, 2007; Johnston and Miles, 2010). Potential assessment lenses to bring the collaboration into focus, like self- and peer- assessment, also come loaded with challenges, particularly when it comes to summative assessment (Leach, 2012, Meijer et al., 2020; Rasooli et al., 2024).

This dissertation charts a constructivist Grounded Theory (GT) (Charmaz, 2014) study into EAP practitioner beliefs and practices surrounding assessing collaborative assignments. GT was the preferred research approach as both assessment of collaboration and EAP practitioner approaches to collaborative assignments are areas of scant research. This means a GT approach was appropriate since there were no established theories to test (Charmaz, 2014; Hadley, 2017). The focus is particularly on collaborative written assignments as these pose significant challenges. In EAP, a group essay may be developed over a series of weeks, with work taking place inside and outside the classroom, synchronously or asynchronously (Bolster & Levrai, 2022; Storch & Knoch, 2023). This necessarily means the teacher is not going to be cognizant of the bulk of the collaboration that takes place and will also not be fully aware of every contribution made by students to their group, particularly when it comes to intangible contributions, like maintaining motivation or group harmony.

Experienced EAP practitioners working in a variety of contexts completed a pre-interview task evaluating four models of assessing collaborative assignment and then participated in an in-depth interview. Through this it became clear that they held a range of attitudes and beliefs, with justifications made for and against each assessment scheme. It also brought to light EAP practitioner collaborative assessment identity, where the three key principles of fairness, transparency and simplicity encompass three guiding orientations: reward, practical and defensive. The reward orientation means practitioners are looking for opportunities to reward students for the work that they have done while the practical orientation means a prime concern is that assessment is feasible and manageable for both teachers and students. The most complex orientation, the defensive, speaks

to the need for assessment to be justifiable to a range of stakeholders and is influenced by environment, experience and expectations.

It was also possible to identify collaborative assessment archetypes that encapsulate prevailing attitudes towards collaborative assessment. The Never-everers resist the idea of collaborative assessment, considering it essentially invalid. For I-in-Teamers the individual takes precedence, with the prime concern being the opportunity to reward the individual for what they have done. Differentiation between students is essential for an I-in-Teamer. Real-worlders are product-focused, drawing a parallel with collaboration outside the academy, where it is the product that matters, not how it was produced. Pragmatic Real-worlders also focus assessment on the product, but this is more through necessity than belief. For Pragmatic Real-worlders there is not the time, resource, or means to assess collaboration more closely. The All-I-can-doer accepts their approach to assessing collaboration may not be ideal and is open to change, giving them a close relationship with the archetype inhabited by most of the participants: the Would-if-I-coulder. Would-if-I-coulders know their current assessment practices for collaborative assignments were not fully fit for purpose but were unsure how to improve them. After training and collaborative assessment literacy training, practitioners could shift to inhabit the Best-I-can-doer archetype, delivering principled and rationalised collaborative assessment. The final archetype, not encountered in the data, is the logical end of the spectrum of archetypes. The All-seeing-eye would have full confidence in their ability to assess a collaborative assignment.

Further to the theoretical outputs of this dissertation, collaborative assessment identity and collaborative assessment archetypes, there was also a practical output. This is the collaborative assessment algorithm, which is the considerations and decision trees that have to be navigated when designing a collaborative assessment scheme. There is no simple solution to assessing collaborative assignments, but the process can be made fairer for students through careful planning and consideration of the motivation for and the purpose of a collaborative assignment. The readiness of teachers and students to engage in such assignments needs to be considered and then appropriate

assessment lenses selected to encourage positive collaboration practices and ensure students get the rewards they deserve. This algorithm was actualised in a ‘Starter Guide for Assessing Collaborative Assignments’, which was shared with research participants for review. Feedback on the guide shows such a guide is necessary and the guide as is would help develop EAP practitioners’ collaborative assessment literacy, as well as the assessment literacy of others in the HE sphere, when engaging in collaborative assignments.

The assessment of collaboration is complex, but can be leveraged to encourage students towards positive collaboration and contribute to the development of various learning aims. While there is no simple answer to collaborative assessment there are good questions to ask, which can develop robust, justified and fairer assessment schemes.

Resumen

En un mundo complejo, la colaboración se valora cada vez más como una habilidad global (Oxford University Press, n.d.) para ayudar a abordar la multitud de desafíos que enfrentamos. Esto ha hecho que la colaboración se convierta en un atributo valorado de los graduados (Foro Económico Mundial, 2016) y, por lo tanto, en un resultado común del aprendizaje en la Educación Superior (ES). Las tareas colaborativas se utilizan en todas las facultades y disciplinas para facilitar el desarrollo de habilidades colaborativas.

El inglés para fines académicos (EAP) cumple una función preparatoria al preparar a los estudiantes para sus estudios académicos y carreras futuras (Ding y Bruce, 2017; Gillett, 2017; Hyland y Hamp-Lyons, 2002). Parte de esto es preparar a los estudiantes para la colaboración, por lo que los programas EAP también utilizan evaluaciones colaborativas, como ensayos y presentaciones grupales (Bolster & Levrai, 2022). Además de lograr los resultados de aprendizaje deseados, la colaboración aporta beneficios en términos de ganancias de aprendizaje, socialización académica y mayor autonomía del alumno (Dörnyei y Malderez, 1997; Storch, 2013; Storch, 2021, Thomas, 2014; Wigglesworth & Storch, 2009; Wilson et otros, 2018). Esto no quiere decir que las tareas colaborativas no estén exentas de tensiones, siendo el parasitismo (donde un miembro de un grupo no contribuye pero se beneficia del trabajo de otros) un excelente ejemplo (Strijbos, 2016).

Sin embargo, la principal fuente de tensión en las tareas colaborativas es la evaluación, que es un área plagada de dificultades (Berry, 2007; Meijer et al., 2020; Strauss & U, 2007; Strijbos, 2016). No existe un enfoque acordado para evaluar las tareas colaborativas, basando la evaluación únicamente en el producto, complementando la tarea colaborativa con tareas individuales y evaluando todas las posibilidades de la colaboración en sí (de Hei, Strijbos et al., 2016; Meijer et al., 2020). Evaluar únicamente el producto ha sido criticado como injusto, ya que los estudiantes pueden haber puesto diferentes cantidades de esfuerzo o contribución a la tarea y aún así recibir la misma calificación (Meijer, et al., 2020). Un enfoque centrado únicamente en el producto tampoco reconoce el

aprendizaje sobre la colaboración que podría haber tenido lugar pero que no se refleja en el producto (Boud y Bearman, 2022). La evaluación individual en una tarea colaborativa también ha sido criticada por desalinear potencialmente el comportamiento de los estudiantes y comprometer la colaboración (Meijer et al., 2020). La evaluación de la colaboración en sí misma probablemente genera la mayor preocupación debido a una concepción insuficiente de lo que se está evaluando y a la falta de medios claros para hacerlo (Gammie y Matson, 2007; Johnston y Miles, 2010). Los posibles lentes de evaluación para centrar la colaboración, como la autoevaluación y la evaluación por pares, también vienen cargados de desafíos, particularmente cuando se trata de evaluación sumativa (Leach, 2012, Meijer et al., 2020; Rasooli et al., 2024).

Esta disertación traza un estudio constructivista de la Teoría Fundamentada (GT) (Charmaz, 2014) sobre las creencias y prácticas de los profesionales de EAP en torno a la evaluación de tareas colaborativas. GT fue el enfoque de investigación preferido, ya que tanto la evaluación de la colaboración como los enfoques de los profesionales de EAP para las tareas colaborativas son áreas de escasa investigación. Esto significa que un enfoque GT era apropiado ya que no había teorías establecidas para probar (Charmaz, 2014; Hadley, 2017). La atención se centra particularmente en las tareas escritas colaborativas, ya que plantean desafíos importantes. En EAP, un ensayo grupal puede desarrollarse a lo largo de una serie de semanas, realizándose el trabajo dentro y fuera del aula, de forma sincrónica o asincrónica (Bolster & Levraj, 2022; Storch & Knoch, 2023). Esto necesariamente significa que el profesor no será consciente de la mayor parte de la colaboración que se lleva a cabo y tampoco será plenamente consciente de cada contribución hecha por los estudiantes a su grupo, particularmente cuando se trata de contribuciones intangibles, como mantener la motivación o el grupo. armonía.

Profesionales experimentados de EAP que trabajan en una variedad de contextos completaron una tarea previa a la entrevista evaluando cuatro modelos de evaluación de la asignación colaborativa y participaron en una entrevista en profundidad. A través de esto quedó claro que tenían una variedad de actitudes y creencias, con justificaciones a favor y en contra de cada esquema de evaluación.

También sacó a la luz la identidad de la evaluación colaborativa de los profesionales de la EAP, donde los tres principios clave de equidad, transparencia y simplicidad abarcan tres orientaciones rectoras: recompensa, práctica y defensiva. La orientación de recompensa significa que los profesionales buscan oportunidades para recompensar a los estudiantes por el trabajo que han realizado, mientras que la orientación práctica significa que la principal preocupación es que la evaluación sea factible y manejable tanto para los profesores como para los estudiantes. La orientación más compleja, la defensiva, habla de la necesidad de que la evaluación sea justificable ante una variedad de partes interesadas y está influenciada por el entorno, la experiencia y las expectativas.

También fue posible identificar arquetipos de evaluación colaborativa que resumen las actitudes predominantes hacia la evaluación colaborativa. Los Never-everers se resisten a la idea de la evaluación colaborativa, considerándola esencialmente inválida. Los profesionales del mundo real se centran en el producto, estableciendo un paralelo con la colaboración fuera de la academia, donde lo que importa es el producto, no cómo se produjo. Los pragmáticos del mundo real también centran la evaluación en el producto, pero esto es más por necesidad que por creencia. Para los pragmáticos del mundo real no existe el tiempo, los recursos ni los medios para evaluar la colaboración más de cerca. Para los I-in-Teamers, el individuo tiene prioridad, siendo la principal preocupación la oportunidad de recompensar al individuo por lo que ha hecho. La diferenciación entre estudiantes es esencial para un I-in-Teamer. La mayoría de los participantes eran personas que sabían que sus prácticas de evaluación actuales para tareas colaborativas no eran totalmente adecuadas para su propósito, pero no estaban seguros de cómo mejorarlas. El arquetipo final, que no se encuentra en los datos, es el final lógico del espectro de arquetipos. El Ojo que todo lo ve tendría plena confianza en su capacidad para evaluar una tarea colaborativa.

Además de los resultados teóricos de esta tesis, la identidad de la evaluación colaborativa y los arquetipos de la evaluación colaborativa, también hubo un resultado práctico. Este es el algoritmo de evaluación colaborativa, que son las consideraciones y los árboles de decisión que deben navegarse al diseñar un esquema de evaluación colaborativa. No existe una solución sencilla para evaluar las

tareas colaborativas, pero el proceso puede hacerse más justo para los estudiantes mediante una planificación cuidadosa y la consideración de la motivación y el propósito de una tarea colaborativa. Es necesario considerar la disposición de los maestros y estudiantes para participar en tales tareas y luego seleccionar lentes de evaluación apropiados para fomentar prácticas de colaboración positivas y garantizar que los estudiantes obtengan las recompensas que merecen. Este algoritmo se actualizó en una "Guía inicial para evaluar tareas colaborativas", que se compartió con los participantes de la investigación para su revisión. Los comentarios sobre la guía muestran que dicha guía es necesaria y que la guía tal como está ayudaría a desarrollar la alfabetización en evaluación colaborativa de los profesionales de la EAP, así como la alfabetización en evaluación de otros en la esfera de la educación superior que participan en tareas colaborativas.

La evaluación de la colaboración es compleja, pero se puede aprovechar para alentar a los estudiantes a realizar una colaboración positiva y contribuir al desarrollo de diversos objetivos de aprendizaje. Si bien no existe una respuesta sencilla a la evaluación colaborativa, hay buenas preguntas que formular, que pueden desarrollar esquemas de evaluación sólidos, justificados y más justos.



INTRODUCTION

INTRODUCTION

Collaborative assignments are a feature of Higher Education (HE) and of English for Academic Purposes (EAP) courses. These might include assignments such as students writing an essay or report together, or developing and delivering a group presentation. The purpose of this dissertation is to examine EAP practitioners' attitudes to different approaches to assessing written collaborative assignments and determining an acceptable means of assessing such assignments. The fundamental question is if it is possible to assess collaborative assignments fairly.

The short answer is ... that it is not. A more complete answer is that the assessment of collaboration is widely recognised as a complex area and this complexity is compounded in the HE and EAP context where students may be working on extended collaborative assignments taking place over the course of a semester, working synchronously and asynchronously, both inside and outside the classroom. Consequently, EAP teachers will lack the necessary insight into the collaboration to assess it fairly and the assessment lenses that could be employed to gain greater insight, including student reporting, self-assessment, peer-assessment and reflection, are all problematized in some way. A further complication is that EAP conducts high-stakes summative assessments, which could impact students' progression through university, meaning any assessment scheme has to be robust and justifiable. If collaboration is a learning outcome for a course it should be demonstrable that that learning outcome has been met but the construct of collaboration itself complex and it is questionable if the ability to collaborate is stable and actually measurable.

The central dilemma is that even though the summative assessment of collaborative assignments is problematic, it is also necessary, since collaborative assignments do and should exist in HE and in EAP programmes. The aim of this dissertation is to explore EAP practitioner attitudes to assessing collaborative assignments, consider the factors which impact their collaborative assessment identity, and to generate a collaborative assessment algorithm to guide for EAP practitioners when

they are designing collaborative assignments. Although such assessment will necessitate compromises, the algorithm should allow for adaptable assessment of collaborative assignments, depending on the importance of collaboration as a learning outcome. Absolute fairness may be unachievable but fairer collaborative assessment practices should be and must be the aim.

It is vital to bring the practitioner voice into the assessment of collaboration as this has hitherto been marginalised. As noted in Coffin's (2020) study which looked at student and teacher perceptions of implementing collaborative writing,

Though the process and outcome of CW [collaborative writing] are appreciated, the issues of fairness of teamwork contribution and assessment continue to be questionable. These problematic issues may link to teachers' competence in relation to their roles and teaching strategies which require training. The question of how to prepare L2 [second language] writing teachers to be competent in handling lessons, students, and classroom management in L2 collaborative writing classrooms is the issue that needs to be further developed and studied. (p. 189)

This call for research into collaborative writing and practitioners is to help address the problem identified by Zheng et al., of there being "a scarcity of research on classroom teachers' perceptions and knowledge of collaborative writing, which are likely to determine whether and how teachers implement collaborative writing" (2021, p. 1-2). For student collaboration to be effective, practitioners need to be able to both facilitate it effectively and assess it fairly.

An opening vignette

Before fully entering into the topic of this dissertation, and to locate the discussion that follows, the following vignette illustrates how students might navigate such an assignment and how problematic the assessment of collaborative assignments can be.

Tomas, Maria and Sara are undertaking a foundation year at an English medium university and are taking an English for Academic Purposes course in their first semester. In Week 3 of the 10-week course they are grouped together for a problem/solution essay assignment, deliverable in Week 9.

Week 3

They discuss the problem they would like to address in class and decide to focus on food insecurity. Maria sets up a Whatsapp group so they can stay in touch and they agree to post resources that could be useful to read and discuss later. That afternoon Tomas spends quite a bit of time searching online and finds some potentially useful articles which he posts to the Whatsapp group. Neither Maria nor Sara respond but they have a different class together a couple of days later and talk about the topic and agree it might be better to consider the problem of food waste because they think there are better solutions to evaluate. Sara sends a Whatsapp to Tomas, telling him about the change, and Maria starts a Google Doc where they can write the essay and shares it with the others.

Week 4

In the next lesson, there is some time to talk about the essay and think about the organisation but Tomas does not really contribute much because he is annoyed about the change of topic and the time he wasted looking for resources, which are no longer relevant. He suggests Sara and Maria search for sources and he will start drafting the introduction because he does not want to do any more research. Maria thinks it would be better to outline first but Tomas wants them to get started with the writing and thinks the organisation can come later. Maria insists they need a plan before they write anything. Sara intervenes, saying they have some good ideas so it's ok if Tomas writes some things up, like describing the problem, which they can reorganise when they are editing.

Maria reluctantly agrees but, after the meeting, spends some time thinking about the structure of the essay. She does not do any research but posts to the Whatsapp about what she thinks their general thesis could be and areas they need do research to be able to support their argument. Sara spends quite a bit of time on Google Scholar and the library database, searching for sources and posts a selection of ones that could be useful to the

Google Doc, arranged in the topic areas Maria suggested. Tomas does not contribute anything to the shared document.

Week 5-6

They all get busy with other things and the essay takes a backseat for a while, although Maria does read the articles Sara found and thinks about the organisation of the essay some more. In Week 6, Tomas does write something for the first part of the essay, albeit without any support.

Week 7

They realise the first draft is almost due. Sara has an exam in another course but Maria and Tomas meet in the library, where they start working on the essay in earnest. Maria shares her plan with Tomas and he starts writing while Maria reviews the resources Sara found, finding bits they can use as support. As Tomas is typing, he stops to check some things with Maria, leading to a brainwave for a completely new solution, for which Tomas quickly finds an article they could use as support. The following day, Sara goes through what is written and makes some changes, mainly to do with paraphrasing, adding citations, and putting the reference list together.

Week 8

The feedback from the teacher on the first draft highlights quite a few problems, including both content and language. Tomas does not really agree with the feedback and so does not help with the editing, as he thinks the essay is good enough as it is but agrees to do a final proofread before submitting the essay. Sara and Maria go through the essay together, making some changes to the organisation and use of support, strengthening in the areas the teacher indicated. Sara does the typing while Maria has the teacher feedback open on another computer.

Week 9

Tomas takes a final pass over the essay, checking the language and doing the proofreading and submits the essay, three hours after the deadline, which incurs an automatic 10% penalty.

The vignette exemplifies some of the issues surrounding the assessment of collaborative assignments. How should the assignment be assessed? Should all the students get the same grade? Did they all contribute equally? Based on classroom observation, Tomas did not seem overly involved but looking in the revision history of the Google Doc, all the typing was done by Tomas and Sara, with no record of Maria contributing to the document. Tomas improved the language accuracy but Sara made sure they avoided plagiarism and followed referencing conventions. Maria came up with the main thesis and organisation of the essay, while Sara and Tomas found the support. Sara also headed off a major disagreement, which could have seen a breakdown between Tomas and Maria. Sara and Maria both engaged with the feedback process, while Tomas resisted it. Based on incomplete observation, is the teacher informed enough to differentiate between the students? Should they even try? How can their contributions be weighed? Is it number of words on the page? Number of articles found? Logistics, planning, motivating others? Should they all be penalised for the late submission? Would peer-assessment provide a clear and accurate view of the collaboration? Would self-assessment provide better insight? Would Tomas's understanding of the collaboration be the same as Sara's or Maria's if they were to reflect on it? Would those reflections be a true record of what happened and, if so, true to who?

Researcher motivation

During my undergraduate studies, I had no experience of collaborative assignments, which may be reflective of the time. In my later studies, the capstone assignment of my International Diploma was a collaborative assignment and two modules of my MA were assessed collaboratively. Collaboration has been an important part of my higher education and has also featured on EAP programmes I have taught on and on faculty programmes I have provided support for. Collaboration happens in HE, is necessary, but is under-discussed, especially when it comes to assessment.

When I was initially introduced to collaborative assignments in the form of a group essay, I had a level of scepticism toward them, a scepticism shared by a number of colleagues, in part because it

seemed more like a logistical decision by management than a pedagogical one. However, through reading and experience, the potential gains and benefits of collaborative assignments became apparent. When I moved to another institution, I had joint responsibility for course design of a mandatory course taken by all students to prepare them for English-medium instruction. For that course we developed a group essay assignment, with an accompanying framework to support and scaffold students through the process and help them navigate the collaboration process and the paper “A framework to support group essay writing in English for Academic Purposes: A case study from an English-medium instruction context” is included in Appendix 8 as a motivation study, meaning the study that motivated this dissertation. The paper describes an assessment scheme for the collaborative assignment, which is illustrated in Table 1 of that paper (see Appendix 8, p. 9-11). It was a multi-task assignment to enable differentiation between students and to provide adequate insight into the collaboration. The weighting of the various components is illustrated in Table 0.1.

Table 0.1 Grade allocation for the collaborative essay assignment (Levrai & Bolster, 2019, p. 194)

Component		
Individual grade	Forum discussion contributions	5%
	Annotated bibliography	20%
	Peer feedback	15%
	Group contribution	10%
Group grade	Drafting	10%
	Final essay	40%

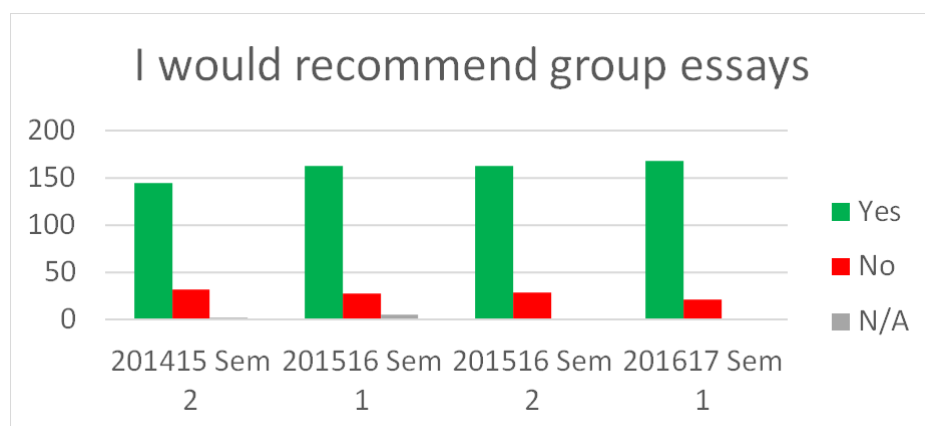
Part of the original motivation for this dissertation was to demonstrate that the above was a suitable assessment scheme for a written collaborative assignment, drawing together assessment of the product, individual complementary assignments, and assessment of the collaboration itself. While the assessment scheme has its merits, a 13-element assignment with a mark scheme across six areas lacks ease of implementation in other contexts and may dissuade people from taking on collaborative

assignments before they even begin. Furthermore, what has become clear through this dissertation, is that there is no single answer or approach to assessing a collaborative assignment and a variety of assessment approaches could be justified, despite all of them being compromised in some way.

Origins of this dissertation

This dissertation grew out of practical experiencing of developing and introducing collaborative assignments in a national university in Asia after being introduced to the concept in a branch campus of a British university, also based in Asia. To gauge the student attitude to the course in the second institution, and the collaborative writing assignment specifically, an end of course survey was conducted over four consecutive semesters (from spring semester 2015 to autumn semester 2016) with four different cohorts of students. That survey showed that the majority of students had a positive attitude to collaborative assignments, which gradually increased as the course was refined. Figure 0-1 shows student response to the question “I would recommend group essays” and there is a clear positive response. Overall, an average of 84% of students would recommend group essays and in the last cohort, from autumn 2016, this figure rose to 87.9% (Levrai & Bolster, 2017).

Figure 0-1 Student attitude to group essays (Levrai and Bolster, 2017)



While this positive response was encouraging, what was more encouraging was the reasons why students did, and did not, recommend group essays. The main reasons students gave in open comments for why they would recommend group essays is that they recognised collaborative writing led to better ideas, better essays and helped them develop soft skills. The most common justification for not recommending group essays was the difficulty in developing a coherent essay when there are multiple voices (reflecting good awareness of what makes for an effective essay) and a recognition that group essays were a beneficial experience but students just preferred writing on their own (Levrai & Bolster 2019).

That study was limited however, as it only surveyed the students at the end of the course so any changing attitudes could not be tracked. There was no initial baseline to see how the experience of the assignment affected student attitude toward collaboration, if at all. The question regarding attitude to collaborative assignments was also presented as a binary yes/no answer rather than a Likert scale, which would have shown more subtleties, and there were no follow-up interviews or other research tools used to probe deeper into the student reasoning. The teacher voice was also missing in terms of their attitude to collaborative assignments. Through experience, it was clear there was a mix of attitudes amongst teachers when it comes to collaborative assignments. Collaborative assessment can engender particularly strong feelings and, indeed, scepticism if not outright resistance. Nonetheless, these initial survey results identified collaboration and collaborative assignments as an area warranting further study, leading into the current research project.

Research focus

As will be discussed in 2.3, the research plan leading to this dissertation has undergone significant changes since its original conception, not least because of the impact of the COVID-19 pandemic. What has remained constant, however, is the desire to better understand how to most effectively and fairly assess collaborative assignments in EAP contexts. Allied to this is the covert intention to identify how collaborative assignments can be more fairly assessed within HE more

generally, since a mechanism of assessing collaborative assignments in EAP should have utility in faculty assessments as well.

Understanding the issue

Grounded Theory (GT) (Glaser & Strauss, 1967) was selected as a means of examining the issue since GT allows for the exploration of under studied areas and a constructivist GT approach (Charmaz, 2014) was adopted. Since the summative assessment of collaborative assignments in EAP is an area that receives scant attention in the research literature it became necessary to search for answers in the field, with the intent of emerging with a theory of practical value to the area under question.

Data collection comprised an initial survey, a pre-interview task, in-depth interviews, and follow-up email correspondence with EAP practitioners working in different contexts across Europe, the Middle East and Asia. This variety of participants illuminated different approaches to facilitating and assessing collaborative assignments, as well as commonalities, particularly in terms of an appetite to know more.

Though the assessment of such assignments will always contain an element of compromise, that compromise has to be navigated as fairly as possible. While there is no single assessment approach that fits a collaborative assignment in EAP, a multi-lens approach to assessment could give greater insight into the opaque construct of collaboration, incorporating a range of assessment tools like self- and peer-assessment, reflection and reporting. While this would not make the process of collaboration fully transparent, it could give confidence to a teacher that they are assessing more fairly. However, there is also a strong argument for not assessing collaboration itself at all. A conclusion of this dissertation is that there are justifications for a light or heavy assessment approach to collaborative assignments, depending on the learning outcomes of the course. What is consistent, however, is that the rationale for the assessment approach, and what it entails, has to be transparent and rationalised to the teachers involved before implementation and to students from the outset.

When considering collaborative assignments in EAP, it is also important to note the quote which opens this dissertation from van der Vleuten et al. (2015), which looks not at assessments in isolation but across a programme, therefore developing a programmatic assessment scheme “that optimised both the learning function and the decision making,” (Van der Vleuten et al. (2015, p. 641). Rather than deciding how to assess a collaborative assignment by itself, it is necessary to consider how that assignment fits into the full range of assessments and learning outcomes for a particular course and programme of study.

The purpose of this dissertation is to better understand the concerns of EAP practitioners in relation to assessing collaborative assignments and to develop a practical tool which will help them navigate the process. The ultimate aim is to dispel some of the tensions around collaborative assignments by better equipping practitioners to set up a clearly rationalised and justifiable assessment scheme, depending on the learning outcomes of the course in question.

Overview of this dissertation

Chapter 1 of this dissertation provides an overview of the contextualising literature for this dissertation. It considers the nature of collaboration and it locates collaborative learning in Sociocultural Theory (Vygotsky, 1978). It moves on to discuss collaborative learning in education and language learning, before looking at the rationales for using collaborative assignments in EAP courses, providing the twin motivation of learning gains and employability. There is then clarification of the meaning of ‘collaboration’ in this dissertation, referring to an indivisible piece of work each group member has equal ownership of and responsibility for. The nature of EAP is then explored, charting the development of the conception of EAP and its role in HE institutions. This incorporates how EAP teaches language development alongside, or arguably subservient to, academic literacy and transferable skills. Consideration will then turn to collaboration in EAP contexts, drawing on the background study to this dissertation, highlighting the gap between the conception of collaboration in the literature and the conception held by EAP practitioners.

The chapter then turns to the issue of assessment, parsing the concepts of Assessment of Learning, Assessment for Learning and Assessment as Learning. From the latter two concepts, evaluative judgement will be explored, with self- and peer- assessment embedded within that. Assessment will be discussed in relation to sociocultural theory and collaborative assessment will be further located at the nexus of classroom-based assessment, authentic assessment and dynamic assessment. The particular challenges of assessing collaboration will be reviewed, as will the concept of practitioner assessment literacy and assessment identity. The chapter closes with a review of the methodological approach of this study. It will introduce the GT approach and its core concepts. Attention will also be paid to critiques of GT and the rationale for using GT in this study.

Chapter 2 explains the philosophical grounding of the study and the methodology followed. It opens with an explanation of the methodological eclecticism of the researcher, mirroring a similar eclecticism in teaching approach, driven by a pragmatism to use the most appropriate approach in a given situation. It also outlines the philosophical framework of the study and how it shifted from postpositivist to social constructivist. Considerable time is spent discussing the evolution of the study from what was principally a postpositivist validation study to a more sociocultural exploration. This is important to show the shift in focus and approach of the study, which has implications for how research questions evolved and hypotheses developed. The impact of the COVID pandemic will also be discussed, recognising how that encouraged the turn to a GT approach and how the research was conducted in practice. There will be mention of the complementary study by Averil Bolster, as this dissertation is part of a wider study into collaborative assessment in EAP and her research data serves as secondary data in this dissertation. The ethical concerns will also be explained, with an overview provided of the technological aids used during this dissertation.

The chapter will then describe the data collection process, including the different stages of sampling, from convenience to theoretical and the research tools utilised in the dissertation will also be described. The participant profiles will also be introduced. Once there is a clear overview of the data collection process, the chapter will go on to discuss data analysis. As GT was adopted as the

methodological approach, memoing played a central role and the various rounds of coding, from open to focused to theoretical will also be explained and the importance of diagramming as a part of theorising highlighted. The chapter closes by discussing the last round of data collection, conducted after the majority of the theorising for this dissertation had been completed. One of the outputs of this dissertation is a practical starter guide to help EAP practitioners design assessment schemes for collaborative assignments and the final data collection was gathering feedback on the guide from participants.

Chapter 3 moves on to the Findings. It opens with participant discussion of the role of EAP, which demonstrates a clearly informed cohort, conversant with major themes and concerns explored in the literature about EAP. There will then be discussion of their understanding of collaboration and its use in the EAP context. Attention will then turn to the participant evaluation of the pre-interview task, which asked them to consider four different approaches for assessing a collaborative assignment. This brings to the fore some of the main concerns and considerations of EAP practitioners when it comes to assessing collaborative assignments, not least of which their desire for fair assessment practices.

The chapter will then move on to discuss the three key principles EAP practitioners hold in relation to collaborative assignments, those of simplicity, transparency and fairness. These in turn encompass three guiding orientations: reward, practical and defensive. The reward orientation encourages practitioners to ensure students benefit from and get due credit for the work they do, while the practical orientation directs attention to ensuring assessment is manageable and reasonable for teachers and students. The defensive orientation is the most complex construct, developing from practitioners' experience, expectations and the HE environment in which they operate, and raises the importance of assessment being robust and justified. The defensive orientation also leads to three behaviours when it comes to collaborative assignments. Defensive avoidance refers to not doing something in anticipation of problems it could bring, defensive acceptance means acknowledging

there could be a problem but continuing anyway, and defensive adaption means changing approach, either pre-emptively in light of anticipated problems or retrospectively, after experiencing challenges.

The chapter then turns to discussion of the pertinent findings of the companion dissertation (Bolster, 2025), with attention drawn to issues of particular relevance to this dissertation, such as the jagged expertise of practitioners when it comes to collaborative assignments and the need for a practitioner-led pedagogy of collaboration. The chapter closes with reviews of the draft starter guide that was developed as product of this research, which was evaluated by participants in terms of need, coverage and usefulness. As will be seen, the guide was well received and speaks to a clear need.

Chapter 4 contains the discussion and theory generation of the dissertation. The principles and orientations identified in Findings are brought together to shape the theoretical output of this dissertation: EAP practitioner collaborative assessment identity, a novel construct of this dissertation. So too are eight collaborative assessment archetypes identified, meaning teachers who would share the same approach to assessing a collaborative assignment placed on a continuum of those who feel collaboration should never be assessed (the Never-everers) to those who are fully confident they are able to assess collaboration accurately and fairly (the All-seeing-eyes). Note, the latter archetype is a purely theoretical construct not encountered in the participants in this dissertation, but which would be the logical opposite end point to the Never-everers.

After proposing the theoretical outputs of this dissertation, collaborative assessment identity and collaborative assessment archetypes, attention turns to the practical output, the collaborative assessment algorithm. This outlines the five layers of considerations that are needed before embarking on a collaborative assignment, starting with fundamental issues like what is motivating the use of collaborative assignments and, in layer two, how the results of the assessment are going to be used. The third layer concerns the readiness of teachers and students to engage with collaborative assignments and the fourth leads to selection of appropriate assessment lenses, incorporating assessing the product, complementary individual assignments and assessing the collaboration itself. Such a selection involves navigating the decision-trees central to designing an assessment scheme for

a collaborative assignment, guided by the importance of collaboration as a learning outcomes and other course learning outcomes. The final layer of the algorithm is evaluation in which the theoretical output of this dissertation, the dimensions of collaborative assessment identity, form the evaluation criteria of collaborative assignment assessment schemes. The collaborative assessment algorithm is then applied to the opening vignette, demonstrating how it could be used in practice.

The algorithm was actualised as a starter guide to assessing collaborative assignments, which was well received by participants and indicates the value of such a work both within EAP and across HE more generally. The chapter closes with evaluation of the theoretical and practical outputs of this dissertation, considering them through a *prima facie* lens and through classic GT criteria before moving into the main evaluation of the work through the constructivist criteria of credibility, originality, resonance and usefulness.

Chapter 5 concludes the dissertation, summarising the main findings. It then reflects on the research questions and hypotheses of this dissertation, and on those from the original postpositivist research plan. The chapter goes on to discuss the limitations to the dissertation, not least of which is the absence of the student voice. It then points to future directions of research, which would involve determining if collaboration is a static construct which can be assessed, how teacher training interventions could impact a practitioners' collaborative assessment identity, and how students may evaluate differing models of assessment for collaborative assignments. There is then discussion of the implications and significance of this research, highlighting the rich data gleaned from EAP practitioners and the value of the collaborative assessment algorithm to EAP and the wider HE community. It concludes with reflection on the need to better understand collaborative assessment and the contribution this dissertation has made to that aim, bringing the previous side-lined practitioner voice forward as a necessary step in developing our understanding of the assessment of collaborative assignments.

There are 13 appendices attached to this dissertation, accessible through a secure UPV/EHU shared folder - Peter Levrai PhD Appendices

- Appendix 1 contains the ethical approval and data collection tools used in this dissertation
- Appendix 2 contains the survey data, redacted to protect anonymity of participants
- Appendix 3 contains the pre-interview task data, also redacted for anonymity.
- Appendix 4 contains the interview transcripts for the interviews conducted as part of this dissertation.
- Appendix 5 contains the secondary research data of transcripts of the interviews conducted by Averil Bolster.
- Appendix 6 contains the open, focused and theoretical codebooks.
- Appendix 7 contains the data collection tools for the original study.
- Appendix 8 contains two articles, charting the motivation and background studies.
- Appendix 9 contains a draft version of the starter guide to assessing collaborative assignments.
- Appendix 10 is the consent form for new participants for the final round of data collection reviewing the starter guide.
- Appendix 11 contains the reviews of the starter guide.
- Appendix 12 contains the full size graphic Figure 4-13 of the collaborative assessment algorithm in action.
- Appendix 13 contains the revised version of the starter guide, edited after initial feedback.



CHAPTER 1

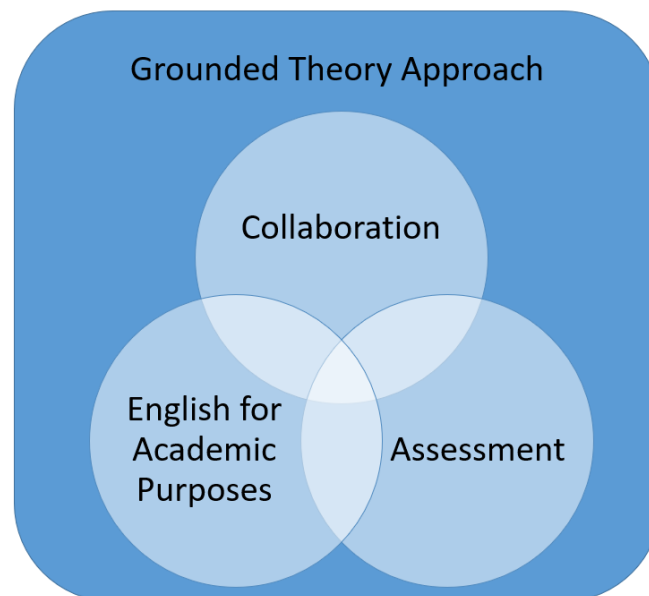
LITERATURE REVIEW

1 LITERATURE REVIEW

1.1 Chapter overview

This chapter will establish the knowledge framework for this study and will address three main areas: collaboration, EAP, and assessment (see Figure 1-1). These will be examined independently and the connections between them elucidated. There will also be discussion of the Grounded Theory theoretical framework through which these ideas will be explored in the context of this dissertation.

Figure 1-1 *Theoretical framework for examining three core concepts*



The chapter will begin by establishing the importance of collaboration and its role in modern HE settings. The rationale for using collaboration in university courses will be discussed, considering at both the employability and pedagogical rationales. As part of the learning theory rationale, sociocultural theory (Vygotsky, 1978) will be examined and the most pertinent aspect for the purpose of incorporating collaboration in education – the Zone of Proximal Development – will also be reviewed. After consideration of *why* students working together is beneficial, attention will turn to *how* students can work together, clarifying the meanings of “groupwork”, “cooperation” and

“collaboration”. Specific attention will be given to collaborative writing, providing an overview of the learning gains.

Once there is a clear conception of collaboration there will be an overview of the EAP context, considering the role of EAP in supporting student’s language and academic literacy development, as well as their academic socialisation. There will be consideration of EAP as a mature academic field and then the role collaboration in EAP will be examined, drawing on the background study to this dissertation, leading into a definition of collaborative assignments in the EAP context.

After the role of collaboration in EAP is understood there will be a review of the relevant assessment literature, looking at the concepts of assessment of, for and as learning. There will be discussion of evaluative judgement, encompassing self- and peer- assessment. Assessment will also be examined through a sociocultural lens, which will lead into discussion of dynamic assessment, classroom-based assessments, and authentic assessment, locating collaborative assessment at the nexus of those assessment approaches. Attention will turn to the assessment of collaboration specifically, which is a complex area and the review will help explicate the challenges to assessing collaboration, as well as identifying tentative proposals from the literature for navigating collaborative assessment successfully.

As an assessment agent, teachers play an essential role in assessments, so attention will turn to teacher’s relationship with assessment, considering both assessment literacy and assessment identity as the success of a collaborative writing approach will depend in part on the competence and knowledge of practitioners to implement and assess it effectively. The chapter closes with a review of the methodology employed in this study – Grounded Theory. It will introduce the theory and expand on four key concepts: theory, emergence, theoretical saturation and the process of conducting a Grounded Theory study. There will also be consideration of the critiques of Grounded Theory and a final rationale for its utilisation.

1.2 Collaboration

We live in a complex world, with amazing possibilities. Project CETI (Cetacean Translation Initiative) is developing the equivalent of Google Translate for whales, which would allow us to understand and, potentially, communicate with the largest mammals on the planet. What sounds like pure science fiction is a collaboration of “artificial intelligence and natural language processing experts, cryptographers, linguists, marine biologists, roboticists and underwater acousticians” (projectceti, 2023, para. 2). It is through collaboration that the impossible becomes tantalising possibility. It is how planes get built, how companies function, how great music gets made and shared with the world. There are very few truly individual pursuits or careers, making collaboration a fact of life.

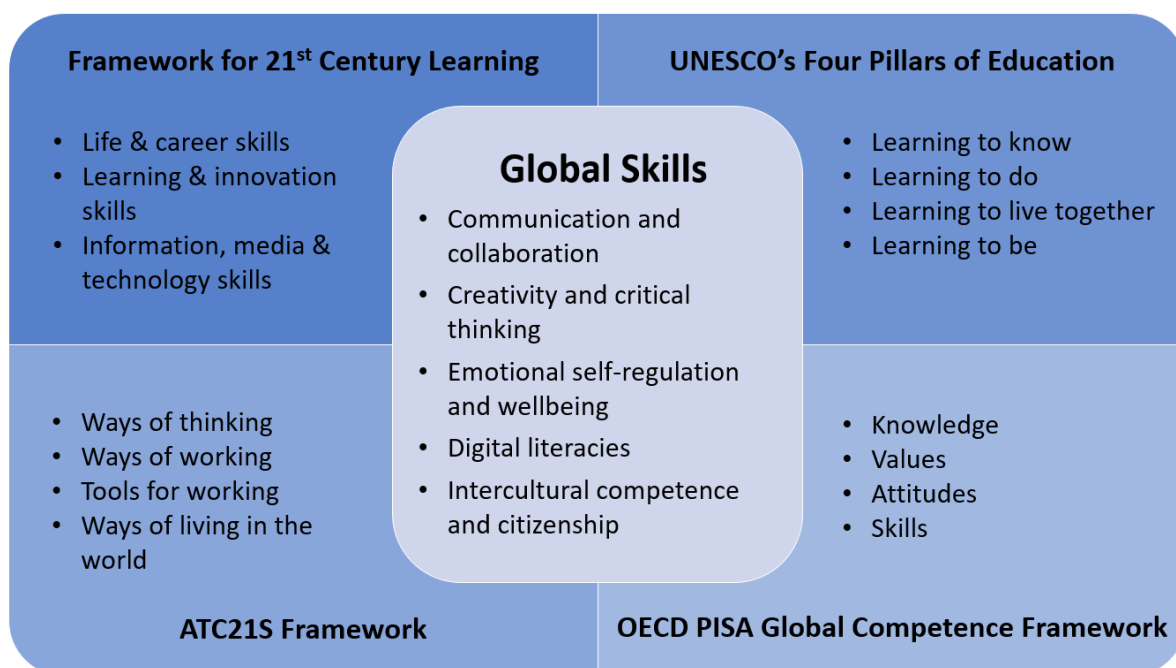
1.2.1 Collaboration as a valued global skill

One of the aims of education is preparing students for work and life. This is the cause of some tension, with concern over the neoliberalisation of education (Ding & Bruce, 2017; Hastings, 2019), but it is now common for university programmes to indicate graduate attributes. These attributes are “the qualities, skills and understandings a university community agrees its students would desirably develop during their time at the institution” (Bowden et al., 2000, as cited in Wong et al, 2022, p. 1341). These graduate attributes may be referred to as generic or soft skills and came under the umbrella of 21st Century Skills, with the World Economic Forum (2016) recognising that

to thrive in the 21st century, students need more than traditional academic learning. They must be adept at collaboration, communication and problem-solving, which are some of the skills developed through social and emotional learning (SEL). Coupled with mastery of traditional skills, social and emotional proficiency will equip students to succeed in the swiftly evolving digital economy. (p. 4)

Not only did collaboration come first in the above quote, it was also identified as one of the key 4Cs, which also include creativity, critical thinking, and communication, which were central to the Framework for 21st Century Learning (Battelle for Kids, 2019). While the Framework for 21st Century Learning was focused principally on employability, it has been integrated with UNESCO's 4 pillars of education, the OECD PISA Global Competence Framework, and the ATC21S Framework, to form Global Skills (see Figure 1-2) which look beyond the world of work to the whole person.

Figure 1-2 *Global Skills and their four foundational frameworks (adapted from Oxford University Press (n.d.))*



Collaboration runs through all of the frameworks which build into the Global Skills, and it holds a central role, as explained in the Global Skills position paper.

Collaboration refers to the ability to work effectively with others towards shared goals. It requires individuals to demonstrate openness to learning from others as well as to sharing their own resources with others. As such, it is a cornerstone of communicative ELT, which typically builds language skills through learners working together on activities designed to achieve communicative goals. (Oxford University Press, n.d., p. 8)

While the position paper was written to demonstrate how Global Skills could be integrated into English Language Teaching (ELT), it also contends that the Global Skills, “can and should be taught and assessed in every educational setting. Transferable across subjects and to life outside school, global skills are for *all* learners, whatever their educational context” (Oxford University Press, n.d., p. 5, emphasis in original). Given the importance of collaboration, it has become a stated learning outcome and collaborative assignments are a feature of HE at undergraduate (Plastow et al, 2010) and post-graduate (Gammie & Matson, 2007) level, across degree programmes as diverse as social psychology (Johnston & Miles, 2004) to law (Berry, 2007).

1.2.2 Collaborative learning in education

Although there is a lot of discussion about graduate attributes, the rationale for collaboration in education is not solely related to employability. Panitz and Panitz (1998) considered collaboration a personal philosophy which would inform a teacher’s practices both inside the classroom and through other aspects of their lives (p. 162). In contrast, Smith and MacGregor (1992) consider collaborative learning as

an umbrella term for a variety of educational approaches involving joint intellectual effort by students, or students and teachers together. Usually, students are working in groups of two or more, mutually searching for understanding, solutions, or meanings, or creating a product. (p. 11)

and, indeed, their chapter discusses a range of different approaches to collaborative learning, encompassing cooperative learning, peer teaching, simulations and writing groups. Through this range of activities, the common thread is that students are working together and learning from each other. This belief, and Panitz and Panitz's contention that collaborative learning is "the most effective form of learning" (1998, p. 161), has its root in sociocultural theory.

1.2.3 Sociocultural Theory

This dissertation is grounded in Sociocultural Theory (SCT), which posits that learning is a social process, meaning that it is through interaction with others that we learn. Our understanding of the world is informed by socialisation and culture. SCT grew from the work of Vygotsky, which located social interaction as the place where learning occurs. As Vygotsky wrote in *Thinking and Speech*, it is a "well known fact that with collaboration, direction, or some kind of help, the child is always able to do more and solve more difficult tasks that he can independently" (Rieber & Carton, 1987, p. 209).

A core concept of SCT is that our understanding of ourselves, the world, and others is mediated through symbolic tools or artefacts, such as mathematics, art or language systems (Lantolf, 2000). As illustrated in the Ted Talk *How language shapes the way we think* (Boroditsky, 2017), the language system we use influences how we perceive ordinal directions, colours and the very nature of things. This demonstrates how our understanding of the world is fundamentally impacted by the mediation of language. As discussed in the Ted Talk, a German speaker may describe a bridge as "beautiful" or "elegant", contrasting with a Spanish speaker who may use words like "long" or "strong" instead. It is not the bridge which differs, but the language culture, where "bridge" is a feminine noun in German but a masculine noun in Spanish, changing the way speakers of those languages perceive and describe things. We are not born with a gendered understanding of a bridge, but learn it through the language we speak and culture we are brought up in. With language as a mediator between us and our understanding of and interaction with the world, it plays a pivotal role in our development, acting as "the carrier of what Vygotskians see as the powerful scientific concepts that allow people to think

beyond the limits of their immediate everyday experiences and to work collaboratively with others to build new knowledge” (Edwards & Daniels, 2004, p. 108).

1.2.3.1 Zone of Proximal Development and the More Knowledgeable Other

SCT is a rich field, which has led to Activity Theory (AT) (Bakhurst, 2009), been used as a justification for task-based learning (Ellis, 2000), and which has had varied impacts on our understanding of second language acquisition, the breadth of which can be seen in the 35 chapters of the Routledge Handbook of Sociocultural Theory and Second Language Development (Lantolf et al., 2018). Of most relevance to this dissertation is the concept of the Zone of Proximal Development (ZPD), which is not atypical. As Chaiklin (2003) and Ohta (2013) recognised, the ZPD “is arguably the most widely applied construct from sociocultural theory, not only in educational research, but also in inquiry related to second and foreign language acquisition” (p. 648).

The ZPD is often described as the potential of what an individual can achieve, if supported by someone else or a cultural artefact (Lantolf, 2000). For example, there is only so far we can go with mental calculation, but with the aid of pen and paper we can go further and with the aid of a calculator we are able to process even more complex mathematical problems. Exascale computers hold vast computing power that can be turned to complex problems across different fields to help with analysis of complex systems such as viruses or hurricanes (Scoles, 2023). Tools allow us to surpass our unaided abilities and so too does learning from someone who knows more than us. Education systems are built on the concept of a teacher helping a student unlock their potential. In terms of development, we do things with aid until we internalise how to do it so, in future, we will be able to do those things unaided.

Vygotsky’s proposal of the ZPD arose as a critique of traditional IQ testing, arguing traditional IQ testing simply provided a static snapshot of matured knowledge and proposing the ZPD as an alternative. The ZPD focused on the potential of children and explained why students who scored highly on an IQ test may perform below expectations in future education compared to a student who had an initially lower IQ score (Vygotsky, 1978). The high scoring student could have been operating

near the top end of their potential, with a small ZPD, while the lower scoring student could have a wider ZPD and would consequently be able to develop further (Fani & Ghaemi, 2011).

While Vygotsky initially envisaged the ZPD in relation to child development, it is now understood to be applicable across education settings, including adult education (Ohta, 2013). It has also been appropriated in discussion of teacher education, with the zone of proximal teacher development (ZPTD) (Warford, 2011), emphasising “a three-way conversation that places teachers’ prior experiences as learners and often tacit beliefs about pedagogy into conversation with pedagogical content of the teacher education program and observations of teaching and learning in field placements” (p. 252-253) as the means to teacher development.

Another change in the conception of the ZPD is that while initially it described a novice/expert relationship, with the novice being able to achieve more through working with and learning from an expert (Lantolf & Pavlenko, 1995), it is now understood that the ZPD can also be engaged during interaction between two true peers (Ohta, 2013). This means, when working in groups, students can learn from each other as they will have differing expertise and each can take on the role of expert, or More Knowledgeable Other (MKO), at different times. This can lead to a collaborative scaffold, where students support each other through their learning (Donato, 1994; Ohta, 1995). In this way the concept of the ZPD can be expanded beyond the individual to the potential of the whole group, generating a group ZPD (Nyikos & Hashimoto, 1997). In Chaiklin’s (2003) analysis, what matters is not so much the competence of the MKO so much as how the Other can assist in the development of the learner, which highlights the importance of learners helping each other in their development. A particularly relevant examination of the ZPD was conducted by Smagorinsky (2018), which emphasised the long-term development which could take place through the ZPD. This is especially pertinent in the field of EAP where the aim is not so much to get students through the language course in question but rather to equip students for their future academic studies and the working lives which follow.

1.2.4 Collaboration in language education

As Storch and Wigglesworth (2007) recognise, there is a social constructivist rationale for students engaging in pair and group work in language education. Within the field of language learning, the move to communicative language teaching (CLT) saw a shift from teacher centred classrooms to ones in which students were working more often together through, “games, role plays, pair and other small-group activities,” (Savignon, 1991, p. 265). As approaches to language teaching have evolved, group work remains a vital element, be it in the form of student-centred learning (O’Neill & McMahon, 2005), task-based learning (Littlewood, 2004), or team-based learning (Michaelsen & Sweet, 2008).

An important issue to recognise is that although there are affordances generated by students working together, there are different ways in which students can do so and this has an impact on the potential learning which can take place. Considerable effort has been put into parsing the difference between key terms which can seem synonymous and be used interchangeably: group work, cooperation and collaboration. Indeed, Meijer et al. (2020) place “cooperative learning, collaborative learning, group work, peer learning, team learning, group learning activities, and group-based learning” (p. 1222) under the umbrella term of ‘collaborative learning’. For the purposes of this dissertation, however, the distinction between terminologies is important as it has significant implications for assessment.

1.2.4.1 *Groupwork*

Groupwork encompasses the widest range of ways in which students can work together and “is one of the most widely used and deeply researched teaching approaches in the college classroom” (Wilson et. al, 2018). The importance of groupwork was also emphasised by Stevick (1980), who prioritised the importance of “what goes on inside and between people in the classroom” (p. 4) above other factors like materials and teaching techniques when it came to successful learning taking place.

When students do engage in groupwork there are widely discussed benefits. Dörnyei and Malderez (1997) recognised how students can motivate and support each other, can help guide and

enculturate each other, and how group work widens the resource pool available to all students, providing additional opportunities for learning to take place. Wilson et al. (2018) point to students' higher achievements, deeper learning, and improved communication skills as clear benefits of group work.

When thinking about groupwork, we could be inclined to think of any activity students engage in together, be that a group discussion, reading circle or project. However, not all interactions are groupwork and Jaques identified eight features which helped distinguish groups from random collections of people or interactions (see Table 1.1).

Table 1.1 *Eight features of groups (tabulated from Jaques, 2000, p. 1-2)*

Collective perception	members are collectively conscious of their existence as a group.
Needs	members join a group because they believe it will satisfy some needs or give them some rewards.
Shared aims	members hold common aims or ideals that to some extent bind them together. The achievement of aims is presumably one of the rewards.
Interdependence	members are interdependent inasmuch as they are affected by and respond to any event that affects any of the group's members.
Social organization	a group can be seen as a social unit with norms, roles, statuses, power and emotional relationships.
Interaction	members influence and respond to each other in the process of communicating, whether they are face-to-face or otherwise deployed. The sense of 'group' exists even when members are not collected in the same place.
Cohesiveness	members want to remain in the group, to contribute to its well-being and aims, and to join in its activities.
Membership	two or more people interacting for longer than a few minutes constitute a group.

While this is a good starting point to try and focus the concept of groupwork, further work has been done to distinguish between two broad approaches students could take to complete a group

work task: cooperative or collaborative. In Panitz's analysis, cooperation is more product orientated than collaboration and has a greater role for the teacher in setting the parameters of the group to enable them to achieve the desired end result. Collaboration is seen as a more student-centred endeavour and a more process orientated construct (Panitz, 1999). Paulus (2005) refines this distinction, drawing on the work of Henri and Rigault (1996) and that of Roschelle and Teasley (1995), to see cooperation as an approach where students can divide a task into discrete sections and take responsibility for 'their' individual section. In contrast, collaboration encompasses a closer working relationship where students engage in shared creation and hold a shared responsibility for the final product (Paulus, 2005). This distinction between responsibility for the whole problem (collaboration) and a divide and conquer approach (cooperation) is also recognised by Hathorn and Ingram (2002) with them seeing collaboration as providing an opportunity for a group to take "advantage of *all* the participants" (p. 33 emphasis in original). Within the field of L2 research, Storch (2013) has also done considerable work in refining our understanding of collaboration, making a similar distinction between cooperation (a process where students have distinct roles and responsibilities) and collaboration (a process where there is a shared responsibility and ownership of the work throughout). As Hathorn and Ingram recognise advantages to a collaborative approach, so does Storch, illustrating how collaborative students have a higher level of mutuality and produce better quality work while cooperative students have a lower level of mutuality and do not really engage with each other's ideas, which limits their ability to co-construct (Storch, 2012). It is therefore a collaborative approach which, "maximises the opportunities for language learning" (Storch, 2019, p. 40).

These distinctions may seem pedantic from a lay perspective but do have significant implications for assignment design, facilitation and assessment. For example, the assessment of a cooperative task is potentially fairly straightforward. If students take responsibility for a particular 'part' of the task, they can be individually assessed on that part. To encourage the cooperative element of the task, some parts could be co-constructed (e.g. the group write the introduction and conclusion together) but sections/paragraphs could be written separately. While this would reduce the

opportunity for learning from each other it would provide a transparent assessment scheme and clear opportunities for differentiation between students. In contrast, collaborative tasks result in an indivisible, cohesive whole which makes individual assessment much more problematic, which is part of the focus of this dissertation.

1.2.5 Collaborative writing

The main focus of this study is the assessment of collaborative written assignments. For the purpose of this project, a collaborative written assignment refers to an assignment where students work together in a group to produce a joint piece of work for which they share responsibility e.g. an assignment which requires the students to take a collaborative approach. Storch and Knoch (2023) neatly summarise the three defining features of this kind of task in L2 writing (see Table 1.2).

Table 1.2 *Defining features of collaborative writing tasks (Storch & Knoch, 2023, p. 229)*

- | |
|--|
| <ol style="list-style-type: none"> 1) A single co-authored text that cannot be easily reduced to the separate input of individual writers. 2) Substantive involvement by all co-authors in all stages of the text co-construction process. 3) Shared decision making and responsibility for the text produced leading to a sense of collective ownership of the joint text. |
|--|

Student collaboration on writing tasks has been found to have a range of benefits, including the development of collaboration skills (Thomas, 2014; Wigglesworth & Storch, 2009) and improved oral skills due to the interactive nature of the task (Storch, 2013). Indeed, Storch (2021) went on to argue that a collaborative task has more benefits for participants than dedicated writing or speaking tasks and provide more opportunity for language learning. This concurs with the findings of Fernández Dobao and Blum (2013), who investigated the attitude of intermediate level learners of Spanish to writing in pairs and small groups and found that in groups there were more ideas to exchange, which provided opportunity for more language practice and development. The students also felt that the

quality of the group writing was better than they would have been able to achieve alone (Fernández Dobao & Blum, 2013), which is in line with the findings of Levrai and Bolster (2019) in the motivation study to this dissertation, exploring attitudes of EAP students in Macau to collaboration.

The student perception that they generated better work when collaborating is supported in the literature, where it has also been found that student collaboration leads to better quality writing (Shehadeh, 2011; Wigglesworth & Storch, 2009) with higher levels of accuracy (Fernández Dobao, 2012; Storch, 2005) and better task fulfilment (Mulligan & Garofalo, 2011; Talib & Cheung, 2017), suggesting that collaborative writing is productive and effective. Elabdali (2021) conducted a meta-analysis of studies examining the learning effects in collaborative writing tasks in L2 and the performance of students in individual writing tasks after having taken part in collaborative writing tasks. Although it was hard to come to a firm conclusion regarding benefits in terms of writing fluency or complexity, the meta-analysis found a clear benefit in terms of accuracy and also that students who have collaborated are able to go on to produce better quality individual writing (Elabdali, 2021), who went on to note that collaborative writing “might afford learners opportunities to foster cultural reflection (Angelova & Zhao, 2016), learn discipline-specific content (Rice, 2007), and navigate group projects (Ede & Lunsford, 1992), all crucial skills in their future personal and professional lives” (Elabdali, 2021, p. 13).

That is not to say collaboration is without tensions and, as Meijer et al. (2020) note in their recent review of the state of the art of collaborative learning in HE, these tensions stem from a lack of teacher guidance through the collaboration and the assessment of it. It is these tensions this dissertation intends to explore.

1.3 English for Academic Purposes in Higher Education

Although collaborative assignments feature across academic disciplines in HE, the focus of this study is EAP. EAP helps prepare students for higher-level education, in terms of language competence, academic skills and academic discourse socialisation. EAP is a feature of universities in anglophone

countries, primarily to help international students, but EAP units are also part of universities operating in non-native English speaking (NNES) contexts (such as: the English Language Centre in the University of Macau, Macau, the Centre for English Language Education in the University of Nottingham Ningbo, China or the Centre for Languages and Communication in the University of Turku, Finland). Hamp-Lyons (2011) and Ding and Bruce (2017) chart the rise of EAP globally in response to the globalisation of HE, the increase of international students attending university in anglophone countries, and the growth of English Medium Instruction (EMI) education in NNES contexts.

EAP provision may be a pre-entry requirement for students who do not meet the language requirements of the university hoping to join a degree programme (a pre-sessional course). Alternatively, EAP could form part of a Foundation year programme, which covers broad academic subjects to orientate students into university before starting their degree studies proper. EAP is also delivered in-sessionally, either as part of credit-bearing courses or stand-alone provision. EAP grew out of English for Specific Purposes (ESP) and an excellent overview of the history of the field is given in Ding and Bruce (2017).

Various definitions of EAP have been put forward to distinguish EAP from more general types of language teaching. Hyland and Hamp-Lyons (2002) described EAP as “Language research and instruction that focuses on the specific communicative needs and practices of particular groups in academic contexts. It means grounding instruction in an understanding of the cognitive, social and linguistic demands of specific academic disciplines” (p. 2). Hyland (2006) went on to identify EAP as teaching English “with the aim of assisting learners’ study or research in that language” (p. 1), drawing on Flowerdew and Peacock (2001) and Jordan (1997). The focus of EAP on helping students develop study skills is highlighted by Gillett (2017), who describes EAP as “the language and associated practices that people need in order to undertake study or work in English-medium higher education” (para. 1) and Ding and Bruce (2017) who attest “the overall goal of EAP courses is to enable students to develop a complex integration of knowledge and skills in order to communicate and participate effectively in higher education as students and researchers” (p. 65).

While there is broad agreement as to the overall aim of EAP, complexity is added on closer inspection. A distinction is drawn between English for General Academic Purposes (EGAP) and English for Specific Academic Purposes (ESAP) (Jordan, 1997). EGAP describes an approach that aims to equip students with a general awareness of academic convention, generic academic skills, and language that can be applied across fields of study (De Chazal, 2014). This is in line with what Widdowson (1983) termed wide-angle ESP courses, meaning courses that aim to enable students “to cope with undefined eventualities in the future” (p.6). In contrast, ESAP has a tighter, discipline and genre specific focus, closer to Widdowson’s narrow-angle ESP courses. In practice, this may mean an EGAP course could cover general search strategies for finding sources but an ESAP course may focus on using a discipline specific database. Similarly, an EGAP course may look at argumentation generally, whereas an ESAP course may look at how an argument is made in a legal case, following the specific conventions of that genre. As argued by Alexander, Argent and Spencer (2008), EAP is more focused on needs than language level, with a focus on “what the student is trying to do in the future” (p. 18).

As part of preparing students for their future studies and beyond, EAP also has a role in students’ academic socialisation. Duff (2010) describes academic discourse socialisation as the process through which a student becomes a member of a discourse community, transitioning from an outsider or novice to an insider. As Schmitt and Hamp-Lyons (2015) put it, “Many EAP courses take an expanded view of EAP within which instruction aims both to improve the language proficiency of learners and to bridge any gaps between student’s previous educational experiences and the educational culture they wish to join,” (p. 5). The concept of academic socialisation was expanded by Lea and Street (2006) in the academic literacies model. This model encompasses the study skills model (in which good writing is mainly a question of mechanics) and the academic socialisation model (in which students need to learn the norms and tropes of the discourse in their academic community) but goes a step further by bringing in issues like identity and power. Wingate and Tribble (2012) bring together the concept of EAP and that of academic literacies as “complementary components in an inclusive writing pedagogy for students of all backgrounds” (p. 491). It is important to note this move away from

viewing EAP as a deficit model of ‘correcting’ lower language competence level students to a more inclusive model, being of benefit to any students, regardless of their language competence.

This evolving role of EAP is discussed by Ding and Bruce (2017) as they explore its place in HE and the tensions arising from it being seen as either a support service for academia or as an academic discipline in and of itself. In terms of recognising EAP as a discipline they note it has established a growing body of research, encompassing second language acquisition, genre analysis, corpora studies, systemic functional linguistics and critical EAP (Ding & Bruce, 2017), as well as dedicated research journals, such as the *Journal for English for Academic Purposes*. There has also been a growth in EAP organisations, such as BALEAP (formerly the British Association of Lectures of English for Academic Purposes) or the European Association for the Teaching of Academic Writing (EATAW) and specific EAP qualifications, such as modules on MA courses in Teaching English as a Second or Other Language (TESOL) or the BALEAP TEAP (Teaching EAP) accreditation scheme. This scheme incorporates the BALEAP Competency Framework for Teachers of English for Academic Purposes (CFTEAP), which categorises competencies needed by teachers of EAP in the areas of academic practice, EAP students, curriculum development and programme implementation. Bruce and Ding point to this growing professionalism to demonstrate how EAP is maturing as a discipline and should take on greater status in universities (2017).

As EAP itself matures, so the understanding of the role of the EAP practitioner is also refined. An EAP practitioner may play many roles, including that of syllabus designer, materials developer, and collaborator with subject specialists (Ding & Bruce, 2017). The wide-ranging nature of EAP is also apparent through the CFTEAP (Alexander, Bell, et al., 2008), which outlines EAP practitioner competencies, as illustrated in Table 1.3.

Table 1.3 CFTEAP Competencies

Competency	Academic Practice	EAP Students	Curriculum Development	Programme Implementation
Requires knowledge of	• Academic contexts • Disciplinary differences • Academic discourse • Professional learning, development and autonomy	• Student needs • Student critical thinking • Student autonomy	• Syllabus and programme development • Text processing and text production	• Teaching practices • Assessment practices

Of particular relevance to this dissertation is the role of assessment practices. The CFTEAP highlights the need for EAP teachers to understand the relationship between assessment and teaching and learning, which involves choosing appropriate modes of assessment and using the outcomes of assessment to inform future teaching and learning. Assessment literacy is of particular importance when EAP assessment may have a gatekeeping function e.g. barring entry to degree programmes from pre-sessional courses. It is also necessary to note that in EAP assessment is not limited to language competence, but to wider academic skills. This should be no surprise. As Paran put it in his chapter in *Testing the Untestable* regarding language testing more generally, “Within language education, since often more is taught than only language (...), the case can be made that not only language should be tested” (2010, p. 1) meaning it is within EAP’s remit to test transferable skills of value to students in their other studies and beyond.

1.3.1 Background study: Collaboration in EAP

Collaborative assignments feature on EAP courses, which often serve as preparatory courses for students’ future studies, in the form of a pre-sessional or foundational course during the first year at university. As the role of EAP is to support students in their concurrent and future studies, preparing

students for collaborative assignments becomes an authentic aim, given their prevalence across degree programmes. As a precursor to this study, Bolster and Levrai surveyed EAP practitioners across a range of institutions and found “collaborative assignments feature in 83.3% of courses and contribute to final grades in 76% of cases” (Bolster & Levrai, 2022, p. 97).

What must be recognised, however, is that EAP practitioners do not share the same understanding of the distinction between cooperation and collaboration as that found in the literature. The survey by Bolster and Levrai illustrated the gap between researchers and practitioners as it was widely felt by practitioners that *anything* students do together could be classified as collaboration despite Hathorn and Ingram’s (2002) valid point that students simply working together does not mean that they are collaborating. While there does need to be more focus in the practitioner conception of collaboration it is also the case that the researcher conception is too narrow for the realities of the EAP classroom.

One of the issues is that a collaborative assignment in an EAP class will not be a short, single lesson event but more likely a multi-week or full semester task, which requires students to engage in a process writing approach to develop an academically sound piece of text. Writing may take place in or outside the classroom and, by necessity, this changes the conception of collaboration from that based on short, in-class, synchronous activities (Storch & Knoch, 2023). It must be recognised that during a collaborative assignment (such as a group essay or group presentation) there may be times when students work more cooperatively than collaboratively and Bolster and Levrai proposed the following definition for collaboration in EAP (Bolster & Levrai, 2022).

A collaborative assignment is one where learners work together and make equitable contributions to develop an indivisible artefact for which they share responsibility and ownership. During the development of the artefact, learners may work synchronously or asynchronously, face-to-face or online, but there is interdependence between group members, drawing on all their strengths. (p. 101)

1.4 Assessment

The crux of this dissertation is the summative assessment of collaborative assignments in EAP courses. While there are benefits to collaborative learning and group assignments can play an important formative role in students' development, EAP courses exist in the context of HE and summative assessment plays an essential role. Bloxham and Boyd's (observation about the state of assessment in HE is useful for contextualisation, particularly the latter point about how assessment literacy is developed.

Sadly, though, university assessment practice lags well behind its equivalent in the school sector (Murphy 2006), relying largely on a limited range of tried (but not always tested) methods. It is dealt with in an ad hoc way (Swann and Ecclestone 1999a) and the situation is not mitigated by the 'amateur' status of many academics regarding assessment (Ramsden 2003: 177). We learn the craft of assessment informally through being assessed ourselves and through being part of a community of practice, but lack scholarship regarding assessment (Price 2005). (2007, p. 3)

Assessment plays a pivotal role in EAP, with Manning contending that "assessment can be considered as a prevalent practice associated with the teaching and learning of EAP" (Manning, 2013, p. 23). The same year, Ding raised key questions around assessment in EAP (see Table 1.4).

Table 1.4 *Key questions for assessment in EAP (Ding, 2013)*

1. Are we testing achievement or proficiency?
2. Are we testing four skills or integrated academic literacies and discourses?
3. Are we bound by the need to test in the ways most familiar to us?

All of these questions hold relevance in the consideration of collaborative assignments. Are we testing the output of a group's collaboration or how well they collaborated? Are we looking beyond language to collaboration skills? As EAP operates in a HE context, are we limited by the typical norms of (summative) evaluation?

When discussing assessment in HE more widely, Bloxham and Boyd (2007) note that it "shapes the experience of students and influences their behaviour more than the teaching they receive," (Bloxham & Boyd, 2007, p. 3). The importance of assessment to students and how it focuses their attention is well recognised (Chan, 2023; Schellekens et al., 2021). As Gibbs puts it, the assessment system is "the dominant influence on the way students learn: on how much effort they put in and what they allocate this effort too" (1999, p. 42). That which is assessed is that which students pay most attention to and provides "important motivating pressure," (Brown, 2011, p. 46), which again will hold relevance when later discussing whether to assess collaboration itself.

Assessment has traditionally been an individual endeavour and degrees are awarded for individual work and achievement. Boud and Bearman (2022) discuss the tensions around individual assessment, arguing an "overly individualistic notion of learning profoundly frames and disfigures what is assessed and how it is assessed," (p. 460), particularly given learning is a social, collaborative activity. Traditional assessment also pushes back against collaboration on two fronts; first in terms of fears of collusion (Boud & Bearman, 2022) and second in terms of the competitive element to individual assessment, particularly in contexts where grades are normed on a curve and therefore an individual's grade is affected by the performance of other students. As degrees are still awarded individually, the individual needs to be assessed within the collaborative assignment. Although Meijer et al. (2020) consider the affordance of collaborative assignments being clearly indicated on student transcripts and Boud and Bearman (2022) call for full reform of HE education to reflect the collaborative nature of learning, it seems these developments are a considerable way from becoming common practice in HE institutions.

1.4.1 Bridging assessment and learning outcomes

Assessment does not exist in a vacuum. We assess for a purpose and that purpose is typically to gather evidence to demonstrate students' level of knowledge, if learning has taken place, and/or if learning outcomes have been met, with assessment subordinate to those learning outcomes (Boud & Bearman, 2022). The assessment of learning outcomes is important to a range of stakeholders, be that policy makers, potential students or future employers (Nusche, 2008) and as Queen Mary Academy note (n.d.) "most Higher Education institutions across the world use an 'outcomes-based' approach to learning" (para. 1). The link between assessment and learning outcomes is clear if we consider different approaches to course design.

Richards (2013) reviewed three approaches to course design; forward, central and backward. Forward design is content driven. The content of a course is determined first, with the processes of the course second and the outcomes the last thing to be determined. Central design starts with the process itself, with task-based learning being a good illustration of this. From the outset, the methodology is determined, with content and outcomes considered later. In Backward design, the outcomes are considered first, then the means by which they can be measured (the assessments), and content is the final element put into place. As Richards (2013) notes, the Common European Framework of Reference for Languages (CEFR) is a clear example of Backward design with the outcomes paramount. In all approaches, learning outcomes have a pivotal role and as Wiggins and McTighe (2005) emphasise, the curriculum serves as "a map of how to achieve the "outputs" of desired student performance, in which appropriate learning activities and assessments are suggested to make it more likely that students achieve the desired results" (p. 6). This is crucial as it locates assessment as an active agent in student learning in helping them achieve the course outcomes and desired learning. As Norton explained, "one of the pedagogical benefits of assessment is that it can be used to act as a lever to make students actively engage with a given task," (Norton, 2004, p.688). Indeed, assessment criteria can "drive the student learning," (Norton, 2004, p.691-692)

1.4.2 Assessment of/for/as Learning

There is a wealth of assessment literature and a valuable scoping review by Schellekens et al. provides a good overview of prevalent taxonomy. They chart the current conception of assessment to the Assessment Reform Group (ARG) who, in 1999, reimagined the former functions of assessment of ‘formative’ and ‘summative’ to ‘Assessment for Learning’ (AfL) and ‘Assessment of Learning’ (AoL) (Schellekens et al., 2021). Where Scriven (1967) focused on the temporal difference between formative (during course) and summative (end of course) assessment, for the ARG it is more the intent of assessment which is of concern.

The focus of AoL is the judgement of performance and outcomes, while AfL is assessment intended to improve the learning and teaching processes (Schellekens et al., 2021). In 2013, Earl added the conception of ‘Assessment as Learning’ (AaL), which emphasised the role of the student in assessment practices (e.g. utilising self-assessment and self-directed learning), locating students as “the critical connector between assessment and their own learning” (Earl, 2013, p. 3). While this is the more commonly understood conception of AaL, Torrance (2007) utilised the same term, suggesting assessment could come to dominate learning and could, indeed, be deformative rather than formative (Torrance, 2012). As Paran notes, prevalent models of language competence e.g. the Common European Framework of Reference (CEFR) originate from “the need to specify language competences for testing” (Paran, 2010, p. 2) and Torrance’s argument resonates in the HE context where student competence has to be evidenced, promoting an AoL assessment culture. However, while Torrance’s conception of AaL and the potential negative impact of assessment on learning will be returned to later in this dissertation, it is the Earl conception of student engagement with assessment which will be used when referring to ‘Assessment as Learning’ (AaL), positioning assessment as an arena where learning can take place.

1.4.3 Evaluative judgement and the role of self- and peer- assessment

Growing from AaL and the inclusion of students in their own assessment, is the concept of evaluative judgement. Tai et al. (2018) argue for the role of evaluative judgement in HE since the capacity to judge one's own work, and the work of others, contributes to lifelong learning. They locate evaluative judgement in sustainable assessment, meaning assessment that has a preparatory function in readying students for their futures (Boud, 2000), which chimes with the purpose of EAP. A learner who can identify whether their work meets the expected standards is independent and no longer reliant on a teacher to make that decision for them. A crucial element of evaluative judgement is the role of self- and peer-assessment, which helps students understand quality work and contributes to learning (Tai et al., 2018). Lee (2007) locates self-assessment in AfL and discusses how it can help students become "independent and reflective in learning" (p. 204), which can develop their writing practices. They go on to embrace AaL and advocate for the primacy of student assessment, within a framework of teacher support (Lee, 2016).

Although championed, there have been concerns about the role of self- and peer- assessment in summative assignments, and Leach offers a good summary of the tensions around self-assessment, with students potentially not seeing it as their role, or having too much other work to do, or expecting the evaluation of an expert, or potentially experiencing cultural barriers to accurate self-assessment (Leach, 2012). When it comes to peer-assessment, Rasooli et al. (2024) discuss the how friendship can skew peer-grading, an issue also identified by Meijer et al. (2020), who additionally note peer-pressure, self-interested marking and negative personal relationships as further concerns.

These tensions are echoed by Adachi et al. (2018), when examining academics' perceptions of self- and peer-assessment in HE, although they also value the evaluative judgement benefits of employing such assessment methods. The mixed reception of peer-assessment by students is discussed in Ashenafi (2017), who urges the development of design quality and measurement standards so peer-assessment can be more fruitfully adopted in HE. Storch and Knoch when exploring collaboration in L2 writing (2023) voice the concern about distortion within peer-assessment,

influenced by “individual biases and misconceptions,” (p. 237), which aligns with Meijer et al. (2020), who recognise the potential for the validity of peer-assessment to be compromised for similar reasons. However, it has also been argued peer-assessment can increase students’ perceptions of fairness, engage them more thoroughly in the collaboration process and develop their evaluative judgement (Ion et al., 2023). What is clear, is that self- and peer- assessment have to be supported and considerable learner training is involved if they are going to be implemented (Rasooli et al., 2024; Topping, 2003) with some, such as the King’s Medway, Oxfordshire, Formative Assessment project, seeing the ability to peer-assess an essential foundation for self-assessment (James & Pedder, 2006, p. 116).

1.4.4 Assessment and Sociocultural Theory

SCT not only provides a framework for learning, but also for assessment. AoL is reflective of more traditional summative assessments, where assessment serves as a test of what students know. In contrast, AfL emphasises the formative nature of assessment, following sociocultural theory in arguing that it is through feedback that learning takes place (Brown, 2005), thus making assessment part of the learning process itself. Sadler could be speaking of the ZPD when discussing the role of formative assessment as identifying what a student knows and what they need to know (Sadler, 1989). Lund (2008) also makes a sociocultural argument for assessment, locating assessment as an entity consisting of teachers, students and cultural tools (assessments, assessment guidelines and criteria) engaged in a social endeavour, bringing assessment, learning and instruction together. Davidson (2004), discusses this in some depth when discussing how environment shapes teachers’ beliefs about assessment and advocating that teachers “need time and space to develop a sense of ownership and common understanding of the assessment process and to articulate and critique their often-implicit constructs and interpretations” (p. 328).

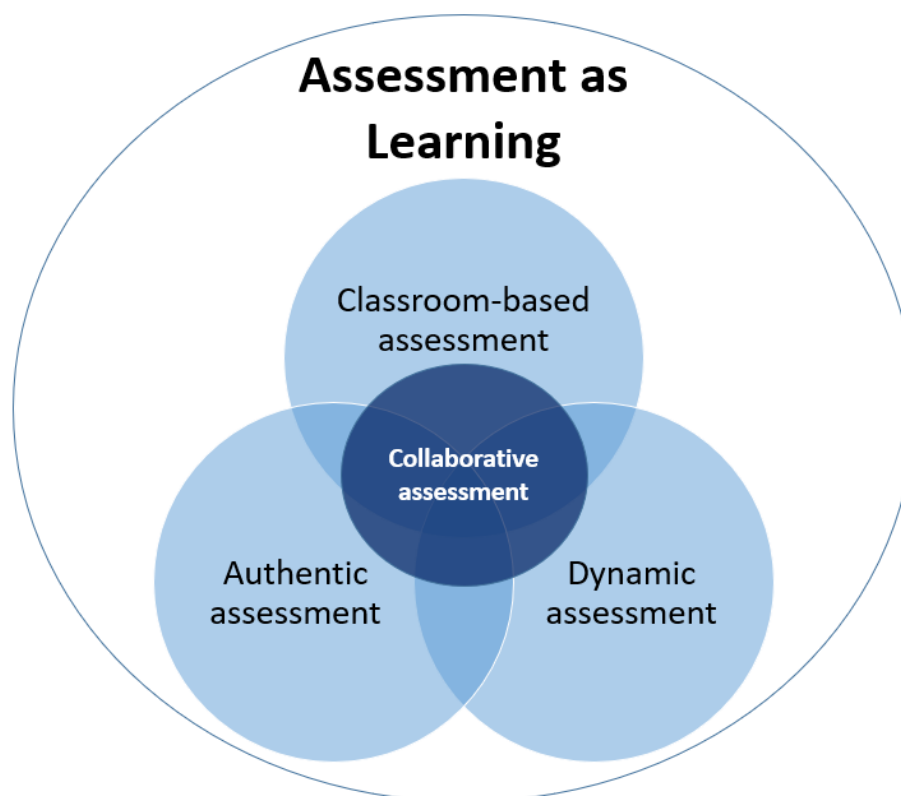
The sociocultural rationale for assessment also informs dynamic assessment (DA) (Anton, 2012), which is the integration of assessment and instruction for the purpose of development. Within DA the

key concern is what students may be able to achieve in the future with the use of scaffolding and feedback to help them get there, with DA, “trying to discern and nurture emergent abilities” (Anton, 2012, p. 107). The aim is to “unify assessment and instruction into a single activity, the goal of which is learner development” (Poehner & Lantolf, 2005, p. 254). The DA approach of the teacher intervening to try and improve student performance matches the process writing approach common in EAP, where the teacher provides feedback during the drafting process to improve the quality of the final essay. DA also resonates with EAP as DA “aims at cognitive development and transfer of skills to future tasks” (Anton, 2012, p. 110), which is the very basis of EAP (see 1.3).

1.4.5 Locating collaborative assessment

Building off the concept of DA, we can locate collaborative assessment within AaL, sitting at the nexus of DA, classroom-based assessments (CBA), and authentic assessments (AA) (see **Figure 1-3**).

Figure 1-3 *Locating collaborative assessment within AaL*



Collaborative assessments can be considered CBA since they are generally teacher-led class assignments. Note that 'class' is not classroom bound as a collaborative assignment may be completed through a blend of in-class and out-of-class activities, potentially mediated through collaborative virtual spaces. CBA is a relatively under-researched area in the assessment literature (Turner, 2012) and, as she notes, we need to be aware of generic labels for assessments. While assessments may have the same method label (in the case of collaborative assignments a 'group essay' or 'group presentation' for example) it is important to recognise that there could be considerable variation in the assessment tasks involved from context to context. When we talk about collaborative assessment we are really discussing a wide range of assessment tools and methods unified by the requirement that students work together and have shared ownership of the final product.

Collaborative assessment can also be considered a type of AA, which falls within the AaL sphere, with assessment "being integrated throughout the learning process" (Searle & Poth, 2021, p. 424). Villarroel et al. provides a neat definition of AA: "Authentic assessment aims to replicate the tasks and performance standards typically found in the world of work, and has been found to have a positive impact on student learning, autonomy, motivation, self-regulation and metacognition; abilities highly related to employability" (2017, p. 840), with those objectives chiming with the aims of EAP. A benefit of AA over more traditional tests of knowledge are that "authentic tasks not only serve as assessments but also as vehicles for such learning" (Mueller, 2005, p. 3). Schmitt and Hamp-Lyons (2015) recognise the appropriacy of AA tasks in EAP, but also how authenticity may have to be compromised in light of other concerns in EAP testing e.g. satisfying the needs of exam boards.

Key principles of an assessment are typically "reliability, validity, practicality and (sometimes) consequences)," Hamp-Lyons (2009, p. 256). However, Turner (2012), concurs with Hamp-Lyons in suggesting concepts like 'validity' and 'reliability' need to be re-examined in the AfL and CBA context, which is qualitatively different to the traditional large-scale summative testing situation from which the understanding of the concepts originates. Davidson also recognises that validity may need to be reconceptualised, so rather than validity being demonstrated by agreement between assessors and

criteria, “it is only through the teacher interpretation and negotiation of judgments that judgments can be made valid and reliable,” (2004, p. 309). This changes reliability and validity from fixed to negotiated and constructed. If we locate collaborative assessment within DA, there are clear challenges to the traditional concept of reliability. The aim of DA is to modify learner behaviour during the assessment to achieve a better performance. However, change in performance undermines the reliability of the assessment (Lantolf & Thorne, 2006) as a traditionally reliable assessment should give the same results every time. This is something dynamic assessment very much tries *not* to do. As Poehner and Lantolf put it: “any assessment that fails to determine the extent to which the person’s performance is modifiable is incomplete” (2005, p. 235).

Validity in testing is a key concept that has undergone scrutiny, moving beyond the general conception of it meaning an assessment measures what it claims to measure. Messick (1989) viewed validity not so much as a quality of the assessment itself, so much as the interpretation of the results of the test and the actions resulting from those implications. For instance, the validity of an IELTS result would not be whether a score of 5.5 fully captures a student’s language ability but whether a score of 5.5 means the student is ready to enter an EMI institution. In an EAP pre-session, the question of validity speaks to using the assessment results to allow a student to enter their degree programme studies. In relation to CBA, validity would therefore refer to how the results of the assessment are used. Messick’s conception of validity is the pre-eminent in the area of language testing (Fulcher & Davidson, 2007).

However, an alternative means of measuring an assessment is “test usefulness”, proposed by Bachman and Palmer in 1996. Although a collaborative assignment is not a test in the traditional sense, the test usefulness framework does seem applicable to collaborative assessment. Test usefulness was discussed by van Daalen (1999) in relation to alternative assessment, with the measure of an assessment being “How useful is the information provided by the test for its intended purpose?” (p. 3). Test usefulness considers six qualities (see Figure 1-4).

Figure 1-4 The six qualities of test usefulness (van Daalen, 1999, p. 3)

- **Reliability** (accuracy of a test in measuring consistently what it is supposed to measure).
- **Validity** (degree to which a test measures what it is designed to measure).
- **Authenticity** (characteristics of a test task).
- **Interactiveness** (topical knowledge, affective schemata and metacognitive strategies, the relationship between input and expected responses).
- **Impact** (how the test affects processes of learning and instruction (e.g., washback)).
- **Practicality** (feasibility or usability of a test in the situation for which it is intended).

For a collaborative assessment, interactiveness seems a necessary measure, as it considers “the degree to which the individual test taker’s characteristics (language ability, background knowledge and motivations) are engaged when taking the test” (Fulcher & Davidson, 2007, p. 15). So too does impact, given EAP’s aim to prepare students for their other studies and consideration of impact necessitates considering washback and trying to design towards beneficial washback. Washback “is rooted in the notion that tests or examinations can and should drive teaching, and hence learning” (Cheng & Curtis, 2004, p. 4). What is assessed focuses the attention of teachers and students and, as Messick held, washback describes the “extent to which the introduction and use of a test influences language teachers and learners to do things they would not necessarily otherwise do that promote or inhibit language learning” (1996, p. 241). Assessment design impacts what happens in the teaching and in the learning, although the relationship between the three is complicated (Cheng & Curtis, 2004).

A key principle in assessment not covered in the usefulness test is fairness, which is particularly pertinent to collaborative assessment, where an individual’s result can be dependent on the performance of others. In assessment circles, there has been discussion on the nature of fairness and

whether, for example, fairness can be seen as a component of validity or if validity is folded inside the concept of fairness (Walters, 2012). Regarding the assessment of a collaborative assignment, Kane's (2010) conception of fairness illustrates the challenges, envisaging fairness as a combination of procedural fairness (meaning everyone is treated the same way and works under the same conditions) and substantive fairness (meaning the evaluation of the assessment, and decisions based on that, are appropriate and reasonable for all). It is questionable if the measure of procedural fairness can be met in a collaborative assignment as people will be working in different groups, with different personalities and while some groups may operate in harmony, others may encounter dysfunction. So too could a collaborative assignment fall on the issue of substantive fairness since an individual's performance on the assignment in question may not mean their future collaborations will go the same way. As Boud and Bearman put it, "we contend that unfairness occurs when a particular collaborative experience is seen as being representative of a student's learning" (2024, p. 464), since not all collaborations will be equal and learning is represented over many collaborative experiences rather than single events. This recalls Vygotsky's critique of IQ tests for being isolated snapshots rather than a complete picture.

That is not to say the situation is hopeless and fairness in collaborative assessment is unobtainable. As Gipps and Stobart (2009) recognise when discussing fairness through a sociocultural lens, "We will never achieve fair assessment, but we can make it fairer" (p.116). This could be achieved by following Tierney's (2014) five needs of fairness: opportunities for students, transparency in assessment procedures, critical reflection about assessment practices, constructive classroom environment, and a balance between equality and equity (p. 62). This may not lead to an objectively fair means of assessment, but in their exploration of fairness in the field of medical assessments, Valentine et al. (2022) make the case that, "Assessment is always an evaluative process and therefore subjective," (p. 355). Absolute fairness may be out of reach but as fair as can be should be a reasonable aim.

1.4.6 Assessing collaboration

One of the tensions around collaborative assignments is the question of assessment, which is recognised as a complex issue (Berry, 2007; Meijer et al., 2020; Strauss & U, 2007; Strijbos, 2016) to which we are still lacking solutions and which remains under-researched (Meijer et al., 2020; Storch & Knoch, 2023). Indeed, assessment is one of the key research agenda areas Li & Zhang (2023) identified in their exploration of current and future L2 collaborative writing research. As such this section draws on relevant literature from both L2 learning and HE as EAP operates at the nexus of both.

Assessment is also one of the components in the Group Learning Activities Instructional Design (GLAID) framework, proposed by de Hei, Strijbos et al. (2016) after conducting a thorough thematic review of group learning activity design in HE, which proposes a framework for developing group learning activities in HE through consideration of the following eight key components: (1) interaction, (2) learning objectives and outcomes, (3) assessment, (4) task characteristics, (5) structuring, (6) guidance, (7) group constellation, and (8) facilities. They held that this framework was meant to provide a “comprehensive design approach, which is grounded in the literature” (de Hei, Sjoer et al., 2016, p. 395). When it comes to assessment, the framework maintains decisions need to be made about whether to employ individual or group assessment, who should undertake assessment (e.g. incorporation of self- and peer-assessment), whether the assessment should be formative and summative, and the “number, timing and integration of assessment measures,” (de Hei, Strijbos et al., 2016, p. 38).

However, it has even been questioned whether collaboration should be assessed, with Nouri et al. (2017) considering collaboration a way of learning rather than a skill to be directly assessed itself within the Programme for International Student Assessment (PISA) assessment scheme. Nonetheless, as Swan, Shen and Hiltz (2006) amongst others argue when discussing online courses, collaboration must be assessed in a collaborative assignment if it is an intended learning outcome, a point echoed by Williams (2017) when discussing collaborative assessment in HE, who stresses the role of assessing the *process* of collaboration alongside the product, even while acknowledging the assessment of

collaboration is messy. Where assessing collaboration is discussed in the literature, it is commonly located as formative rather than summative assessment (Meijer et al., 2020; Storch & Knoch, 2023), not least because all the challenges of assessing collaboration are compounded in the case of summative assessment. Nonetheless, the assessment of collaborative assignments matter, because as Beard et al. (1989) asserted when discussing assessing university student collaborative assignments, “the assessment system, although one of many factors, may encourage members’ full participation and their practice of effective group behaviors. It may help foster members’ full engagement in the collaborative-writing process” (p. 45). Essentially, the way a collaborative assignment is assessed impacts how well students engage with it, which in turn influences what they get from it.

1.4.6.1 Summative grading of collaborative assignments

In terms of grades in HE, collaborative assessments have been found to disadvantage stronger students and advantage weaker students (Almond, 2009) and marks tend to cluster closer together than in individual assessments (Berry, 2007, Meijer et al., 2020), which has serious implications if collaborative assignments are credit bearing as students may pass who would not otherwise do so and those who might reach distinction could fall short (Plastow et al., 2010). The grading process of a collaborative assignment in order to obtain those grades is another area of concern. There are multiple approaches to grading collaborative assignments, as outlined in Race (2001), including the use of self- and peer-assessment, the application of an individual weighting factor based on individual contribution (see Nepal, 2012) and the use of additional individual tasks or assessments. As yet there is no agreed principle in the literature as to who should be grading collaboration (e.g. teacher led-, self- or peer-assessment), how it should be done, and little discussion of how collaborative assignments are assessed in EAP specifically.

What those grades mean also needs to be considered. Strijbos (2016) recognises that, during a collaborative assignment, individuals may not converge to have the same learning gains and, indeed, it is likely that each individual may process the experience differently, leading to different learning

gains, which may take place at different rates and over different timeframes. This is echoed by Boud and Bearman's discussion of collaborative assessment in HE (2022) when they contend "students can learn things that cannot be known in advance through working with others" (p. 464), highlighting unexpected learning outcomes that can occur during collaborative assignments. Storch and Knoch when discussing L2 writing (2023) also caution that it may not be possible to assess individual abilities through a piece of collaborative writing, echoing a point also made by Van Aalst (2013) when reviewing research on assessing collaboration from 1994 to 2009 at both K12 and HE level. This problematizes employing a collaborative assignment as a single summative assessment. Pitt when discussing collaboration in HE (2000) laments how dependent an individual's performance in a collaboration is on the people they are collaborating with and how membership of a different group may lead to a different level of performance. This raises the question of if we can draw any generalizable conclusions from a collaborative assessment and whether any 'score' in collaboration could stand the test of repeatability. A strong performance in one collaborative assignment may not lead to a strong performance in another assignment, on a different topic, with a different group of people.

1.4.6.2 How collaborative assessment can impact student behaviour

The challenges are manifold in assessing collaboration, but they need to be addressed as assessment design has the potential for washback, especially in the sociocultural construct where assessment and instruction are intertwined. As discussed in Wall (2012), assessment washback describes a change in behaviour in reaction to an assessment and so the way collaboration is assessed may impact how both the teacher and students engage with the assignment. Pitt (2000) applied game theory to collaborative assessments and concluded collaborative assessments are essentially unfair as the best course of action by a group is for the students to claim they collaborated when in reality the most proficient students do all the work to ensure a high-quality product. Norton et al. (2001) explore psychology undergraduates use of tactics to try and increase their scores, employing what the paper terms the 'Rules of the Game' (ROG) to describe their strategic approach to assignments.

As Strijbos (2016) highlighted when discussing grading criteria for collaborative assessments, what is valued in grading criteria may undermine the collaborative process itself e.g. criteria which reward participation may encourage greater student participation even if this has a detrimental effect on the assignment in question. In other words, students may try to ensure they are seen to participate, regardless of how useful that participation actually is to assignment completion. Davidson (2004) also problematised criterion—referenced assessment, contending “the nature of achievement described is mediated, even constituted, by the instruments used to describe it,” recalling Lantolf and Frawley’s contention that assessment criteria can be reductive (1985, p. 340). Criteria draw importance to themselves meaning grading criteria may push students one way or another along the collaboration/cooperation continuum or encourage certain collaborative behaviours at the cost of others, thereby “over-emphasizing those aspects of the curriculum, or fragmenting and compartmentalizing learning and teaching” (Davidson, 2004, p. 309). What things are graded, and how they are graded, can influence student behaviour. As Norton (2004) notes, “it is not surprising that students tend to use assessment criteria in a strategic, marks-orientated approach,” (p. 689). Not only this, grading criteria also influence what the teacher focuses on in teaching to match what is going to be assessed (Norton, 2004, p. 689).

In both L2 and HE research, solely grading the product in a collaborative assignment may encourage positive interdependence between group members, but could be unfair in light of freeloaders (Meijer et al., 2020; Storch & Knoch, 2023) and could lead to a performance-orientated approach where students divide the work according to their strengths (Meijers et al., 2020), ultimately undermining the original collaborative intent of the assignment. Nor does the product itself “represent anything about the achievement of outcomes associated with collaboration,” (Boud & Bearman, 2022, p. 462), meaning a lot of potential learning may not be visible through the product alone and the product is not an effective measure of the collaboration. These critiques of solely assessing the product point to the importance of being able to assess an individual within a team effort and the need to assess the process as well as the product. However, individual assessment within a

collaborative task could lead students to focus on their own performance, rather than that of the group as a whole (Meijers et al., 2020). Similarly, adding additional individual tasks a collaborative assignment could dilute the collaborative task and send mixed messages to students about the need to work together (Meijers, et al., 2020). Careful consideration of collaborative assignments is required to have a positive impact on instruction and learning and, as recommended by McDonald when discussing assessing collaboration on online HE courses (2003), the design of the assessment tool should itself facilitate collaboration. If an assignment can be completed individually, there is no authentic reason for students to collaborate.

However, collaboration as a CBA provides flexibility in ‘assessment windows’ (Palmour, 2023), which are opportunities to assess students through, “any actions, interactions or artifacts (planned or unplanned, deliberate or unconscious, explicit or embedded) which have the potential to provide information on the qualities of a learner’s (or group of learners’) performance” (Hill & McNamara, 2011, p. 397). In the context of a written collaborative assignment, these assessment windows could include “peer and individual reflective diaries and teacher observations” (Storch, 2017, p.139). As such, there are multiple pathways we could take to try and determine individual assessment within a collaborative task, although as yet there is no clear recommendation for how to do this. Within this dissertation the term ‘assessment lenses’ is preferred to assessment windows. Rather than independent assessment windows, the assessment opportunities are envisaged as lenses that can be used in conjunction to bring a collaboration into greater focus.

1.4.6.3 Freeloaders, social loafing and suckers

A common complaint with regards collaboration is the occurrence of freeloaders in a team assignment. Strijbos (2016) draws a distinction between the phenomena of ‘social loafing’ (which describes when an individual does not work as hard in a team as they would in an individual assignment) and ‘freeloading’, which describes the situation where a member of the group does not contribute to the work but gains the benefit of the work of others (also referred to as free-riding (Gaur

& Gupta, 2013)). A closely related phenomena is the sucker effect, where an individual lowers their productivity to avoid being seen as, or taken as, a sucker, doing all the work while others freeloader (Davies, 2009).

The concept of effort in a collaborative assignment is an interesting one, especially, as Johnston and Miles (2010) point out, in an individual assignment there is typically no weighting given to effort. Nonetheless, as discussed by Wise and Smith (2016), a lack of effort by a test-taker challenges the validity of the test and the conclusions we can draw from the results. As such, freeloaders could undermine the validity of collaborative assignments. Maiden and Perry (2011) investigate the use of different strategies to deal with freeloaders and found that any explicit attempt to address the issue was viewed favourably by students.

1.4.6.4 Looking inside a collaboration

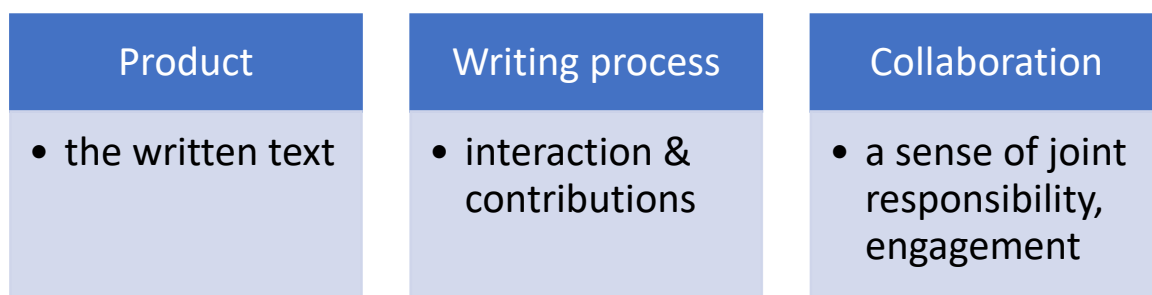
A challenge for the teacher when assessing collaboration is that collaboration is an opaque process, as recognised by Gammie and Matson (2007) and Johnston and Miles (2010), who suggest the use of peer grading for the process of collaboration as the students are the ones inside the process. Kearney et al. (2016) also strongly advocate for the use of peer-assessment in summative collaborative assignments and how peer-assessment in itself is a learning opportunity. However, computer supported collaborative communication (CSCC) allows for greater insight into the collaboration process (Alghasab & Handley, 2017; Elola & Oskoz, 2017; Williams, 2017) making previously invisible interactions visible and thus providing the opportunity for teacher assessment of collaborative practices. An important addendum is that teachers also need to help students develop the digital literacies they will need throughout CSCC (Elola & Oskoz, 2017) and this would be particularly true of EAP programmes, given their role in preparing students for future study. Nevertheless, it is questionable if the strategies used by researchers can scale to classroom environments. For example, Alghasab and Handley (2017), when researching online collaboration by high school students, interlaced changes in a text and chat forum discussion posts together to assess the impact of

comments on edits as a means of measuring collaboration. Such an approach could become overwhelming for an EAP classroom teacher with multiple groups and multiple classes.

1.4.6.5 *What to assess when assessing a collaborative assignment*

Another serious concern is what to assess when assessing collaboration. Storch and Knoch (2023) identify three elements that could be assessed in an L2 collaborative writing activity (see Figure 1-5).

Figure 1-5 *The three elements of collaborative writing (based on Storch & Knoch, 2023, p.234)*



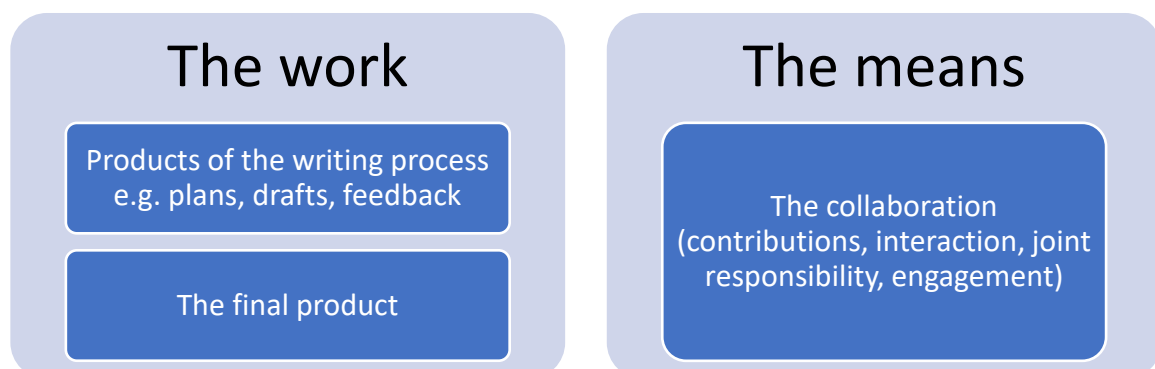
Assessing the product is potentially the least problematic of these as it is a single piece of written work and there are similarities to assessing a piece of writing by an individual, given the nature of collaborative writing as indivisible. However, they caution against using assessment criteria written for assessing individual work as they may not address issues that could arise in a piece of collaboration e.g. problems with coherence and cohesion that could arise if the writers wrote cooperatively and amalgamated the separate pieces at the end (Storch & Knoch, 2023). This concern may be mitigated within EAP, however, as it is likely that coherence and cohesion are clearly addressed in the marking criteria for individual writing, given their importance to argumentation and effective academic communication.

Assessing the process of writing poses more challenges, not least in terms of the visibility of student contribution to the teacher, particularly in an assignment which may be written over an

extended period of time, both synchronously and asynchronously. While technology gives us some insight into the process of online writing, attempting to review version history of online documents to see who did what would be time-consuming (Zhang & Chen, 2022). The information provided would also be questionable, as knowing who typed what does not show the research, interactions and discussions that may have led to that typing. While a computer-based assessment tool could potentially process a text to help determine who did what, it would not be able to capture the quality of each students' contributions, particularly those not captured in keystroke (Van Aalst, 2013).

Within EAP, the nomenclature of 'writing process' could be confusing, with the prevalence of the concept of 'process writing' in EAP (Eckstein, et al., 2011; Khan, 2016; Salter-Dvorak, 2016), which considers writing an iterative process, involving drafting, feedback and editing. The problem is there are concurrent processes in play, those of the writing process itself, and that of the collaboration. This potentially confusing use of 'process' to discuss both the academic writing process and the process of collaboration was something that was also apparent in the data collected for this dissertation, with participants blurring the meaning between the two. For the EAP context it may be preferable to envisage the writing process and the final product as 'the work', and the collaboration as 'the means', as illustrated in Figure 1-6.

Figure 1-6 *Suggested elements of collaborative writing in EAP*



Turning to the third element Stoch and Knoch (2023) identify as assessable, the collaboration itself, or the means in the above illustration, begs the question, “What does ‘good’ collaboration look like?”, which is returned to later in this dissertation. Storch and Knoch (2023) and Van Aalst (2013) would look to mutuality and interaction, how fully students are engaging with each other’s ideas and feedback. Williams (2017) outlines roles students may play in collaboration and it raises questions regarding whether characteristics such as empathy, integrity and flexibility (key elements of a successful collaboration) can, or should, be measured and assessed. As one of the arguments in favour of collaboration at university is preparation for future work it is also pertinent to look at how collaboration is discussed in the world of business.

Belbin (1981) proposed nine team roles, divided into three broad categories; action-orientated roles, person-oriented roles and cerebral roles. While action-oriented roles are highly visible (and therefore measurable) person-oriented roles serve more of a social, facilitation function, developing and maintaining relationships in the team, which is a harder aspect of collaboration to measure. Similarly, cerebral roles may not make many visible contributions but what contributions they do make may have great value. Other attempts to classify team roles result in a similar mix of roles, which contribute to the team effort in different ways (Mathieu et al., 2015). The question in a collaborative assessment is how to weigh the various roles someone may play throughout a process, given that roles may not be static and that some of those roles, although essential, do not lend themselves to measurement, a point made by Swan, Shen and Hiltz (2006), Williams, (2017) and Frykedal and Chiriac (2011).

1.4.7 Assessment literacy

This leads into the question of teacher assessment literacy when it comes to assessing collaborative assignments. Stiggins contributed greatly to our understanding of assessment literacy and gave the following definition.

Assessment-literate educators (...) come to any assessment knowing what they are assessing, why they are doing so, how best to assess the achievement of interest, how to generate sound samples of performance, what can go wrong, and how to prevent those problems before they occur. (1995, p. 240)

It covers the full range of assessment, from conception through to delivery. The assessment literacy and assessment competence of language teachers is a concern (Meijer et al., 2020; Plake, 1993; Şişman & Büyükkarci, 2020) and needs to be improved as teachers must be assessment competent if the assessments are going to be fair to the students (DeLuca, 2012). As assessment agents, practitioners,

find themselves at the confluence of different assessment cultures and faced with significant dilemmas in their assessment practices: sometimes torn between their role as facilitator and monitor of language development and that of assessor and judge of language performance as achievement. (Rea-Dickins, 2004, p. 253)

This is particularly pertinent in the EAP sphere, where practitioners are engaged in high stakes assessments, which could affect their students' academic careers. Assessment is a common area of professional tension and a source of apprehension for teachers (DeLuca, 2012) as it is an area where training is sometimes lacking. Villarroel et al. (2024) discuss the need for improved assessment practices in HE but also note the difficulty in developing assessment practices due to the time and resource limitations. Despite the importance of the assessment EAP practitioners engage with, the assessment literacy of EAP practitioners has been questioned and found wanting (Manning, 2013; Schmitt & Hamp-Lyons, 2015). Schmitt and Hamp-Lyons trace EAP teachers' relative lack of assessment literacy to it being under-taught in training courses, with most assessment literacy developed "on the job" (2015, p. 6), an issue also discovered by Manning (2013).

Meijer et al. (2020) propose the need to consider “collaborative learning assessment literacy” (p. 1224) as a specific type of assessment literacy, with its own demands. This is not least because assessment of a collaboration would need to incorporate literacy in all the assessment lenses used. For example, if reflection was utilised as an assessment lens as part of a collaborative assignment, Chan (2023), who advocates strongly for reflection, would question its role as an assessment lens as “unless teachers and students are reflection and assessment literate, they might find it difficult to fully unite the two together” (p. 183-184). A teacher could potentially need literacy in self-assessment, peer-assessment and reflection, all as part of their collaborative assessment literacy, and would need to be able to develop those same literacies in students. This is especially important as teacher related factors can impact assessment washback (Wall, 2012), both in terms of the focus of their teaching and the washback of the assessment on students. Indeed, when faced with a novel type of assessment, which could be the case with collaborative assignments, Burrows (2012) identifies teachers as ‘adopters’ (those who embrace the new system), ‘adapters’ (those who take what they want from the new system) and ‘resisters’ (those who do not change their practices in the face of new requirements). Practitioner beliefs and assessment literacies are crucial to any assignment as practitioner beliefs impact their practices (Brown, 2011). This will be seen in the Findings and Discussion of this dissertation.

As Turner (2012) notes in her review of the literature, teachers have an appetite for improving their assessment literacy but there is a question of how to best achieve it, a finding endorsed by Meijer et al. (2020). When discussing CBA Turner also highlights the importance of ‘teacher thinking’ research to see inside the process by which teachers make assessment judgements. Yin (2010) explored teacher thinking in the examination of EAP practitioners and determined they employed both strategic and interactive cognitions when undertaking CBA. The strategic cognition encompasses practitioner beliefs about language learning, the learning context and assessment. Interactive cognitions meanwhile were informed by assessment principles and knowledge and expectations of students. What this demonstrates is that practitioners bring a lot into their assessment approach and different

practitioners in differing cognitions could navigate particular assignments in quite different ways (see also Davidson, 2004). While there has been some use of ‘teacher thinking’ research in relation to collaborative assessments (although not in the field of EAP), teachers have difficulty expressing what they assess when assessing collaboration and are unsure of their assessment literacy in relation to assessing collaboration (Frykedal & Chiriac, 2011; Strijbos, 2016).

The development of a teacher’s assessment literacy seems more informed by personal experience than through formal training in assessment practices and teachers can hold contradictory conceptions of assessment simultaneously (Looney et al., 2018). This contradiction or gap between practice and beliefs was also discussed by James and Pedder (2006) when they delved into the issue of assessment, learning practices and values. They discussed how school teachers’ practices and beliefs interinfluence each other and how there can commonly be a gap between teachers’ values and their practices.

Looney et al. (2018) go on to explore the impact of emotion in assessment, drawing on Stiggins’ (1995) work examining practitioner fear of assessment as an element in the ongoing development and evolution of teacher assessment identity, which encompasses “beliefs, feelings, knowledge and skills” (Looney et al., 2018, p. 455). There is a rich field of literature devoted to identity, with Looney et al. drawing on Beijaard et al. (2004), Day et al. (2006) and Mockler (2011), amongst others, to form understanding of identity as a socio-cultural construct, where teacher assessment identity is developed and evolves through context and experience, responding to events and circumstances (2018, p. 446). Within teacher assessment identity they discuss the contradictory feelings teachers may hold in relation to assessment since

when teachers assess more is in play than simply knowledge and skills. They may have knowledge of what is deemed effective practice, but not be confident in their enactment of such practice. They may have knowledge, and have confidence, but not believe that assessment processes are effective. Most importantly, based on their prior experiences and their context, they may consider that some assessment processes should not be a part of their role as teachers and in interactions with students. Teachers can, quite literally, have mixed feelings about assessment. (Looney et al., 2018, p. 455)

As will be seen, this is very much the case when it comes to collaborative assessment and the findings of this dissertation will discuss teachers' collaborative assessment identity, encapsulating their key principles and guiding orientations.

In sum, assessment of collaboration is a complex area and one of particular concern in the field of EAP given its role in academic discourse socialisation and potential gate-keeping role in universities. There needs to be better understanding of the approach(es) EAP tutors make when assessing collaborative assignments, how the grading of those assignments is conducted, and how that process is perceived by both teachers and students to help us move towards developing a principled approach to collaborative assignment assessment. Meijer et al. (2020) called for research into "teachers' actual rationale for and practices of the assessment of collaborative learning in higher education" (p. 1235), which is an area this dissertation hopes to address.

1.5 Grounded Theory

Grounded Theory (GT) is a qualitative research method which originated from Glaser's and Strauss' research into dying patients and articulated in *The Discovery of Grounded Theory: Strategies for Qualitative Research*, originally published in 1967. As noted by Holton (2007), Grounded Theory does not sit neatly in a particular research paradigm but "transcends the specific boundaries of established paradigms to accommodate any type of data sourced and expressed through any

epistemological lens” (Holton, 2007, p. 267). Rather than entering the field armed with research questions and hypotheses, GT is an exploratory process, where the researcher observes a social phenomenon and comes across something unusual or confusing, leading to closer investigation. This investigation is cyclical in nature as the researcher gathers data, analyses it, reflects and gathers more data, guided by emerging issues. The reflective aspect of GT is done through extensive memoing, which is a crucial aspect of conducting a GT research project (Charmaz, 2014; Hadley, 2017) as the researcher is constantly engaging with the data.

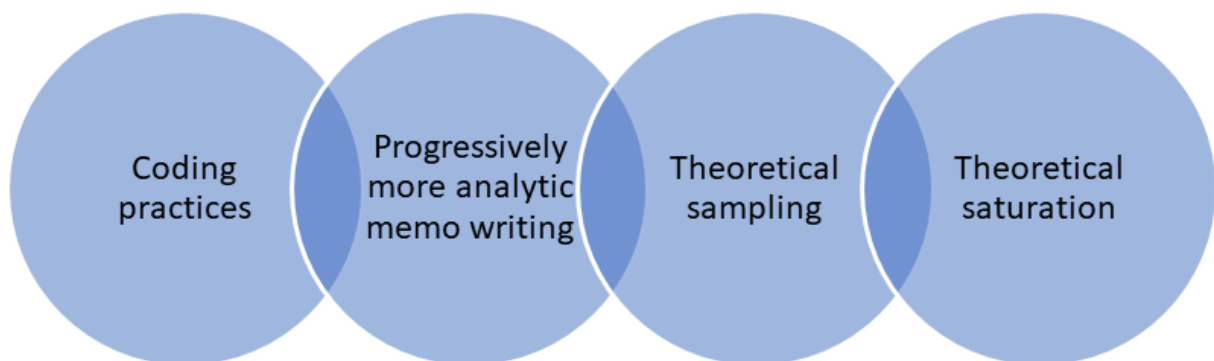
1.5.1 The Grounded Theory family

Grounded Theory could be viewed more as a family of methods/approaches rather than a single discrete method (Bryant & Charmaz, 2007; Cohen et al., 2011; Hadley, 2017; Punch & Oancea, 2014). After publication of *Discovery*, Glaserian Grounded Theory, as explained in *Theoretical Sensitivity* (Glaser, 1978) advocated avoiding the literature and purposive sampling at the outset of a research project and rather entering the research field with an open mind and no preconceptions, with the emphasis being on theory generation, with the theory emerging from the data gathered in the field. Straussian Grounded Theory, as discussed in *Basics of Qualitative Research* (2008), differs in encouraging the researcher to engage with the literature early in the research process in order to better understand the relevant issues and context (while avoiding getting too fixated on discussions in the literature and remaining open to the possibilities of data gathered in the field) and so too is purposive sampling permissible (Corbin & Strauss, 2008). The original Glaser and Strauss conception of GT could be considered postpositivist while the later version of Corbin and Strauss fitting better in the interpretivist paradigm (Levers, 2013). There have also been constructivist takes on GT, most notably Charmaz in *Constructing Grounded Theory* (2014), recognising the role the researcher has in bringing their subjectivity into the research project and in constructing a theory based on their interpretation of the data. That is not to say that a constructivist approach welcomes personal bias. As Charmaz (2008) argues, “... researchers’ perspectives will direct their attention but not determine

their research” (p. 160). Another strand of GT is Situational Analysis proposed by Clarke (2007), which looks beyond the focus on action of classic GT and looks toward the situation in which the empirical situation being studied occurs, meaning in addition to the process under study, the context in which it takes place must also be examined. There are also Critical Grounded Theory approaches, which not only try to determine what is happening in a particular social situation, but why it is happening, bringing in issues such as power, race and gender (Hadley, 2017).

Despite the variation, there is a common core which ties these differing versions together into a recognisable family of methodologies. Theory emerges from the data rather than being pre-determined and tested. Inductive and deductive approaches are used iteratively in order to construct the theory from the data (Cohen et al., 2011). Charmaz (2008) argues the overarching distinctive quality to Grounded Theory, its “comparative and interactive nature” (p. 163). She further breaks this down into four main strategies that characterize Grounded Theory (see Figure 1-7).

Figure 1-7 Charmaz's Four Main Strategies of Grounded Theory



In her landmark work, *Constructing Grounded Theory*, this expanded from four strategies to nine actions as shown below (Figure 1-8).

Figure 1-8 *Defining actions of a Grounded Theory Researcher (adapted from Charmaz, 2014, p. 15)*

1. Conduct data collection and analysis simultaneously as an iterative process.
2. Analyze actions and processes rather than themes and structure.
3. Use comparative methods.
4. Draw on data in service of developing new conceptual categories.
5. Develop inductive abstract analytical categories through systematic data analysis
6. Emphasize theory construction rather than description or application of current theories.
7. Engage in theoretical sampling.
8. Search for variation in the studied categories or process.
9. Pursue developing a category rather than covering a specific empirical topic.

Hadley (2017) broadly concurs on the importance of coding and sampling, constant comparison and theory generation but his conception highlights six core characteristics of a GT method (Figure 1-9)

Figure 1-9 *Core Methodological Characteristics of GTM (Hadley, 2017, p. 59)*

Reflexivity	Grounded theorists are open to the informants, the flow of the academic community and to their own unique research talents, cultural background and philosophical research standpoint.
Coding	The data is fractured and reinterpreted so that the grounded theorist can begin to work with it in a meaningful way. Coding and data collection take place simultaneously.
Constant Comparison	Codes and their interpretations are constantly reflected on in memos and compared with new data as it becomes available.
Theoretical Sampling	The investigation follows the theoretical implications of the data as it has been interpreted by the grounded theorist.
Theoretical Coding	The theoretical concepts are reconstructed and then linked into a framework that will be meaningful to a comprehending discourse community.
Grounded Theory	The theory will be useful. It studies the concerns of a particular group of people and the ways in which they deal with these issues. The theory results in heightened awareness, understanding, ability to predict, and/or a greater sense of control, first for the group that was studied but then later potentially for people in other situations, but who are experiencing similar problems or concerns.

What is clear is that the close interaction of data and analysis is essential in Grounded Theory, where ideas and theory are refined through constant comparison and targeted data collection, with the ultimate aim of theory generation (Hood, 2007). Of particular resonance with this researcher is Turner's (1983) contention that a theory generated through GT should be "accurate, understandable and enlightening to individuals who have some familiarity with the social phenomena under investigation, either as participants or as 'lay' observers" (p. 347). The adoption of Grounded Theory in this dissertation is to generate a theory of practical use to EAP practitioners to better equip them to assess collaborative assignments.

1.5.2 Core concepts in Grounded Theory

Before going further into the process of Grounded Theory it is important to clarify some key terminology, which can be used with quite different purposes and meanings.

1.5.2.1 *Theory*

The aim of Grounded Theory is the generation of theory. Glaser and Strauss (1967) envisaged "theory as process" (p. 32), meaning theory which is not fixed but which continues to evolve and develop, matching the constant change in social interaction. They locate the theory generated from a GT study as middle-range theory, which sits in the middle of the continuum of working hypotheses and grand theories. Apramian et al. (2017) analyse the use and meaning of theory in the GT approaches of Glaser and Holton (2007), Strauss (1987), Clarke (2005) and Charmaz (2006), identifying them as distinct schools of GT. While Glaser and Strauss have different conceptions of theory (for Glaser theory is abstract and general while for Strauss it is conceptually dense and complex) they both view theory as something which helps us describe and understand a single social process. Charmaz and Clarke move from theory to theorising, valuing the voices of participants and bringing to light stories and experiences not typically seen and heard. Rather than a formal theory describing a social

process they generate analytical frameworks to provide understanding of complex sociocultural representations of a situation.

1.5.2.2 *Emergence*

Across the different interpretations of theory there is a commonality in that the theory emerges from the data. Through analysis and coding the researcher induces nascent concepts which can be further explored and examined. The concept of emergence is not discussed in great detail and, curiously, across four foundational texts on GT (*The Discovery of Grounded Theory* by Glaser and Strauss (1967), *Basics of Qualitative Research* by Corbin and Strauss (2008), *Theoretical Sensitivity* by Glaser (1978) and *Constructing Grounded Theory* by Charmaz (2014)) emergence does not feature in the index. Levers (2013) draws a distinction between postpositivist emergence (there is an external truth and an impartial, objective researcher, through a process of constant comparison, can identify patterns and regularities that both transcend and simplify the data), constructivist emergence (the researcher is external from the data but constructs the emerging theory, which is shaped by how the researcher interprets the data) and interpretivist emergence (there is no external truth, the researcher is a constituent element in the data and the emergent theory is filtered through the researcher's interpretation). Emergence occurs through constant comparison, which is a process of joint coding and analysis, looking within and across data, comparing incidences and further refining codes to enable further comparison (Charmaz, 2014; Glaser & Strauss, 1967) and the constructivist view of emergence is held in this dissertation.

1.5.2.3 *Theoretical sampling and theoretical saturation*

To generate theory, data must be sought that can help with the understanding the subject of the study. Once nascent concepts have emerged, the researcher has the beginnings of a way of understanding what is going on. This then informs future sampling, to get data to enable a fuller understanding of the concept or category (Charmaz, 2014; Morse, 2010). The sampling is directed

toward developing the theory more completely (Breckenridge & Jones, 2009) and data collection is “controlled by the emerging theory” (Glaser & Strauss, 1967, p. 45), which may involve returning to participants involved earlier in the study. This differs from purposive sampling in that purposive sampling is driven by the desire to find participants with experience relevant to the study whereas theoretical sampling is directed towards finding data which can shed further light on the developing theory (Breckenridge & Jones, 2009) so, as Hood (2007) notes, “all theoretical sampling is purposeful, but not all purposeful sampling is theoretical” (p. 158). This means sample size cannot be determined in advance, nor is ‘representativeness’ of the group under study the aim of theoretical sampling (Hood, 2007, p.154).

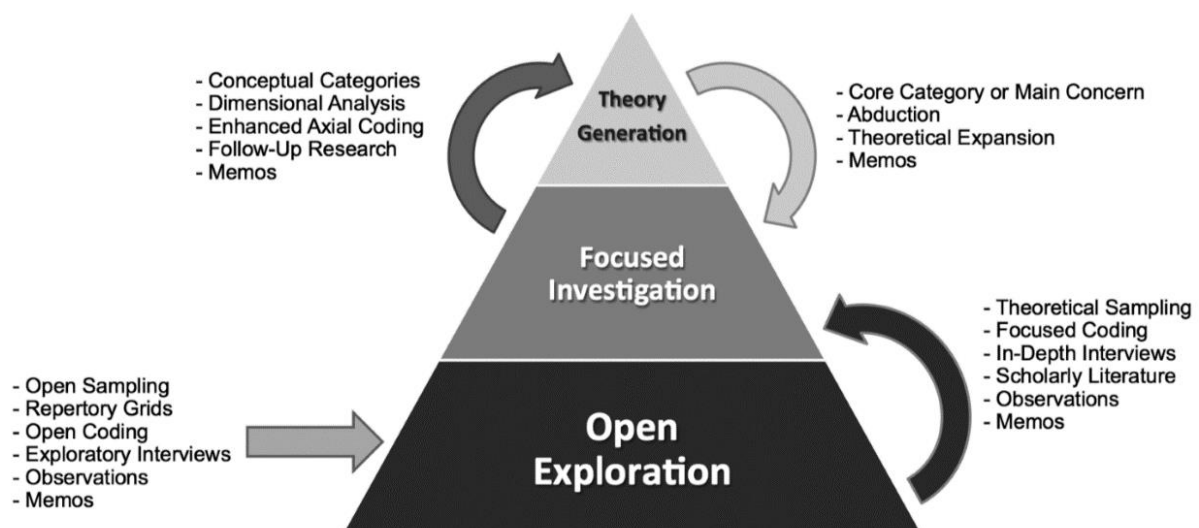
Theoretical sampling, data collection and analysis continues until the point of theoretical saturation. Theoretical saturation is the point when “no additional data are being found whereby the sociologist can develop properties of the category” (Glaser & Strauss, 1967, p. 61). In order to do this, they propose maximising variety in the groups being sampled to gain varied perspectives on the category in question. Charmaz (2014) concurs with this, defining theoretical saturation as “the point at which gather more data about a theoretical category reveals no new properties nor yields any further theoretical insights about the emerging grounded theory,” (p. 345). Dey (1999) preferred the term “sufficiency” (p. 117) as he argued ‘saturation’ suggested no new data could be found, whereas sufficiency indicates more of choice of when to stop data collection. This point occurs when “such an investment is no longer worth the trouble given the likely (theoretical) reward,” (Dey, 1999, p. 117). That is to say, sufficiency is reached when there is little perceived value in conducting further data collection given the effort that would go into it for marginal gains.

1.5.2.4 The process of grounded theory

In GT, the theory emerges from the data and as such is grounded in the data. To do this an iterative process of data collection, analysis, and reflection is followed, leading to further data collection, further analysis and further reflection (see Figure 1-10).

The initial stage of data collection takes place during open exploration (Hadley, 2017) where the researcher is open to the possibilities of the data and goes where the data leads. Through open sampling, open coding and reflection, tentative ideas emerge that are examined more closely through the next stage, that of focused investigation. In this stage the researcher is trying to understand and articulate the ideas emerging from the data and engages in constant comparison, looking for patterns within and across the data, using focused coding and categorising codes. Sampling also changes from open sampling to theoretical sampling, looking for samples and data that can further illuminate the emerging theory. This leads into the final stage of a GT study, that of theory generation, where theoretical coding is used, with further data collection going on until a point of saturation is reached, meaning no new concepts are emerging from the data and all the data fits the resultant theory. It must be noted, however, that not every GT study has to result in a complete theory but may lead to a better understanding of the phenomenon in under investigation or a conceptual framework rather than a fully elaborated theory (Timonen et al., 2018).

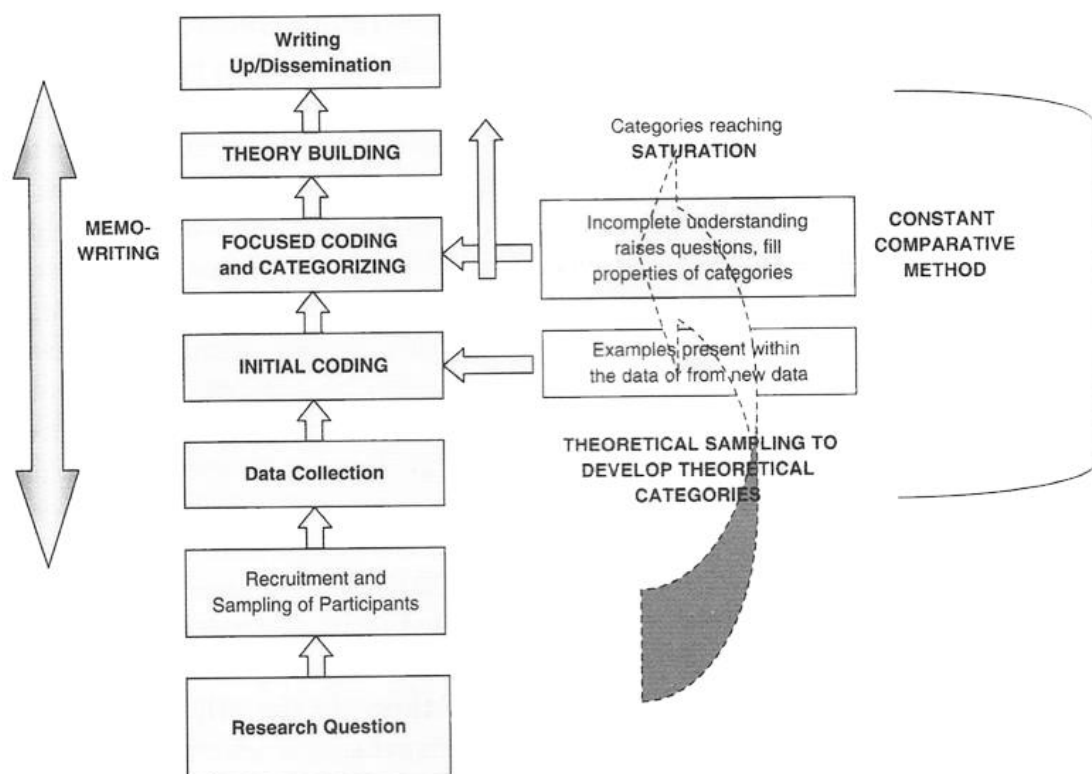
Figure 1-10 Summary of techniques used in GTM (Hadley, 2017, p.131)



This process is more fully illustrated by Charmaz (2014), highlighting the ongoing importance of memoing throughout the process (Figure 1-11). Although iterative, there is a linearity to GT. Once the

move to focused investigation has begun and the researcher is in the process of focused coding and categorising there is not a return to the open coding stage. Glaser (1978) and Hadley (2017) warn of the dangers of getting stuck generating too much data through open coding and becoming overwhelmed.

Figure 1-11 A visual representation of Grounded Theory (Charmaz, 2014, p. 18)



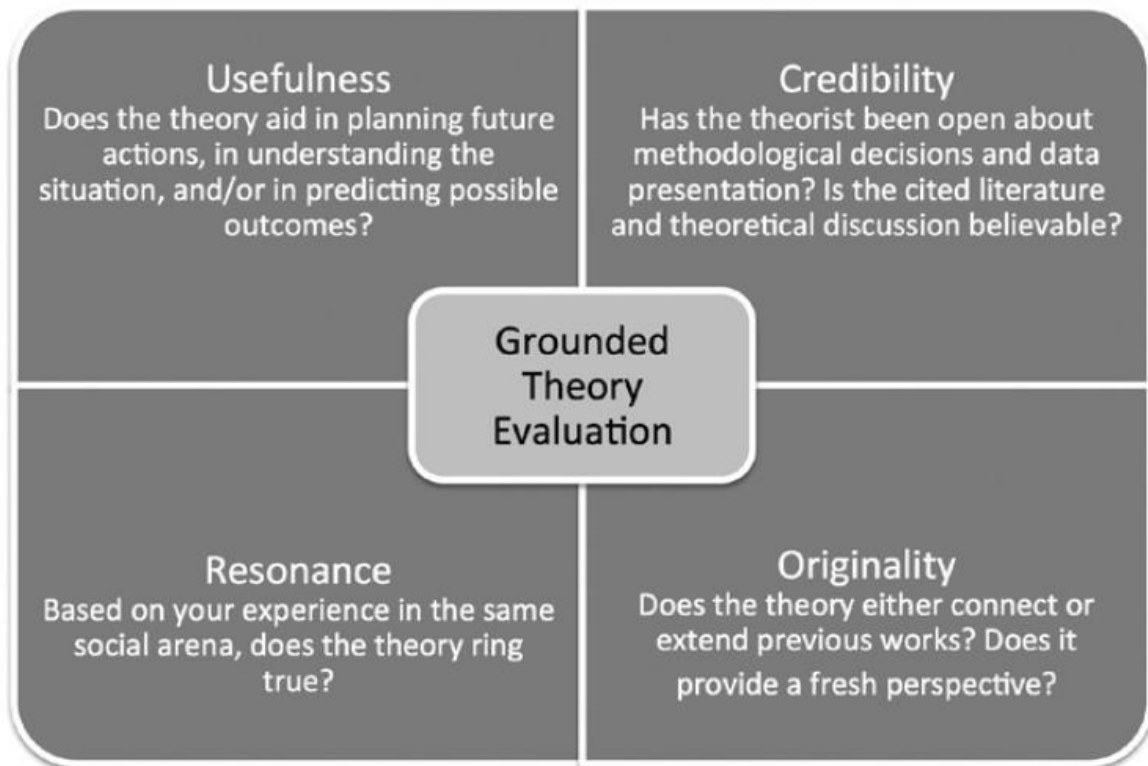
1.5.2.5 Theory evaluation

As there are different schools of GT, so too are there different evaluation methods. Birks and Mills (2023) recommend starting with *prima facie* evaluation, considering the face validity of findings to get “an overall feel for the quality of a piece of grounded theory research” (p. 53). As laid out by Glaser and Strauss in *Discovery* (1967) and expanded on by Glaser in *Theoretical Sensitivity* (1978). In that text, theory must have;

- Grab – the theory is interesting, memorable and people want to use it.
- Fit – the theory is indicated by the data and is applicable to the data and covers all the data.
- Work – the theory can explain, interpret and predict what is happening in the area under study and must be relevant to it.
- Modifiable – the theory can adapt to new data in an ever-changing world

Charmaz (2014) recognises the value of these in evaluating “how your constructed theory renders the data” (p. 337). Within the constructivist strain of GT, Hadley (2017) draws on Charmaz (2008) to highlight appropriate evaluation criteria (see Figure 1-12). He prefaces this with discussion of American pragmatism, with utility being central. As he notes, “Something is true because it is useful; it is useful because it is true.”

Figure 1-12 Matrix for Grounded Theory Evaluation (Hadley, 2017)



The purpose of the study is to generate a theory of assessing collaborative assignments practitioners will find useful and applicable and evaluation will be guided by the evaluation criteria of Figure 1-12. A *prima facie* evaluation will preface consideration of Hadley's and Turner's conceptions of what gives a grounded theory value, leading into evaluation by the classic Glaser criteria to see how the theory covers the data. However, as this is a constructivist GT dissertation, the main evaluation will be based on credibility, originality, resonance and usefulness (see 1.5.1)

1.5.3 Critiques of Grounded Theory

Grounded Theory is not without its detractors. Where a more traditional research design would involve a thorough review of the literature to identify a research gap, formulating a research question and hypotheses in light of that gap, developing research tools to test those hypotheses, ensuring triangulation opportunities to ensure the validity of data, GT can be dismissed as a "smash and grab" (Dey, 1999, p. 119) where the researcher enters the field unarmed and exploits what they can from the participants. For example, in *Theoretical Sensitivity* Glaser (1978) discusses using overhead conversations as data. Tools are always changing, and the methodological flexibility inherent in GT could lead to "methodological indifference," (Bryant, 2002, p. 25). It could be argued that in GT there is no consistency or any checks on validity. Inter-coder reliability, for example, is not a feature of GT as the aim of coding is for the researcher to fully engage with the data to allow for the theory to emerge rather than to reach an objective agreement about how a particular excerpt of data should be coded. This locates the researcher as central to the GT process, as recognised by the constructivist conception of GT.

Hadley (2017) discusses some of the main critiques of GT (it is overly complicated, overemphasises inductive reasoning and underutilizes literature) but suggests these criticisms do not take into account developments in the GT field, which has evolved beyond its origins. Indeed, Bryant's scathing critique of GT revolves around the strong positivist strain running through GT (2002) but as the field continued to develop he subsequently went on to co-edit *The Sage Handbook of Grounded*

Theory, suggesting more acceptance of the approach. Hadley recognises the strongest criticism of GT is to do with an over-reliance on data which is filtered through the researcher and could therefore, consciously or not, overlook certain aspects of the topic under consideration (Hadley, 2017, p. 46). Charmaz has rebutted this, contending it could be viewed more a failing of the researcher rather than of the method itself and that GT theory should also contend with issues to do with power and social justice (Charmaz, 2014). A researcher armed with the knowledge of the potential pitfalls of a GT approach should be able to navigate the research effectively. GT also enables research of practical application that are otherwise under studied, which makes it the ideal approach for this dissertation.

1.6 Chapter conclusion

This chapter has reviewed pertinent literature in the areas of collaboration, EAP and assessment. Collaboration is a valuable and necessary skill, which HE tries to instil in its graduates. There are twin benefits in terms of soft skill development, which can lead to greater employability, and also the affordances of group work, leading to greater learning gains for all involved.

Despite the value of collaboration, it holds an uncertain place in the academy, where its formative and development role can be in contrast to, and indeed at odds with, the need for individual summative assessment in degree programmes. Van der Vleuten et al. re-envisage standard distinction between 'formative' and 'summative' "as a continuum of stakes, ranging from low- to high-stakes decisions," (2015, p. 641), which is also highly relevant to the final findings of this dissertation in terms of informing approaches to assess collaborative assignments, which suggests the means of assessment is contingent on the importance of collaboration as a learning outcome.

The most effective collaboration, one in which students work closely together to produce a piece of work in which their individual contributions cannot be clearly identified, is the most problematic to assess. Assessment lenses which could help look inside the collaboration, such as the development of evaluative judgement or self- or peer- assessment, are time consuming in terms of

learner training and the validity of peer- and self- assessment in high-stakes summative assessment has to be considered.

Despite the challenges of assessing collaborative assignments, it is a necessity in EAP. As EAP has the function of helping students succeed in their studies, and beyond, it consequently has to engage with collaborative assignments. However, teachers may lack the knowledge and academic literacy to navigate such assignments effectively, a likely outcome as even within the research literature there is little clarity as to the best way forward. As part of this, reconsideration would have to be given to classic assessment concerns like 'validity', 'reliability' and 'fairness' to see how they can be applied in a collaborative assignment. So too is the practitioner voice missing in discussion of approaches to assessing collaborative assignments, an issue this dissertation intends to directly address.

The assessment of collaborative assignments in EAP will be examined through a GT approach, as this is an exploration rather than validation study. It is necessary to engage with EAP practitioners to discover their beliefs and attitudes to collaborative assignments in order to determine the way or ways in which collaborative assignments could justifiably be assessed. GT offers a pathway to develop a practical outcome of value to the EAP community and its students, as well as wider implications in the academy.

The next chapter will explain the methodology employed to better understand the assessment of collaborative assignments in EAP.



CHAPTER 2

METHODOLOGY

2 METHODOLOGY

2.1 Chapter overview

This chapter is going to locate this dissertation in its philosophical framework and discuss the evolution of the research from its original conception to its current form, addressing the substantial impact of the COVID-19 pandemic on the nature of the research. The complementary study will also be introduced. Although atypical, this dissertation has a symbiotic relationship with the dissertation of Bolster (2025). They both explore teacher attitude to collaboration, with the focus of this study being on the summative assessment of collaborative assignments.

The application of GT in this study will then be elucidated, identifying the guiding tenets and processes followed. The role of the writer as an insider researcher will also be explored. Once the methodological approach is clarified, the main research questions and hypotheses of the study will be outlined. The chapter will then review the ethics of the study and how the safety and anonymity of the participants was ensured. The technical aids used throughout the study will also be introduced.

The chapter will then move on to discuss the data collection process, including the sampling procedure and the research tools which were used, including those from the earliest conception of the study. At this stage, participant profiles will also be introduced. Once it is clear how the data was collected and who was involved, the approach to data analysis will be explained, charting the progression of analysis from open exploration to focused exploration.

The chapter will then turn to the latter part of the study, with theoretical coding and diagramming being central to the generation of the theory of this dissertation. The chapter closes with a return to immersion as a means of recontextualising coded data to ensure the findings of this dissertation fully chime with the lived experience of the participants and discussion of the development of the practical output: the starter guide to assessing collaborative assignments.

2.2 Philosophical grounding

The intention and shape of this research project was significantly impacted by the global coronavirus pandemic, which had a seismic impact on HE in 2020 and continues to resonate. Although the pandemic severely disrupted the original research plan, it also resulted in a wider reaching yet more focussed study. This section outlines the evolving philosophical underpinning of the study and its development from the original conception to the final Grounded Theory methodology which was followed.

2.2.1 Philosophical eclecticism

Earlier in my career when asked about my teaching approach I would have confidently stated “communicative”, which would, broadly speaking, have been true. But as my career progressed, that answer evolved, recognising the post-method condition posited by Kumaravadivelu (1994) with my teaching approach more properly described as principled eclecticism (Cushing-Leubner & Bigelow, 2014). Rather than dogmatic devotion to any one teaching method, I draw across different methods and techniques, depending on the students, context and particular teaching point. No single method will work in every classroom, every lesson, for every student and pragmatic decisions have to be made as to the best approach or approaches in each situation. Similarly, I lack devotion to a single philosophical view of the world, as at different times, each can have a value and merit.

As an illustration, consider a solid gold sphere and a crystal glass. If a solid sphere made of gold is held in the air and released, it will fall. It will not float, or fly away. It will fall. The sphere will drop, whoever holds it and whoever observes the fall. The rate of speed at which it will fall can be calculated in advance, as can the exact moment when it will hit the floor. Some things are true, measurable and predictable.

If the sphere is dropped from a height into a wine glass, the glass will shatter. This is broadly true, although the durability of the glass and the height from which the sphere is dropped may lead to an unexpected outcome. Some things are broadly true, with exceptions.

If struck by a falling sphere we know the glass will most likely shatter but the exact nature of that shattering cannot be predicted. A glass identical in appearance and composition will shatter in a different way and there is no way to know how the same glass would shatter a second time. We know roughly what would happen but the exact spray pattern of glass shards will be unique to each glass. Some truths cannot be predicted with certainty, some are both unknowable and unpredictable.

When the glass shatters, it will make a noise. This sound could be met with a flight or fight response, the shock of the unexpected, has a fight begun? Or it could be met with laughter and a round of applause and jeers. These responses to the same event could happen in the same room at the same time and be equally true and valid. To some, the sound of a breaking glass could be joyous, signify the end of a wedding ceremony, to others it could recall the accident that changed everything. The world can and will be seen and interpreted in different ways by different people.

If the sphere dropping were filmed in slow motion, the resulting video could be considered a disposable YouTube clip by some, or a deep artistic commentary on capitalism destroying itself by others. Different, opposing interpretations of the truth can be true at the same time and truth can be constructed by the shared response of people. What is true in one context may not be true in another and may not be true for every actor in the original context.

Although contradictory, there is a measurable truth, a general truth, an unpredictable, unknowable truth, multiple truths, socially-constructed truths, and context-dependent truths. There is a reality and each of us views that reality through our own perception, perception that can be changed through socialisation with others. All those realities, those which exist, those which we perceive, and those which we learn to perceive, are simultaneously real. This straddles philosophical paradigms and so, as in the classroom, my approach to research is ultimately pragmatic. What works, works or, as Saldaña put it, “I myself take a pragmatic stance toward human inquiry and leave myself open to choosing the right tool for the right job,” (Saldaña, 2011, as cited in Saldaña, 2021, p. 4-5). In relation to this dissertation, the stated problem is determining the best way to assess a collaborative assignment and this research hopes to go some way toward addressing that problem.

2.2.2 Philosophical framework of the study

The philosophical paradigm of this study has evolved from initial conception to final form. Aligning to Creswell and Poth's (2017) interpretive frameworks, the initial study design would have been postpositivist, in the sense of following a scientific method to 'prove' a theory, with elements of social constructivism, with that 'proof' being constructed by the research participants. The study was going to be deductive in nature, identifying the variables that influence attitude toward collaborative assessment approaches. In contrast, the current study design still maintains a social constructivist element, looking to discover a 'theory' rather than trying to validate or test a theory, and is interpretivist, as that theory is going to emerge from the research data. There is also a strong pragmatic element, with the important thing being that the ultimate outcome of this study is of practical use to the EAP community.

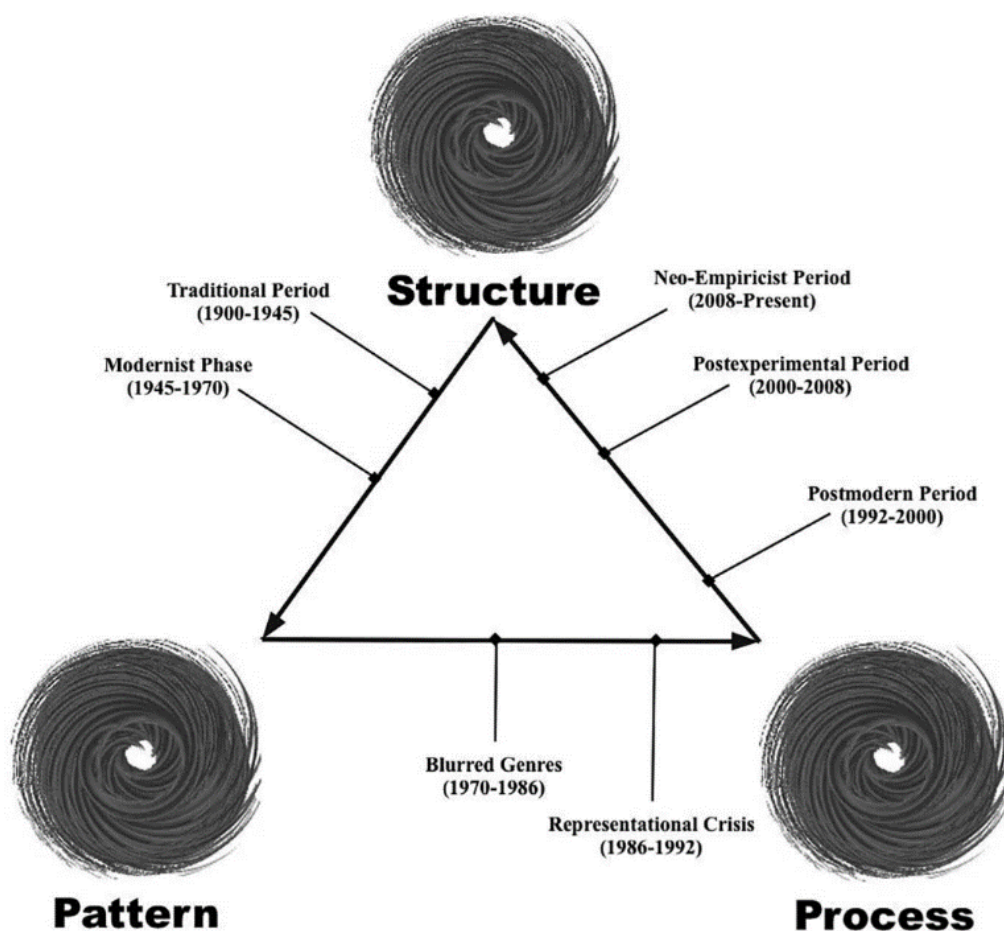
This straddling of approaches raises the issue of incommensurability, as some scholars have argued that research methods fit within a particular philosophical paradigm and are not compatible with other paradigms (see Hadley, 2017, for a review). This concern can be addressed through the lens of paradigm mapping (Lowe et al., 2005). Paradigm mapping builds on the concept of autopoiesis, which was originally developed in the field of biology as a means of describing the nature of living things through the constant interaction of structure, pattern and process. This in turn provides a metaphorical framework to interpret philosophical paradigms, as illustrated in Figure 2-1 below.

Figure 2-1 Aligning structure, pattern and process to philosophical paradigms

The paradigm of structure	Relates to positivism in the sense that there is an external truth that can be discovered. It is objectivist in that that external truth is discoverable and will be objectively true to all.
The paradigm of pattern	Relates to constructivism, that truth is constructed and is interpretivist, in that knowledge is created through consensus.
The paradigm of process	Relates to deconstructivism and critical theory, in that there is no one generalizable truth.

Hadley (2017) demonstrates how different qualitative research methods have moved, over time, from structure, through pattern, to process and back toward structure (see Figure 2-2). All of the paradigms are useful and have their place and to focus on or utilise one exclusively cannot reveal a full picture of what is happening. The world is rich and complex, which is a theme that will reoccur in the discussion section of this dissertation.

Figure 2-2 *Autopoietic Movement within the History of Qualitative Research (Hadley, 2017, p. 27)*



To locate this study in the above metaphorical framework, the initial conception had elements of structure, in that it was attempting to demonstrate a nomothetic theory, that there was one ‘best’ way to assess collaborative assignments that would be generally applicable across different contexts. However, there were also elements of pattern, in that that ‘best’ way would be socially constructed

by EAP practitioners. Once the original plan was disrupted by the global pandemic, the scope of the research widened. Rather than examining one particular context it would draw more deeply on the experience and knowledge of EAP practitioners in different teaching contexts. This maintains the pattern element from before but the intention is no longer to find a demonstrably 'best' way to assess collaboration. As per process, there are multiple valid ways such assessment could be done, multiple realities the assessment must fit. Rather than looking to provide an answer, an assessment method EAP practitioners can pick up and use, this study is looking to generate guiding questions, the questions EAP practitioners must ask themselves to generate a sound assessment for collaborative assignments in their particular context.

2.3 The evolution of this dissertation

This section provides an overview of the evolution of the study, from first origins to final form. It is important to note from the outset that this dissertation is part of a wider study into collaboration in EAP, conducted with Averil Bolster, a fellow PhD candidate, author of the complementary dissertation. The paucity of literature on practitioners and student collaboration meant the potential scope of looking into the issue would be too broad for a single dissertation since practitioner attitudes, beliefs, approaches and assessment practices are all worth investigation. As such, the two dissertations work in concert, looking at different aspects of practitioners and student collaboration (see Table 2.1 below). These studies are paired together because there is a synergy between them and to also try and mitigate research fatigue in potential participants. Both dissertations are looking at the same population, EAP practitioners, and could go through similar avenues to find participants e.g. overlapping professional networks. Given the shared interest in student collaboration it was therefore practical to present the dissertations to participants as two parts of a greater study and use a shared initial survey, which would enable the researchers to then independently select participants for interview. To make it clear to participants that data collected would be available to both dissertations, Averil Bolster was listed as a secondary researcher for this dissertation although, in

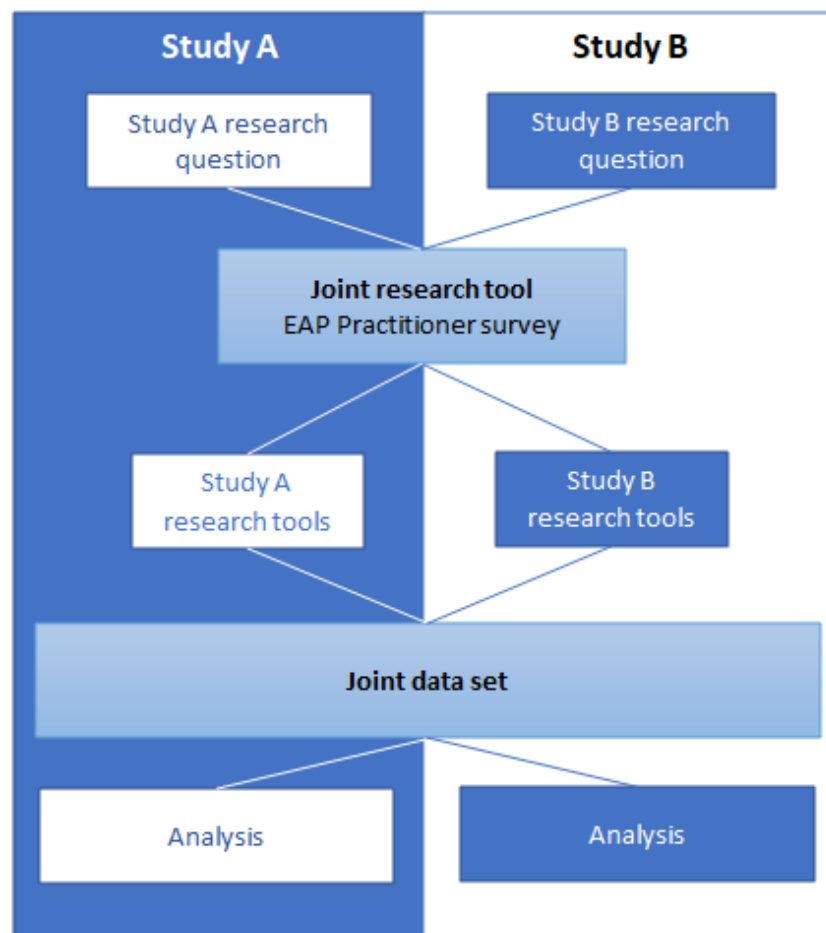
practical terms, this was limited to having access to data collected but playing no role in data collection, analysis or for this dissertation.

Table 2.1 *Collaborative assignment complementary dissertations*

Dissertation A		Dissertation B	
Title	A qualitative Grounded Theory study of EAP practitioners and collaborative learning: Identity and beliefs	Title	Exploring practitioner collaborative assessment identity to develop a principled multi-lens approach to assessing collaborative assignments in EAP
Lead researcher	Averil Bolster	Lead researcher	Peter Levrai
Focus:	EAP Practitioner attitude to and beliefs around collaborative assignments	Focus:	EAP Practitioner attitude to assessment of collaborative assignments

While there was cooperation on the design of the initial survey, all subsequent data collection tools were generated separately, and the researchers would then go on to independently gather data toward their own research questions. This generated a large data set which was available to both researchers for independent analysis and interpretation (see Figure 2-3), with transcripts from the complementary dissertation serving as secondary interview data in this dissertation (see Appendix 5).

Figure 2-3 *Research tools and data in the complementary studies*



Across these dissertations, it is hoped to gain a better understanding of EAP practitioner attitude to student collaboration and what informs those attitudes, from conception, through facilitation, to assessment. Although related, there is a clear delineation between them. Bolster's dissertation is looking more toward the start of a student collaboration, examining EAP practitioner attitudes and what informs those attitudes. Assessment may be discussed as a factor which influences practitioner attitudes towards collaborative assignments, since assessment is recognised as a contentious area as previously discussed in 1.4.6, but only as one of many factors.

The focus of this dissertation is very much on the assessment of collaborative assignments and finding a means to do so fairly. Teacher attitude to collaboration will have a role, as will how collaborative assignments are set up and facilitated, but only inasmuch as they influence beliefs about and approaches to assessment.

2.3.1 The rise and fall of the original research plan

The original research plan intended to re-examine the context of the initial survey into student attitude toward collaborative assignments (see *Origins of this dissertation*) using more robust research methods. Although the researchers had left the institution, collaborative assignments were still in use. Research in this institution would potentially provide access to multiple classes of students, each containing multiple essay groups, moving through collaborative assignments with different teachers. Approval was gained for conducting the study from the head of the English Language Centre (ELC) in University of Macau (UM) on 12th June 2019, pending approval from the institution's Ethics Board.

The application to the Ethics Board was made on 9th August 2019, requiring the submission of the research applications, consent forms and all research tools (see Appendix 7). There was clear delineation between the research plans for the two dissertations, with my focus on the assessment side of collaborative assignments. Prior to approval of the Ethics Board, the Dean of the Faculty of Arts and Humanities had to give approval, which happened on 6th September 2019 and the Ethics Board granted their approval on 9th October 2019 (Appendix 5). This was too late in the semester to start data collection as the initial survey was designed for use with students before starting the collaborative assignment to gauge their baseline opinion. Consequently, it was planned that data collection would start in the spring semester of 2020.

2.3.2 The approach of the original research plan

The aim of this research, which has remained steady through changing circumstances, is to develop a better understanding of how collaborative assignments are assessed in EAP, and move towards best practice in the assessment of collaborative assignments. In the original conception, this would be achieved through two main channels of data collection. The first would be examining a specific EAP context with a mixed methods approach and the second would be interviewing teachers from the wider EAP community. Examining a specific EAP context would give access to students and assessment results, speaking to the research need identified by Meijer et al. (2020) in determining the

impact of collaborative assignments on students' course grades. The teacher voice could be enriched by engaging with EAP practitioners in varied contexts.

2.3.3 Original research questions and hypotheses

The original research plan outlined four main research questions:

1. Which models of grading collaborative assignments used in EAP are perceived by teachers and students to be the fairest?
2. How can an individual (best) be assessed in a summative collaborative assignment?
3. Can, and should, contribution be assessed in a collaborative assignment and, if so, how and by whom?
4. To what extent does computer mediated communication provide the teacher with the affordance to assess individual contribution to a summative collaborative assignment?

and suggested a number of hypotheses

- Attitude to collaborative assignments will depend on the type of assignment. Teachers will perceive oral collaborative assignments to be fairer than written collaborative assignments due to the greater visibility of the individual in an oral assignment.
- Different models of grading collaborative assignments will be perceived differently by teachers.
- When comparing assessment models, including individual tasks and clear grading of contribution will be favoured by teachers and students
- Teachers and students will support the assessment of contribution and that this should be done by both tutor and peers

- Contribution will mainly be measured by the teacher in terms of visibility of production rather than acknowledging various ways an individual can contribute to a team effort
- Peer-assessment (both inter- and intra-group) will be perceived as more problematic the higher the stakes of the assignment, particularly for credit bearing or gate-keeping summative collaborative assignments.
- Collaboration itself is difficult for the teacher to assess and although CMC increases the possibility of teachers assessing contribution to collaborative tasks it will not be felt to be adequate (some interactions and collaborative processes will remain invisible to the teacher) or the workload will be onerous.

It must be noted that the original research plan was essentially a validation study in terms of justifying and supporting an approach to collaborative assessment previously developed by the researchers. It was also envisaged as quite a traditional research design, an approach encouraged by the ethics approval process, which did not have a lot of scope for emergent methods or tools. The research plan, and research tools, were fixed at the time of application.

2.3.4 Original research plan data collection tools

In the original study a sequential integrated mixed methods approach was envisaged (Cohen et al., 2011), with different research tools gathering qualitative and quantitative data, which would be synthesised in the findings. Neither qualitative nor quantitative data would be favoured, and it was intended that the data gathered through different tools would be “mutually illuminating” (Bryman, 2007, p. 8). The tools are briefly outlined below and are collated in Appendix 7.

2.3.4.1 *Teacher and student surveys*

Broadly speaking the surveys aimed to gather bio data, general attitude to collaboration, and experience of, and attitude to, collaborative written assignments specifically, gauged through both quantitative and qualitative questions. This would provide insight into issues and areas to discuss in the subsequent interviews/focus groups. The use of multiple surveys (three with students, two with teachers) was designed to allow for the identification of any change in attitude as the collaborative assignment progressed.

2.3.4.2 *Teacher Interviews*

Teachers would be interviewed about their attitude to collaborative assessment once the assignment was complete and would follow a 'think aloud' protocol during the interview to evaluate different models of collaborative assessment. A "think-aloud" approach was selected to avoid any conscious or subconscious guidance of the participants by the interviewer through questioning. A think-aloud protocol can also be a means of potentially accessing higher-level thinking processes (Olson et al., 1984). The think-aloud was planned as part of the interview rather than as a separate task, principally on logistical grounds. As research would be done remotely, the material for the think-aloud would have had to be sent to the participant electronically and the participant would have had to follow the protocol, record themselves, and return that recording to the researcher. Incorporating the think-aloud into the interview streamlined the process and lowered the burden on the participants.

2.3.4.3 *Student focus groups*

Students who had been through a collaborative assignment would be invited to a focus group to discuss collaborative assignments generally and consider different ways collaborative assignments could be assessed. A focus group format was preferred to a 1:1 interview so students would have support from other students during the interview and could also build on each other's ideas. Rather

than a group interview where there is more interviewer/participant discussion a focus group leads to more participant/participant discussion (Brown & Edmunds, 2011), which was considered a more fruitful means of finding out what students really think.

2.3.4.4 *Gradebook analysis*

The course in question consisted of a number of assignments, individual and collaborative, and grades were kept in a Gradebook in Moodle. This would allow statistical analysis of students' grades in individual assignments and collaborative assignments and potentially enable analysis of the extent to which collaborative grades diverge from individual grades. Of particular interest would be the performance of high performing students and if their overall grade were reduced because of collaborative assignments and also that of lower performing students and if their overall grade were increased because of collaborative assignments. Essentially, it was hoped to discover whether collaborative assignments have a significant impact around important thresholds, such as achieving a distinction or passing a course. If collaborative assignments could be shown to have a meaningful impact on overall grades it would justify the need for further study and research into assessing them as soundly as possible.

2.3.5 *The pandemic interruption*

However, on 31st January 2020 the target institution announced the decision to move into distance learning in light of the worsening COVID-19 pandemic developing in China (University of Macau, 2020) and shortly after the ELC made the decision to move to individual assignments rather than collaborative assignments in the new teaching mode. This meant the original research plan had to be abandoned and a new plan generated which would be practically feasible during a worsening pandemic situation and the massive shift in EAP delivery from face-to-face to remote learning that was taking place worldwide. This required a rethink of the whole research plan and necessitated the pivot into the current study.

2.4 The current research plan

After the withdrawal of the target university in the study, the research plan had to be reconsidered. An important decision made at this time was to focus the study onto teachers only. The rationale for this was both practical and ideological. On the practical side, teachers would be more accessible than students and institutional ethical approval would not be needed or would be less arduous. The plan adhered to the ethical principles of research with human participants and ethical review in the human sciences in Finland (Finnish National Board on Research TENK, 2019) since the researcher is based in University of Turku (UTU) and also with those of the University of the Basque Country, so teachers themselves could give informed consent to participate. Ideologically, the aim of the research is to generate a framework for assessing collaborative assignments and teachers would have an essential voice in that. As per the original plan, the researcher's network of colleagues and acquaintances through conferences could be used as an avenue to find research participants to provide a wider range of attitudes to student collaboration. The change in focus would mean fronting the teacher voice, rather than it adding depth to the findings from a particular research context. This change unfortunately means no longer incorporating the student voice, which would also have great value, but this could and should be sought later, in future studies. Student voice is also already more present in the research literature than that of the teacher (Zheng et al., 2021).

As a legacy of the original research plan, there was an interest in comparing teacher attitude to collaborative assignments within the same teaching context. This could perhaps identify whether attitude to collaboration was part of a "teaching culture" of a specific context or whether individual attitude trumps context culture. Given the disruption caused by the pandemic in HE in Asia, it was decided to refocus onto a geographically more accessible location and approach teachers working in institutions in Finland. In May 2020 eleven Language Centre Heads in universities around Finland were contacted to ask if teachers in their teaching unit could be approached and seven of them gave their approval. There was, however, very little uptake from teachers. In part this was due to the ongoing pandemic situation, with teachers adjusting to remote teaching and also, potentially, preparing for

the much-needed upcoming summer break. This meant that the examination of the impact of teaching culture on teacher attitude could only be accessed through interview rather than survey comparison.

Focussing the study on the experience and attitudes of teachers changed the nature of the study. Rather than doing a deeper study into one localised teaching context, a broader study could be carried out across different teaching contexts. As the intent of the study is to determine a framework of assessing collaborative assessments in EAP, bringing in voices from different EAP contexts is a strength. Rather than a study whose findings may be highly context bound, the opportunity opened up to conduct a study which could be more widely generalizable. This led to the adoption of GT as the driving methodology of the study as it has the advantage of being “coordinated, systematic but flexible,” (Punch & Oancea, 2014, p. 169). It is also an approach well-suited for examining phenomena which are under theorised or uncharted in the literature (Charmaz, 2008; Chun Tie et al. 2019; Thurlow, 2020) and for finding practical solutions to real world problems (Hadley, 2017), lining up with the pragmatic intent of this study.

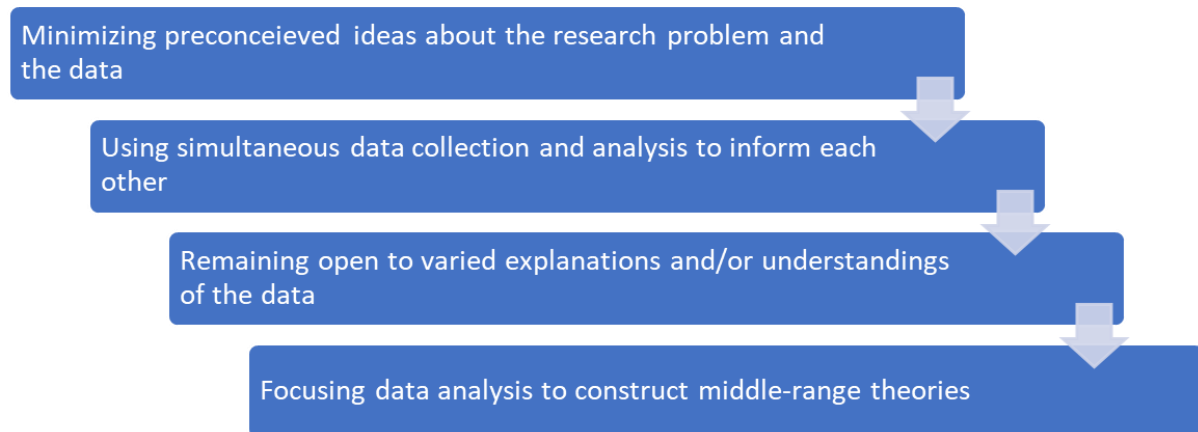
2.4.1 Application of Grounded Theory in this dissertation

This dissertation is going to follow a Straussian flavoured constructivist GT approach. It is Straussian in the sense that there is a role of literature throughout the process and constructivist in that I am a practitioner of EAP with experience of collaborative assessments and am, therefore, necessarily subjective and involved in the interpretation of the data (Charmaz, 2014). Being aware of the subjectivity and being reflexive operates as a check on biases, ensuring I not only look for things that suit the preferences I have. This can be illustrated in how early in the process of data collection I came to realise the limitations of the assessment approach I originally favoured.

GT is not being viewed as a prescriptive step-by-step method that has to be adhered to, but rather as a framework of “flexible yet systematic guidelines” (Charmaz, 2008, p. 161) or “constellation of methods,” (Charmaz, 2014, p. 14). Those guidelines are the core methods and characteristics

identified by Charmaz (see Figure 1-8 in 1.5.1 above) and Hadley (see Figure 1-9 in 1.5.1 above), operating within what Charmaz called the fundamental tenets of GT (see Figure 2-4).

Figure 2-4 *Fundamental Tenets of Grounded Theory* (Charmaz, 2008, p. 155)

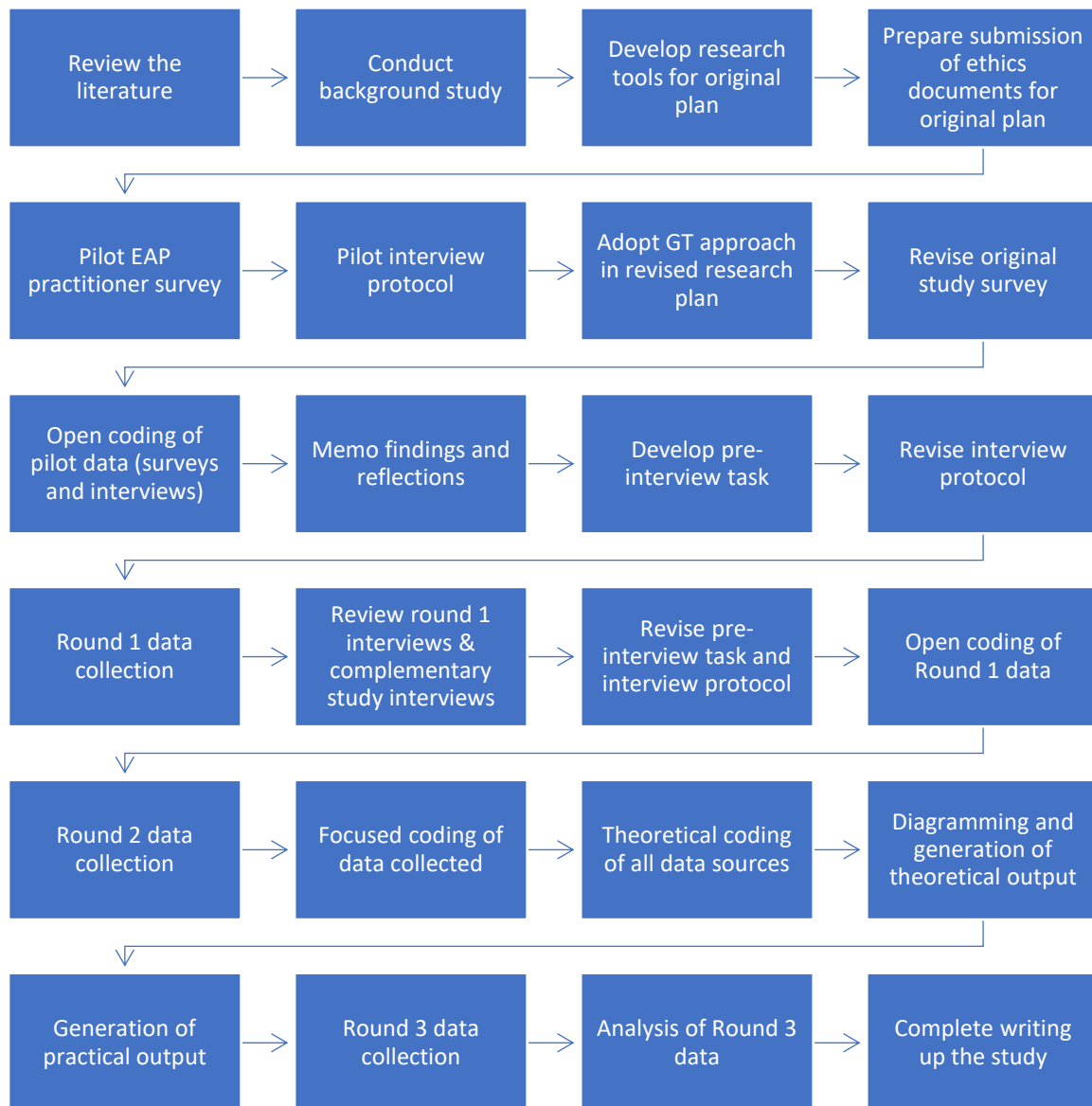


It was these tenets that guided the methodology of this study, along with methods, characteristics and strategies outlined earlier. Of particular importance is Action 6 in Figure 1-8 (see 1.5.1 above) in terms of emphasizing theory construction, “Emphasize theory construction rather than description or application of current theories” (Charmaz, 2014, p. 15).

2.4.1.1 The process of this dissertation

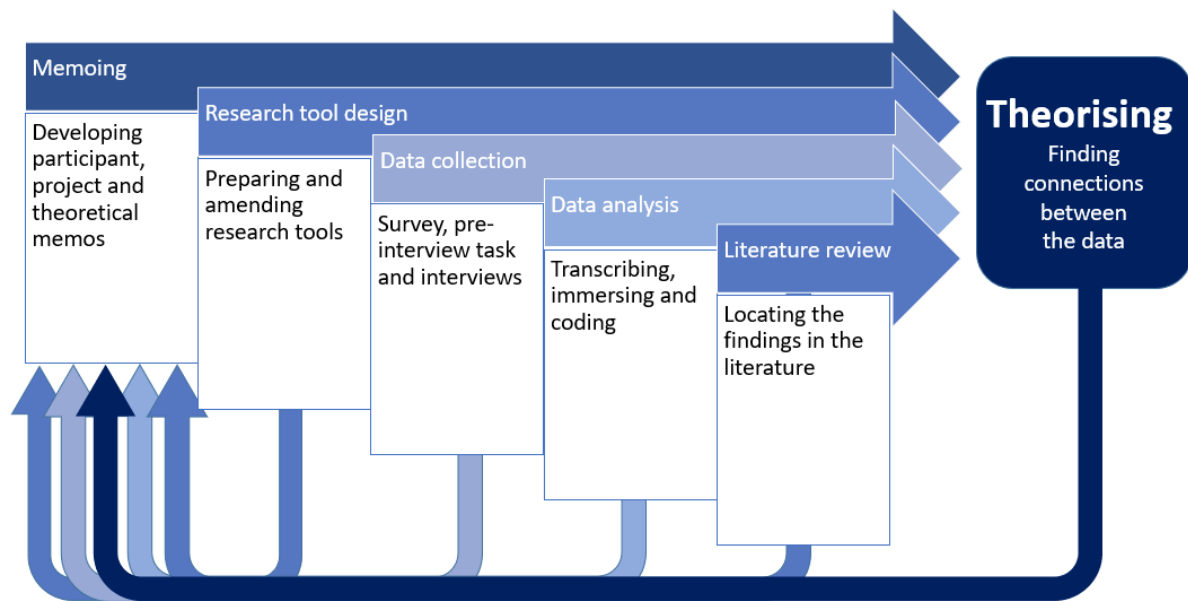
The process followed in this study is illustrated in Figure 2-5 below and broadly follows the processes illustrated by Hadley (2017) and Charmaz (2014) but also covers the events preceding the adoption of a GT approach.

Figure 2-5 *The process of this study*



An issue with Figure 2-5 is that it suggests a linear process, when it must be noted that many processes were concurrent, with memoing happening throughout. During Round 1 and Round 2 of data collection, for example, several processes happened simultaneously, as illustrated in Figure 2-6. An important element to note is how memoing encompasses all other processes.

Figure 2-6 Concurrent processes in Grounded Theory



Research tools were developed to gather data, utilised, the data analysed, leading to theorisation, which resulted in the tools being refined or new tools being developed for further data collection, leading to further analysis and reanalysis of previously gathered data, enriching the emerging theories, which were then considered in the context of the literature.

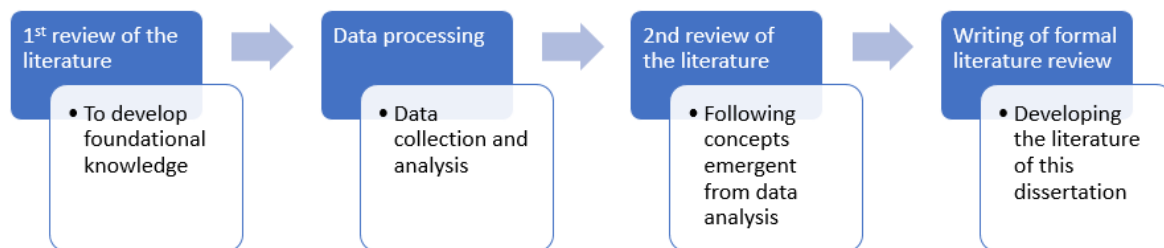
2.4.1.2 Approach to literature

The original conception of this study was not going to follow a GT approach. Therefore, to identify potential research gaps and help with the formation of research questions and hypotheses an extensive range of literature was consulted on relevant areas like EAP, collaborative learning and assessment during the first stage of the study. This reading was not written into the formal literature review for this thesis at that stage, but informed the development of the original research tools and research plan.

With the change of methodology to GT there was a concern that, in the classic form of Grounded Theory, review of the literature should follow after data collection and analysis (Glaser & Straus, 1967). However, the Straussian approach to GT allows for the review of literature earlier in the research

process, to help provide the researcher with the conceptual knowledge to increase their theoretical sensitivity (Corben & Strauss, 2014) and gain what Charmaz (2008) termed “intimate familiarity” (p. 162) with the phenomenon under study. As such, the reading of the literature was not at odds with a GT approach but the decision was made to delay formal development of the literature review until after iterative cycles of data collection and analysis had been completed. One benefit of this is, as Glaser posits, everything is data, including the literature (Glaser, 1998). This means the literature could be reviewed a second time and folded into the developing theory emerging from the data. This also meant that concepts which had not been envisaged in the initial research for the study e.g. teacher assessment identity, could be incorporated in the literature. The approach to the literature is illustrated in Figure 2-7 below.

Figure 2-7 *Approach to Literature in the study*



2.4.1.3 *Role of the insider researcher*

A constructivist approach to GT recognises the role of the researcher in the study and a key feature of this dissertation that must not be overlooked is the identity of the researcher as an EAP practitioner and, therefore, an insider through being part of the target population under study. The researcher has also had practitioner experience of collaborative assignments and part of the prompting for this study is awareness of the practical concerns practitioners have around collaborative assignments. This provides two of the benefits of the insider researcher identified by Brannick and Coghlan (2007), those of access to colleagues and pre-understanding. The advantages of the insider

researcher are also summarised by Breen (2007, p. 163) when reflecting on her own position in her research project;

- Greater understanding of the group under study
- More natural interaction with members of the group under study
- Greater relational intimacy with the group under study
- Less power differential between the researcher and the group under study

She also recognises that each of these advantages carries a disadvantage in terms of losing objectivity, of making assumptions about sameness and difficulty of balancing the roles of insider and researcher and ethical concerns of being a member of the context under study. For example, Humphrey (2013) notes how no student declined to participate in a study run by an insider researcher.

Some of the dangers of the insider researcher are mitigated by the broad nature of this research project. Rather than being focused on one specific institutional context of which the insider researcher is a part, it is a study across different institutions in different teaching and country contexts. The insider status is that of a general domain practitioner rather than a colleague working in the same institution. This also balances out any power differential and lowers sensitivity when talking about specific institutional examples. Perhaps more pertinently, is the observation made by Dwyer and Buckle (2009) when discussing the insider/outsider debate, “the core ingredient is not insider or outsider status but an ability to be open, authentic, honest, deeply interested in the experience of one’s research participants, and committed to accurately and adequately representing their experience” (p. 59), which neatly summarises the position of this practitioner-researcher.

2.4.2 Research questions and hypotheses

An important feature of GT is the different approach to research questions and hypotheses compared to other research approaches. Phase 2 of this study started with an overarching research question, keeping to the theme of the original study.

Research question: How can we, if we can, best assess collaborative assignments in EAP?

The original hypotheses have no role in this study. Rather than looking for affirmation or anticipating possibilities, the purpose of the study is to see what is really there. It also became clear through the pilot interviews that my biases had to be set aside. Even early discussions on different models of assessment highlighted that no one method, even that which I would advocate for, was without its problems. More importantly, models which I would instinctively dismiss were also shown to have a purpose and value to the EAP community that I had not seen in them. By not having preconceived hypotheses I was able to fully engage with the perspectives and responses of other practitioners.

Through the process of data collection and review, other areas of interest emerged.

- Is collaboration a stable 'skill'?
- Is it possible to assess collaboration since it is a complex construct and can never be fully transparent, even to those within it?
- What is needed to make teachers more comfortable and confident assessing collaborative assignments?

This research is a process of inquiry and is not looking for the 'right' way to assess collaborative assignments but is more looking for possible ways to assess collaborative assignments and the rationale and justification for each, with the hope of suggesting fairer ways to assess collaborative assignments.

2.4.3 Ethics

This study was carried out in adherence to the guidelines of the Finnish National Board on Research TENK. Participants were informed of the purpose of the study at the outset, participation was entirely voluntary, and participants were free to withdraw at any time. Participants were given the option in the original survey whether they wanted to be considered for interview for either Bolster’s research, or mine, both, or neither. It was also made clear to participants that data collected would be available to both researchers, with the data collected by Bolster (2025) serving as secondary data in this dissertation.

2.4.4 Consent

Participants gave their consent to participate prior to completing the survey, when they could also choose to consent to be invited for interview to this dissertation, the complementary dissertation, both dissertations or neither (see page 1 of the surveys in Appendix 1). The majority of participants who completed the survey agreed to take part in both interviews.

2.4.5 Anonymity and confidentiality

Participants were guaranteed confidentiality and all data is anonymised to prevent identification (see Table 2.2 below). Data was kept in a password protected laptop and tools operating within the more secure UTU environment were preferred.

Table 2.2 *How data was anonymised*

Research participants	given a pseudonym, taken from the Atlantic tropical storm name list from the World Meteorological Organization
Institutions	identified by the country they operate in
Colleagues or students	names have been redacted and replaced with a letter identifier e.g. “Colleague A”, “Student B” etc

Data collected in this study was used for research purposes only and there are no financial factors in play.

2.5 Technological aids

This research was conducted with the aid of technology for word processing, data collection, transcription and data analysis. No large language models (LLMs) or generative AI were consulted.

2.5.1 Microsoft Office Suite

The Office365 suite was accessible through a secure UTU account. This gave access to Word and PowerPoint, for the drafting of research tools and the final thesis. Microsoft Forms was also used for the pre-interview task. The reason it was preferred to Webropol, which was the main survey tool available in UTU, is due to additional design features in Microsoft Forms, which made inputting images and formatting around them easier. There was also a question type, the star ranking system, which was not available in Webropol, but is an intuitive rating system for the models of collaborative assessment under consideration. All communication with participants went through the official UTU email account via Outlook.

2.5.2 Webropol

Webropol is a Finnish survey and reporting tool, which could be accessed using UTU credentials. This gave an extra layer of security and privacy. It was used for the student and teacher surveys in the original research plan and the EAP practitioner survey used throughout this study.

2.5.3 Zoom

Zoom is a video conferencing software, which could be accessed through UTU credentials to a more secure version of Zoom than a typical public license. It was used for conducting interviews and

the interviews were recorded using the native Zoom recording feature. Even when interviewing practitioners in the local Finnish context a Zoom interview was preferred, given both the ongoing pandemic and recording capability.

2.5.4 NVivo12 & NVivo20

CAQDAS software NVivo12 (Lumivero, 2024) was used as a tool to assist with open coding. Focused coding and analysis were conducted in NVivo20 (Lumivero, 2024b) as there was a cloud service option so the project could be accessed through secure UTU credentials from either an office based computer or a work issued laptop.

2.5.5 WhatsApp

Due to a technical difficulty the main interview with one of the participants, Jerry, was not recorded successfully. At his suggestion, audio of key questions was sent via WhatsApp and he send audio recorded answers back. Although not ideal, and outside of the UTU ecosystem, it was a more effective method of recovering some of the interview data then relying on recall. WhatsApp was also employed for his participation in the final round of data collection.

2.5.6 Otter.ai

The pilot interviews were transcribed by hand but later interviews were transcribed using Otter.ai (Otter.ai, 2024a), a tool offering speech to text transcription. While this certainly had a speed advantage over fully manual transcription, texts still needed careful proofing and editing in line with the audio recording of interviews. This was invaluable in giving the researcher a first run through of the data on a global level. The Otter.ai (Otter.ai, 2024b) mobile phone app was also used for memoing. This meant orally recited thoughts and ideas could be captured in a written text format in real time. Again, this text would need reviewing, which was an opportunity to flesh out and refine ideas.

2.5.7 Samsung Notes

Otter.ai mobile app (Otter.ai, 2024b) provided the possibility of recording oral memos and another mobile phone app, Samsung Notes, was also used to capture important ideas when they occurred away from a computer or notebook. These forms of memoing complemented handwritten and word-processed memos.

2.6 Data collection

This section will discuss the process of data collection, the research tools used, and provide a profile of the participants.

2.6.1 Sampling

A driving force behind GT is the use of theoretical sampling, that is sampling that helps you develop the theory. You go where the data takes you and search for opportunities and participants who can cast further illumination on the emerging theory or theories. Prior to this, at the outset of the study, initial sampling is targeted towards participants who can help you understand the phenomena under study.

2.6.1.1 Initial sampling

For the first step in data collection, the EAP practitioner survey, purposive sampling was used as there was not a discrete context under observation. This is not uncommon in GT, as recognised by Birks and Mills (2023). Purposive sampling is a form of non-probability sampling where participants are selected for the knowledge, experience or judgement they can provide (Cohen et al., 2011). The potential participants invited to participate in the study were all active EAP practitioners, who may or may not have had experience of collaborative assignments.

Four participants were selected for the pilot round of data collection, two local colleagues and two colleagues working in institutions in the UK. In Round 1 of data collection invitations were sent

via mailing list to participants of the background study (1.3.1), Heads of language centres in Finland, and individual contacts. In Round 2 of data collection, only individual invitations were sent. The number of invitations sent and responses are summarised in Table 2.3 below.

Table 2.3 *Summary of invitations to participate and responses*

	Channel	Invitations Sent	Invitations accepted	Surveys completed	Interviews conducted
Pilot	Direct individual invitation	4	4	4	4
Round 1	Mailing list invitation received by multiple recipients	1*	5	5	2
	Institution invitation received by multiple recipients	11*	7	8	6
	Direct individual Invitation	15	8	8	4
Round 2	Direct Individual Invitation	8	11	8	4

**These invitations were received by multiple members of the EAP community e.g. everyone on the mailing list or in the EAP team in the institution so one invitation could result in multiple responses.*

The most successful channel for reaching participants was through direct individual invitations rather than through institutional or mailing list invitations. Purposive sampling continued when determining which of the survey respondents to invite to interview. Teachers with experience of collaborative assignments, or who had strong feelings about collaboration, were invited to the next stage of the study.

2.6.1.2 Theoretical sampling

After the pilot and first rounds of data collection, nascent concepts had emerged, which enabled the switch into theoretical sampling. The aim of theoretical sampling is to gather data which will help with understanding a nascent concept or emerging theory. In practical terms, this means choosing

where to get data (e.g. who to interview) and what data to collect (e.g. what to ask in the interview) (Birks & Mills, 2023; Charmaz, 2014). It is the Birks and Mills (2023) conception of “identifying and pursuing clues” (p. 140) which perhaps resonates most strongly. This can be seen in changes to the pre-interview task and in the loosening of the interview protocol, which brought in questions to ask interviewees to reflect on what previous participants had said, enabling investigation of concepts like performative collaboration and the stability of collaboration.

2.6.2 Research tools

As the nature of the research changed in the evolving pandemic situation, so too did the research tools. The research tools relating to student experience were no longer required, nor would statistical analysis of the Gradebook be a feature of the final studies. A mixed method approach was still preferred, although this time a fully integrated mixed methods design rather than a sequential mixed methods approach. A fully integrated approach suits the iterative nature of GT, where tools can be refined and new tools can evolve as the study progresses. While GT developed in part as a defence of qualitative research methods, quantitative data can and does play a role in GT studies (Bryant & Charmaz, 2007; Corbin & Strauss, 2008; Glaser, 1978).

2.6.2.1 EAP Practitioner survey

The EAP practitioner survey was co-developed with Averil Bolster, lead researcher of the complementary study. The survey was designed to gather data in three general areas across eight categories (see Table 2.4 below). The general shape of the survey stayed the same across different iterations, although for Round 2 data collection the digital literacy section was collapsed into the professional profile and the focus of the third section of the survey encompassed collaborative oral assignments as well as collaborative written assignments.

Assessment was deliberately not a major feature of the EAP practitioner survey. This was partly due to concerns about length of the survey and also the potential washback from considerations about

assessment on the rest of the results of the survey. Assessment of collaborative assignments can be a contentious issue and raising awareness of the difficulties/challenges through a survey could have an impact on attitude to collaborative assignments, skewing the data for the complementary dissertation. An additional consideration was wanting to get some qualitative data from practitioners about assessment through interview which could then, potentially, inform a potential qualitative research tool.

Table 2.4 *Survey Coverage*

General area	Rationale	Category
Contextual	To get an understanding of the participant and their teaching context	Personal profile Professional profile Student body profile Digital literacy
Collaboration generally	To gauge participants' general attitude to collaboration	Working with other people
Collaborative written assignments specifically	To gauge participants' attitude to collaborative written assignments specifically	Collaborative written assignments (experience / expectations) Collaborative written assignments – practical issues Attitude to collaborative written assignments

The survey (included in Appendix 1) was used to inform theoretical sampling to identify participants to invite for interview. While inexperienced teachers would have useful input as well it was determined in the first instance to focus on teachers who had experience of collaborative assignments, or had strong feelings about them. There is potentially rich qualitative data in the surveys, but data analysis is focused on the development of theory arising from the pre-interview task and interview transcripts. There is no statistical analysis of the survey as this dissertation is not searching for statistical significance.

2.6.2.2 Pre-interview task

The aim of the pre-interview task was for teachers to consider different ways a collaborative assignment could be assessed and to evaluate them. This includes consideration of both different models of collaborative assessment and the role, if any, of peer- and self-assessment. The participant pre-interview tasks are collated in Appendix 3.

2.6.2.3 Pre-interview task coverage

The coverage of the pre-interview task stayed broadly the same across the different iterations but there was some reorganisation and expansion, as summarised below, in Table 2.5.

Table 2.5 Summary of the general coverage of the three iterations of the pre-interview task

Pilot Round Pre-interview task coverage	Round 1 Pre-interview task coverage	Round 2 Pre-interview task coverage
• Task Overview • Model 1 – Flat grade • Model 2 – Flat grade plus individual assignments • Model 3 – Flat grade plus contribution grade • Model 4 – Weighted contribution grade • Own model of assessment • Teacher, peer- and self-assessment	• Co-operation vs collaboration • Task overview • Model 1 – Group grade • Model 2 – Group grade + complementary individual assignments • Model 3 – Group grade + collaboration grade • Model 4 – Group grade x weighted collaboration grade • Teacher, peer- and self-assessment	• Co-operation vs collaboration • Task overview • Model A – Group grade • Teacher, peer- and self-assessment for Model A • Model B – Group grade + complementary individual assignments • Teacher, peer- and self-assessment for Model B • Weighting of marking for Model B • Model C – Group grade + individual collaboration grade • Teacher, peer- and self-assessment for Model C • Weighting of marking for Model C • Model D – Group grade x weighted collaboration grade • Teacher, peer- and self-assessment for Model D • Own model of assessment • Criteria for assessing collaboration • Evidence of collaboration

The co-operation vs collaboration section was added for discussion in the interview, considering how teachers *expected* students to approach a collaborative task and how, in their experience, students *did* approach a collaborative task. Names of the models changed slightly, to clarify the nature of the approach to assessment. The discussion of the role of teacher, peer- and self-assessment was also brought into discussion of each model because the Round 1 interviews indicated that who conducts assessment can have an impact on perceptions of how the task can be assessed. Leaving the discussion of peer- and self-assessment until the end of the task meant some models were dismissed on the assumption the assessment would be teacher-led. Drawing on results from Round 1, another section was added to the Round 2 pre-interview task, investigating potential criteria for assessing collaboration and tools which could be used to evidence collaboration for assessment purposes.

An important change from the pilot stage to the data collection stage occurred in the instructions for the task. In the pilot stage, participants were directed to consider the models in light of familiar assessment features: validity, transparency, reliability, fairness and practicality. However, it was felt that this was overly prescriptive and potentially limiting in terms of practitioner response. What was of interest in the study was seeing what the prime concerns of practitioners are when it comes to assessment and which features arose spontaneously.

2.6.2.4 *Models of collaborative assessment*

The models of assessment were based on Race (2001), which summarises seven approaches, and Nepal (2012) which discusses different means of applying a weighted contribution grade to a collaborative assignment. The seven approaches outlined by Race (2001, p. 18-20) are:

1. Group mark for the product
2. Cooperative assignment where students work on specific sections of the task and get a grade the quality of their section
3. Group mark for the product and students negotiated contribution grade

4. Group mark for the product and peer-assessed contribution grade
5. Group mark for the product with additional individual tasks
6. Group mark for the product and oral exam graded individually
7. Group mark for the product with individual written exam

Employing seven models in the task would have been too demanding for the participants. A cooperative approach to assessment was dismissed as an option as there is a fundamental difference between a collaborative and a cooperative assignment. That is not to say there is no place for cooperative assignments or assessment, but rather that that is not the area of interest of this dissertation, which is focussed on collaborative assignments. Options 6 and 7 outlined by Race were also omitted from this study as they could be seen as variations on Option 5, with the nature of the individual task changing but the general principle of students completing a main collaborative task and complementary individual tasks remaining. As such it was decided to use four models of assessment in the pre-interview task. The four models of assessment used in the pre-interview tasks, as per the nomenclature of the task used for Round 2 data collection, can be seen in Table 2.6.

Table 2.6 *Models of collaborative assessment used in the pre-interview task*

Model A	Group Grade	All students in the group receive the same grade, based on the quality of the group product
Model B	Group grade + complementary individual assignments	All students in the group receive the same grade, based on the quality of the group product + an individual grade for complementary assignments.
Model C	Group grade + individual collaboration grade	All students in the group receive the same grade, based on the quality of the group product + an individual collaboration grade, based on criteria
Model D	Group grade x weighted collaboration grade	Students receive the grade for the group project, adjusted by their <i>relative</i> contribution to the other group members.

For the purpose of the models, the group product of the collaborative assignment received a grade of 70%. This allowed for illustration of the results of the students in the essay group under each model, showing the potential for differentiation within the different models.

As the models, particularly Model D, may be difficult to grasp at first sight it was decided not to include a more complex model, combining individual complementary assignments and some form for collaboration grade. However, when teachers were suggesting their own model of assessment there was opportunity to combine different assessment lenses as they saw fit.

2.6.2.5 *EAP Practitioner Interview*

Interviews are the principal source of data collection in this study, as a means of “exploring people’s perceptions, meanings, definitions of situations and constructions of reality” (Punch & Oancea, 2014, p. 182). Intensive interviews are also a means to study a specific topic by engaging with participants who have gone through the experience (Charmaz, 2014). The interview transcripts are collated in Appendix 4.

The interview protocol for the original study was piloted on four participants in late summer 2019. Interviews were conducted remotely, via Zoom, and recorded. Although conceptualised as semi-structured interviews the submission of the protocol to an Ethics Board created doubt as to how unstructured an interview could be. As such, the interviews were conducted as standardised open-ended interviews (Cohen et al., 2011), with close adherence to the protocol, which was an attempt at objectivity and standardisation, echoing the positivist intent of the original study.

The move into Phase 2 of the study and the turn to a constructivist approach changed the nature of the interview. Rather than ‘researcher-researched’ there was recognition that the researcher is part of the study rather than an objective outsider (Roulston, 2010). This is something recognised by Glaser, who argued that researchers should be aware of their own biases and treat them as another variable (1978). Although the general coverage of the interview protocol did not change, the interview became more reflexive in nature. Rather than ‘researcher-researched’ the

relationship became more ‘practitioner-practitioner’, more dialogic and conversational (Charmaz & Belgrave, 2012). This helped with building rapport and creating a space where the participant could freely express ideas and opinions. Assessment of collaborative assignments is an area that has a lot of ambiguity and uncertainty about it and some of the uncertainties of the interviewer were shared with the participants through question stems like “One of the things I struggle with is ...” There was also an effort to create dialogue *between* participants through question stems like “Some of the other people I’ve been talking to have said xxx. How would you take that?”

The interview protocol became more of a general guide than a script and questions were brought forward or dropped in light of participant responses. There was also more in the way of spontaneous questioning, following leads of inquiry, with new questions generated during the discussion. This in itself is reflective of the GT approach, with an element of ‘live’ data analysis and research tool adaptation taking place during the interview in light of emerging observations (Charmaz & Belgrave, 2012), resulting in changes to the research tool, the questions being asked.

From the pilot interview to the second round of data collection the interview protocol expanded and evolved. The full protocols for each iteration of the interview protocol can be found in the Appendix 1 and the general coverage of the interviews is shown in Table 2.7 below. The think-aloud portion of the interview proved problematic in the pilot round of data collection, although potentially insightful. Participants had different levels of thinking time upon seeing a model and responses differed in length and depth, in some cases not necessarily articulating thought processes thoroughly enough to be of research use, an issue with think-aloud protocols recognised by Charters (2003). The task was likely too demanding, requiring participants to read an explanation of the model and interpret a graph. Tasks with a high cognitive load have been found to be unsuitable for think-aloud protocols as they can hinder verbalisation (Ericsson & Simon, 1980). As such, the think-aloud section of the interview was repurposed as the pre-interview task.

Table 2.7 *Summary of the general coverage of the three iterations of the interview protocol*

Pilot interview	Round 1 data collection	Round 2 data collection
• Experience of collaborative written assignments • Assessment of collaborative written assignments • Motivation for using collaborative assignments • Think-aloud protocol ○ models of collaborative assessment ○ teacher, peer- & self-assessment • Closing questions	• Role of EAP in Higher Education • Personal experience of collaboration • Collaboration in Higher Education • Experience of oral collaborative assignments • Experience of written collaborative assignments • Attitude to collaborative assignments • Cooperation vs collaboration • Models of assessment • Teacher, peer- & self-assessment • Own model of assessment for collaborative assignments • Closing questions	• Role of EAP in Higher Education • Personal experience of collaboration • Collaboration in Higher Education • Experience of oral collaborative assignments • Experience of written collaborative assignments • Attitude to collaborative assignments • Cooperation vs collaboration • Models of assessment • Teacher, peer- & self-assessment • Criteria for assessing collaboration • Evidence of assessing collaboration • Closing questions

As can be seen in Table 2.8 these changes led to noticeably longer interviews after the pilot stage, with interviews lasting, on average, just over an hour.

Table 2.8 Interview length

Pilot round data collection		Round 1 data collection		Round 2 data collection	
Pseudonym	Interview length (mins)	Pseudonym	Interview length (mins)	Pseudonym	Interview length (mins)
Barry	44:06	Ava	57:39	Fay	57:36
Andrea	40:50	Fernand	58:56	Gonzalo	63:46
Tanya	36:20	Humberto	70:58	Kyle	62:10
Dorian	33:47	Karen	89:45	Laura	68:05
		Jerry	20:01*		
		Olga	60:20		
		Lorenzo	57:58		
		Wendy	59:27		
		Bertha	75:21		
		Van	56:13		
		Cristobal	85:37		
		Dolly	87:43		
Average length	38.46	Average length	69:05	Average length	62:54

**original interview failed to record – this is the cumulative length of WhatsApp recorded responses and is excluded from the average length of interview shown.*

2.7 Participants

The focus of this study is the experience, attitude and beliefs of EAP practitioners. It is not context bound beyond practitioners teaching EAP in HE institutions. The profiles of the 20 participants who participated in the interview exploring assessment of collaborative assignments are outlined in the next section, drawing on the survey data. As will be clear, the participants come from diverse background, are well qualified, have a range of experience and are working in a variety of EAP contexts (see Table 2.9).

Table 2.9 Participant biodata from the survey

	Pseudonym	Nationality	Working in	Highest qualification	EAP experience (in years)
1	Andrea	Greek	England	MA	11-15
2	Ava	Chinese	Macau	MA	1-2
3	Barry	British	England	Diploma	3-5
3	Bertha	British	England	MA	6-10
4	Cristobal	British	UAE	MA	Over 15
5	Dorian	British	Finland	PhD	11-15
6	Dolly	American	Macau	PhD	11-15
8	Fay	Central American	Israel	MA	Over 15
9	Fernand	British	China	MA	11-15
10	Gonzalo	American	America	MA	11-15
11	Humberto	British	Finland	MA	3-5
12	Jerry	British	Middle East	MA	11-15
13	Karen	New Zealander	Finland	MA	6-10
14	Kyle	British	England	MA	Over 15
15	Laura	Turkish	Turkey	MA	11-15
16	Lorenzo	American	Finland	PhD	3-5
17	Olga	Taiwanese	Finland	PhD	11-15
18	Tanya	American	Finland	MA	6-10
19	Van	British	Scotland	PhD	11-15
20	Wendy	British	England	MA	6-10

The following sections provide a profile of each participant, covering their biodata, attitude to collaboration and accompanying rationale, as provided in the survey. Relevant data that was not captured in early iterations of the survey (e.g. experience working in different EAP contexts) but which

came to light in the subsequent interview is also included in this section. Participants who participated in the pilot interviews are clearly identified.

2.7.1 Andrea (pilot round)

Andrea is a Greek national, working in a UK based university, delivering courses from pre-session through undergraduate and up to PhD level. She has been working there for 6-10 years in the role of EAP lecturer and has a total of 11-15 years teaching EAP and previous experience teaching ELT and holds a Master's degree in Education. She has limited experience of collaboration in her own studies and mixed feelings towards collaboration. She would describe collaboration as "insightful, educational, creative, developmental" and sees collaboration as positive for students. While she is positive about collaboration, she feels assessment is an area that has a strong negative impact on her attitude toward it and she also finds students hold a more negative attitude toward it. For both teachers and students, she sees time management as a major challenge in a collaborative assignment. At the end of the survey when asked if there was anything she wanted to add, the focus of her two questions were on the assessment of an individual within a collaboration.

2.7.2 Ava

Ava is a Chinese national, working in a HE institution in Macau, teaching an academic English course. She has obtained an MA in the field of ELT and has been teaching EAP for 1-2 years, with 1-2 years teaching ELT preceding that. She feels positively toward collaboration and has received occasional training in collaborative learning along with some experience of collaborating as a student. However, her attitude to collaboration has shifted toward more neutral due to issues experienced with free-riders and mixed level groups, which has seen her introduced more scaffolding into her teaching of collaborative assignments. While she values the learning gains in collaborative assignments, the issue of grading in light of some students contributing less than others is a concern.

2.7.3 Barry (pilot round)

Barry is a British national, working in a UK based institution in a management position. He holds a diploma level qualification and has 3-5 years of EAP experience, preceded by 11-15 years of ELT. He had no experience of collaboration as a student and no training toward it in his studies but feels positively toward collaboration as it equips students with necessary real-world experience of working with others. However, he is concerned about the added administration of group assignments, the potential technologies involved and coordinating students.

2.7.4 Bertha

Bertha is a UK national working in a UK institution. She has been working in EAP for 6-10 years (preceded by 6-10 years of teaching ELT), with previous experience working in a China based HE institution and she holds an MA in the field of ELT. She teaches across EAP programmes, from pre-sessional to PhD. She would describe collaboration as variable, meaning the experience depends on who you are working with, and had some personal experience of collaboration as a student. She has a positive attitude and positive experience of collaboration and values the opportunities for peer learning and peer feedback, alongside the transferable real-world skills developed. However, it is difficult for a teacher to determine individual students' abilities and to monitor a collaboration and for students they may experience difficulties of working with others who do not contribute and may not see such assessment as fair.

2.7.5 Cristobal

Cristobal is a UK national working in the UAE teaching undergraduate students, and has over 15 years' experience, holding an MA in the field of ELT. He is very positive about collaboration, viewing it as stimulating and challenging, and had some experience of collaboration as a student himself. For the teacher he sees the benefit that they are able to co-construct with students, developing better

relationships and deeper understanding of individual students. Students in turn benefit from learning from each other, although there is the negative issue of free-riders.

2.7.6 Dolly

Dolly is an American national, working in a HE institution in Macau teaching undergraduate students. She had some experience of collaboration as a student, but never formally studied collaborative learning as part of her studies up to and including PhD. She feels positively about collaborative assignments, but this has been negatively impacted by students' pessimism toward groupwork assignments and she has also become more defensive in her teaching in reaction to free-riders. While she sees real world benefits to students collaborating with each other, the issue of student complaints is a problem.

2.7.7 Dorian (pilot round)

Dorian is a UK national working in a university in Finland in a management position with over 15 years' experience and holds a PhD in an ELT related field. He has no experience of collaboration in his own education and collaborative learning was not included in his studies. He views collaboration very positively, not identifying any drawbacks for teachers or students, instead mentioning the positives of student-centeredness and better-quality work.

2.7.8 Fay

Fay is Central American and is working in an Israeli HE institution, holding a position of responsibility as programme head. She has over 15 years' experience teaching EAP and has worked in five different HE institutions. She has a positive attitude to collaboration, describing it as enriching. She has some experience of collaboration as a student but collaborative learning was not covered as part of her studies. She values collaboration with students as it prepares them for the workplace and

helps foster teamwork, language and soft skills. From the teacher perspective she values the peer learning aspect and the practicality of having fewer papers to assess, although there is clear appreciation that collaborative assignments need carefully considered scaffolding and preparation. For students she noted the benefits in terms of learning gains and increased motivation, but recognised group dynamics could sometimes be difficult to navigate.

2.7.9 Fernand

Fernand is a UK national working at a university in China teaching undergraduate students, with 11-15 years' experience of teaching EAP. He holds an MA but collaborative learning was not covered in any of his studies. He does, however, have some personal experience of collaborating as a student from his MA studies. He has a negative attitude to collaborative assignments, in part because the collaborative writing assignment he was involved with took place in the first semester, which he felt was too early in the students' academic careers. For teachers, he dismisses the notion it could reduce marking load and feels it added complexity to the assessment. For students there was a potential benefit in terms of developing teamworking skills, although he questioned how educational and national culture could curtail this. He also thought students could feel the assessment was unfair and could add to their stress, particularly if they were in a group with a free-rider.

2.7.10 Gonzalo

Gonzalo is an American working in a university in America, although he has been working in EAP for 11-15 years, in five different HE institutions, including one based in China and one in Central Asia. His current role is Instructional Designer, working across the university to help with course and programme development. He holds an MA but rarely experienced or studied collaborative learning during his education. He has a neutral attitude toward collaboration but a positive attitude toward collaborative written assignments, in part because students can be given more challenging work. In terms of benefits, collaborative assignments can be more complex and students can support each

other and engage in peer learning. The drawbacks were more prominent, with difficulties of identifying what an individual has done, unfair workload, resentments between students, lower marks impacting group dynamics and problems in collaborating.

2.7.11 Humberto

Humberto is a British national working in a university in Finland. He has been teaching EAP for 3-5 years delivering courses at undergraduate and MA level, with over 15 years teaching experience prior to that. He holds an MA but rarely studied collaborative learning and had no experience of collaboration himself as a student. He views collaboration positively but has recognised the need for tasks to genuinely foster collaboration and to ensure they are not dominated by one or two of the stronger students. He appreciates the learner-centred nature of collaboration and the opportunities for peer learning. However, this is balanced against the need for the teacher to ensure students are engaged with the collaboration and helping students see the benefits of collaborating and also voiced concern that students could feel short changed as they were not getting teacher attention.

2.7.12 Jerry

Jerry is a British national, working in a management position in a university in Oman. He has 11-15 years' experience teaching EAP and has worked in a variety of contexts, including China and the UK. His feelings about collaboration are neutral. He had little experience of it as a student and it was not covered in any of his studies, up to and including his MA. Despite this he had a positive attitude toward them but his experience in management has brought to the fore some of the issues and challenges, with collaborative assignments sometimes poorly conceived, which can lead to poor feedback from students. Students can also produce poor quality work as they lack the autonomy to navigate the collaboration successfully and the performance of free-riders or weaker students could be a source of frustration for stronger students, particularly in the case where all students receive the same grade. Nonetheless, he sees benefits in terms of fostering students' soft skills, the development

of communication skills, and the opportunities for peer learning. At the end of the survey he commented on wanting to know about approaches to assessing collaborative assignments.

2.7.13 Karen

Karen is a New Zealander working in a university in Finland. She has 6-10 years' experience working in EAP and more than 15 years' experience teaching English previous to that. She holds an MA and has studied collaborative learning as part of her studies. She holds mixed feelings towards collaboration and would describe it as demanding and transformative. She appreciates the reduced teacher work load in collaborative assignments, although the time management can be an issue and students wanting individual feedback can be problematic. For students she sees the twin benefits of peer learning and a reduced study load, although a problem could be equal access to online tools. She is interested in the kind of knowledge which can be created during student collaboration.

2.7.14 Kyle

Kyle is a British national working as the Director of a language centre in a UK university. He has over 15 years' experience teaching EAP and has worked in more than five different HE institutions, both in the UK and overseas. He learnt to collaborate through trial and error, with little experience of collaboration as a student but a great deal more in working life. He holds a positive attitude to collaboration and is interested in what students learn from the process of collaborating rather than the product of the collaboration. In terms of benefits, for teachers he appreciates less marking and more efficient feedback processes as the same feedback is not being given to individually to multiple students. He also feels students can learn to work together and share responsibilities. On the negative side, there is considerable organisation for teachers to manage, and potential conflict resolution. In relation to students, both comments were connected to fairness, with one pointing out unequal distribution of work and the other regarding marks not being equitable with student input. When asked what would help him facilitate collaborative assignments more effectively, he commented on

assessment frameworks, which focus on the process as well as the product and which would balance individual input and output against that of the group. His final comment suggested consideration needed to be given to student needs in the context of a collaborative assignment.

2.7.15 Laura

Laura is a Turkish national, working in a university in Turkey. She has 11-15 years' experience and has worked in four different HE institutions. She holds an MA and while she had some training in collaborative learning she mainly learnt through observation and trial and error. She holds a neutral attitude to collaboration, considering it proactive, challenging and requiring negotiation. Her approach to facilitating collaboration has changed as she gained more experience and she notes she is good at spotting potential problems, based on her previous experience. For a teacher, student collaboration helping students develop real-life skills and requires the students to be more active and engage in more interaction. However, managing groups does take skill and teachers must ensure no student is disadvantaged and must also deal with any group conflict. Students also benefit in terms of developing valuable real-life skills and engaging in peer feedback, but it can be problematic working with people they have not met before, particularly in an online learning environment and an unresponsive group member can negatively impact both group dynamics and the final work. She feels that having students with prior experience of collaboration helps with facilitation.

2.7.16 Lorenzo

Lorenzo is an American working in Finland. He holds a PhD and has 3-5 years' experience teaching EAP, with 6-10 years' experience teaching English prior to that. He considers collaboration to be skill-enhancing, although he had limited exposure to collaborative learning as a student or as part of his studies. His attitude has also shifted from positive to more neutral as students were less enthusiastic about collaboration than he expected. This seems to be due to student concerns about unequal workload and contribution in group work, leading to some of them preferring to work alone.

He appreciates having fewer papers to read in collaborative assignments but considers having to deal with conflict resolution a drawback. He feels students benefit in terms of career preparation but groupwork can have a negative impact on efficacy and speed.

2.7.17 Olga

Olga is a Taiwanese national teaching in a university in Finland. She holds a PhD and has 11-15 years' experience teaching EAP, including time working in the UK context. She describes collaboration as inspiring, albeit sometimes stressful. She held a very positive attitude to collaboration, which has shifted to a more neutral stance. A possible factor in this is her recognition of how different the educational and assessment cultures in the UK (where she felt more positively toward collaboration) and Finland, where academic freedom means it can be more difficult to agree on assessment practices and criteria with colleagues. This resulted in her feeling a lack of support and, consequently, student collaboration became less rewarding. In terms of benefits for teachers, she appreciates having different types of assessment, although collaborative assessments could lead to a higher workload and stress. For students, she sees collaboration as an opportunity for students to learn by doing, although some students do not like working with other people. She expressed curiosity about the results of this research, particularly as it pertains to the Finnish context, which she is finding challenging.

2.7.18 Tanya (pilot round)

Tanya is an American working in a university in Finland. She has 6-10 years' experience teaching EAP and holds an MA. She had some experience of collaboration as a student, but has never formally studied collaborative learning. She holds mixed feelings toward collaboration and sees the reduced marking load as a benefit, although recognises that the marking becomes more complicated in collaborative assignments. She sees collaboration as being a valuable life skill for students and

suggests a collaborative assignment could reduce or focus student workload. However, she notes the potential for an uneven workload as a potential drawback for students.

2.7.19 Van

Van is a British national, working for a Scottish university, with extensive experience working in China. He holds a PhD and has 11-15 years' experience teaching EAP. In his survey he had had no experience of collaboration as a student, although in the interview he did refer to some collaborative experiences during his MA. Collaborative learning did not feature as part of his studies at any level, and he holds a neutral attitude to collaborative assignments. He was initially more sceptical but saw the opportunity for peer scaffolding as a positive. He sees this scaffolding as a potential benefit for teachers, alongside the reduced marking load. However, freeloaders present problems in the assessment of a collaborative assignment as they may get a grade that overestimates their contribution and ability. Freeloaders also pose a problem for students, as others may end up doing more work. However, for students the benefits lie in peer learning, developing collaborative skills, and building social relationships.

2.7.20 Wendy

Wendy is a British national, working in a university in the UK. She holds an MA and has 6-10 years' experience teaching EAP, with a part of that being at a branch campus of a UK university operating in China. She rarely had experience with collaboration as a student and collaborative learning did not feature as part of her studies. She holds mixed feelings towards collaboration but would describe it as humorous and educational. She has developed her approach to collaborative assignments through experience and has seen what students are able to get out of the process. She sees collaborative assignments having benefits for teachers in terms of a lower marking load and more student-centred classrooms. She also feels students can benefit in terms of learning from other students but has a concern that stronger students may feel held back and also recognises the negative

impact a difficult group member could have, both in terms of interpersonal relationships and on final grades. She highlights that teachers have a duty of care to ensure students are well prepared to work together and to intervene in the case of group dysfunction. In the final open comment, she questioned the fairness of grading in collaborative assignments, where some students could feel their grade does not reflect their work, or that they would have been able to get a better grade individually. She saw a place for collaboration as an unassessed activity so it becomes a lower stakes activity, but students still get the benefits of collaborating.

2.7.21 Anna, Claudette and Elsa

These three participants did not complete the initial survey as they were final round data collection participants who evaluated the practical output of this dissertation. Anna works in a university in Norway, while Claudette and Elsa work in UK universities, with Elsa working as the manager of the English language centre in her institution.

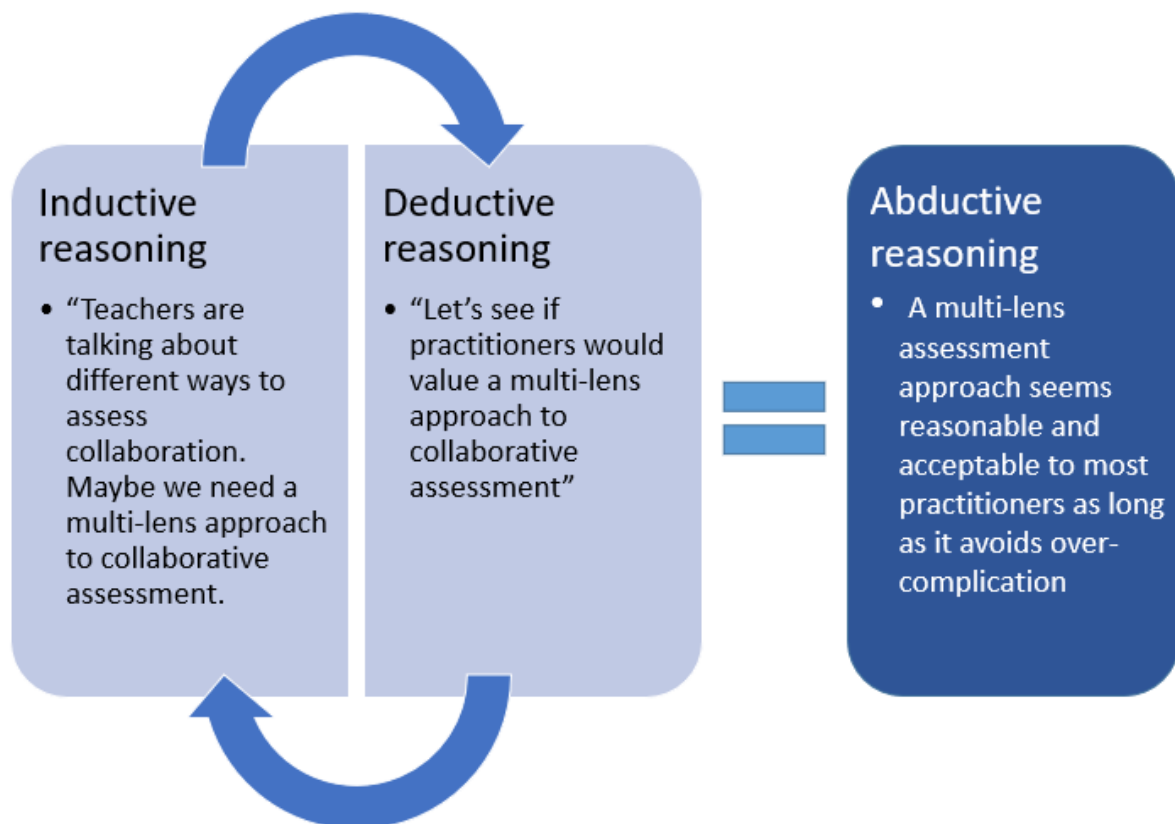
2.8 Data analysis

Data collection and data analysis are conjoined and inviolate in GT, with data acquisition and theorising having a “symbiotic relationship” (Boychuk & Morgan, 2004, p. 607). Analysis occurs in the earliest stages of data collection, informing future data collection, guided by theoretical sampling. GT could be thought of as inductive, generating generalised conclusions from specific data, but goes beyond this. It also involves creativity, conjecture, and deductive checking of that conjecture (Charmaz, 2014). It’s a process of looking at the data, identifying nascent concepts and reaching further into the data and gathering new data to develop those concepts to the point of theoretical saturation or sufficiency.

2.8.1 Types of reasoning

This process of data analysis required different types of reasoning, supporting and feeding into each other. An example of this can be illustrated by one of the early concepts to emerge from the interviews, that of ‘multi-lens assessment’ of a collaborative assignment. This was induced from the different avenues and tools practitioners discussed using to assess collaboration, generating the idea of ‘multi-lens assessment’ of collaboration, using a variety of lenses (assessment tools) to bring the collaboration into focus. This concept was then folded into the pre-interview task, with a section about evidence of assessing collaboration to see how later participants would respond to the idea. This in turn could lead to the abductive conclusion that, with caveats, a multi-lens approach would find favour with practitioners. This demonstrates the journey from inductive through deductive to abductive reasoning, as illustrated below (Figure 2-8).

Figure 2-8 *Illustration of the cycle of reasoning, from inductive to deductive to abductive*



2.8.2 Memoing

Before outlining the process of data analysis, it is essential to emphasise a core tool of data analysis in this study, that of memoing, which is the key arena where thinking takes place. Memoing is a central element of GT and has been since its conception and across different developments (Charmaz, 2014; Glaser & Strauss, 1967; Hadley, 2017). It is more than an intermediate stage of writing, linking data collection and thesis writing. Memoing is the development of ideas, the discovery of concepts, and the crucial step in analysis. Some memos may have ideas that lead nowhere, others concepts that become the centrepiece of the resultant theory. As Charmaz notes, memos can be “partial, tentative and exploratory” (2008, p. 166) and then these can be further developed and interrogated in later memos. Glaser actively discouraged talking about ideas with other people and advocated the researcher stopped whatever else they were doing and memo whenever an idea arose (Glaser, 1978).

During data collection a memo was opened for each participant which contained a review of their survey responses and areas of potential interest during the interview. The memo was expanded after review of the pre-interview task and post interview with key points and thoughts for further research tool refinement or development. Data was not just collected, reviewed, and put aside. Interview transcripts were returned to, prior to coding, to see how effective the interview was, which questions elicited particularly useful and illuminating responses. Participant memos were developed in NVivo12 (Lumivero, 2024a), started after initial survey review and amended after pre-interview task analysis and then again post interview analysis. A global project memo was started in NVivo12 (Lumivero, 2024a) and then moved to Office365 for ease of access as the NVivo12 (Lumivero, 2024a) project was stored locally on a single laptop and could not be accessed remotely. The global project memo had to be accessible whenever inspiration struck. Memos also developed through the process of coding the data which led into concept memos, outlining early concepts, which were gradually refined. Some memos consequently served as the foundation of some of the text in this dissertation.

2.8.2.1 Illustration of memoing

Figure 2-9 (below) is an extract from the project memo, made while doing line-by-line coding on Barry's interview, at the start of the coding process.

Figure 2-9 Extract from near the beginning of the project memo

3rd June 2020

Started coding Barry's interview with a line by line coding. It's good to have to focus on each line and it does highlight some things you could miss anyway. Where he says "That's a really interesting question" makes me think an area to code is the awareness raising of the interview and the discussion of collaboration. Is collaboration something that people just don't think about so any thought about it is beneficial?

Line by line is also good for teasing out small things - you get things like factors impacting student attitude to collaboration and you see how quickly something like one student dominating a collaboration is the default way Barry sees collaboration. The line by line coding is going to generate A LOT of codes and they will need to be organised. Broad categories could be

- > Describing the teaching context & student body*
- > Describing the collaborative assignment (how it works, how it came to be that way)*
- > Teacher attitude to collaboration*
- > Student attitude to collaboration*
- > How students collaborate*
- > How to assess collaboration*

5th June 2020

Finished first coding of Barry's interview transcript. Generated 62 codes and started the organisation of them into groups as coding. There needs to be a clear separation in codes between participant reflection on a model and the model they themselves propose. It might also be better not to force a model for the final proposal e.g. not ask for % for specific assessment elements.

Would it work to send the models in advance, discuss them in the interview and then get email of model proposal after reflection? Would that be better than a model made up on the spot?

In the model review stage it might be good not to give criteria for review e.g. validity to see what spontaneously comes out in the discussion. What are the factors of most concern/interest to EAP practitioners?

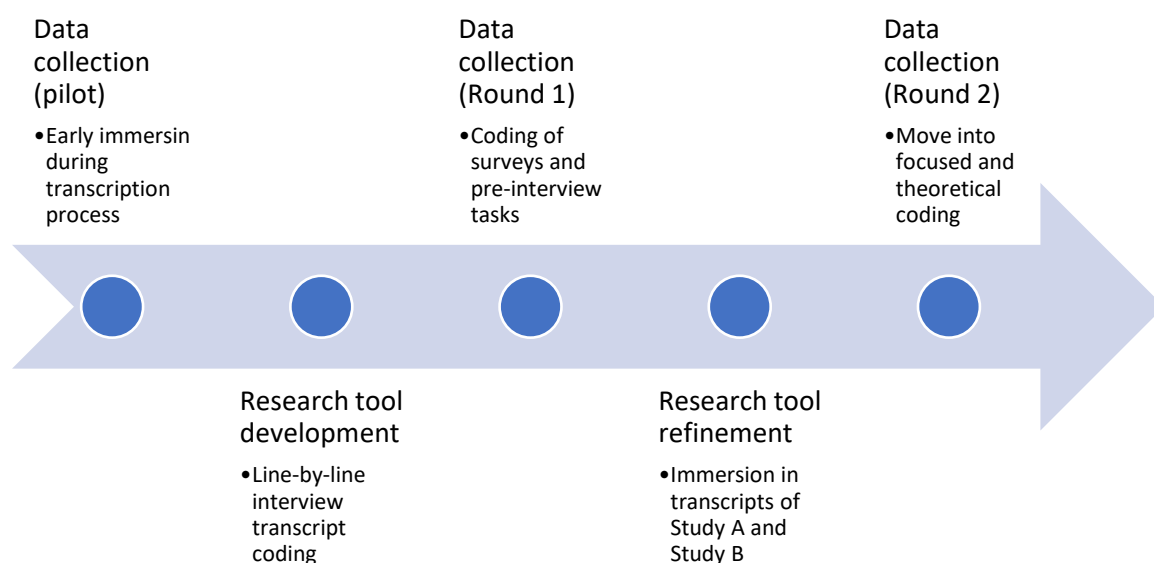
From this early memo the iterative nature of GT is demonstrated, containing ideas for revisions to the interview protocol to glean better data and the genesis of the pre-interview task. There is also early recognition of the value participants placed on discussing the assessment of collaborative assignments, which is a further motivation for developing a practical tool that can help practitioners

in the development of an assessment scheme for collaborative assignments by getting them to ask good questions.

2.8.3 Open exploration

The early stage of data analysis occurs during what Hadley (2017) termed open exploration (see Figure 1-10 in 1.5.2.4) and an iterative combination of immersion and coding was used. Immersion involves reading and rereading data, being submerged in it, reflecting on it and noticing ideas and concepts that crystalize (Borkan, 1999; Green et al., 2007). This holds a commonality with GT where “crystallising” could be seen as synonymous with “emerging”. As can be seen in Figure 2-10, early immersion started in the transcription stage, which prompted memoing of important points raised in the participant and project memos. This was followed by line-by-line coding of the pilot interview transcripts and further memoing, which informed further development of the research tools. After the next round of data collection, coding occurred on the survey and pre-interview task data while close reading and re-reading of the interview transcripts took place. There was an ongoing interplay between data collection and data analysis.

Figure 2-10 *Iterative immersion and coding analysis*



GT is iterative and through this study research tools continued to evolve in light of subsequent rounds of data collection, analysis and reflection. The choice of 'evolve' is deliberate in favour of 'refined' or 'amended'. Refinement and amendment suggest making the tools better at gathering the same data but evolution is meant to evoke the idea of tools developing to do old tasks better and also to take on new tasks, to gather new types of data. As the study moved closer to developing a theory the focus of the tools changed too, gathering data to enrich, consolidate or cast new light on emerging concepts.

2.8.3.1 Approach to coding

The Coding Manual for Qualitative Researchers (Third Edition) identifies coding as an interpretive act "in service of thinking," (Saldaña, 2016, p. 80). Coding generally, and certainly within a GT approach, is cyclical, encouraging a reengaging with the data and the approach to coding in this study was adopted so that data could be looked at in different lights.

The pilot interviews provided the data set used for generating the initial codebook. A line-by-line coding approach was adopted, which would lay the foundation for future coding cycles. While Glaser warns against generating too many codes and getting lost in the data (Glaser, 1978), as does Saldaña (2016), Charmaz advocates for a line-by-line approach for initial coding as a heuristic means for the researcher to interact with the data (Charmaz, 2014). The aim of using line-by-line coding was to force close reading of the interview data and to avoid attention being directed by researcher bias. It also provided a contrast from the holistic reading and rereading of immersion, bringing other aspects of the data to the fore and took advantage of NVivo's (Lumivero, 2024a) ability to bring together coded text from varied data sources together. For the line-by-line coding, which was closer to phrase-by-phrase coding, several coding methods were utilised, shown below in Table 2.10 according to Saldaña's nomenclature (2016). Simultaneous coding was allowed, with a single line of interview transcript potentially carrying multiple codes.

Table 2.10 *Initial coding methods of pilot data in Saldaña's nomenclature (2016)*

Elemental methods	Structural (p. 98)	To gather data linked to the research question together
	Descriptive (p. 102)	To gather data connected to key topics together (e.g. fairness)
	Process (p. 111)	Indicating the action of what is happening, a coding approach favoured by Grounded Theorists (Charmaz, 2008; Hadley, 2017)
Affective methods	Values (p. 131)	Encapsulating values, beliefs and attitudes

These methods were chosen to remain open to the data and to avoid premature theorisation, which could have been driven by bias or the literature rather than the data itself.

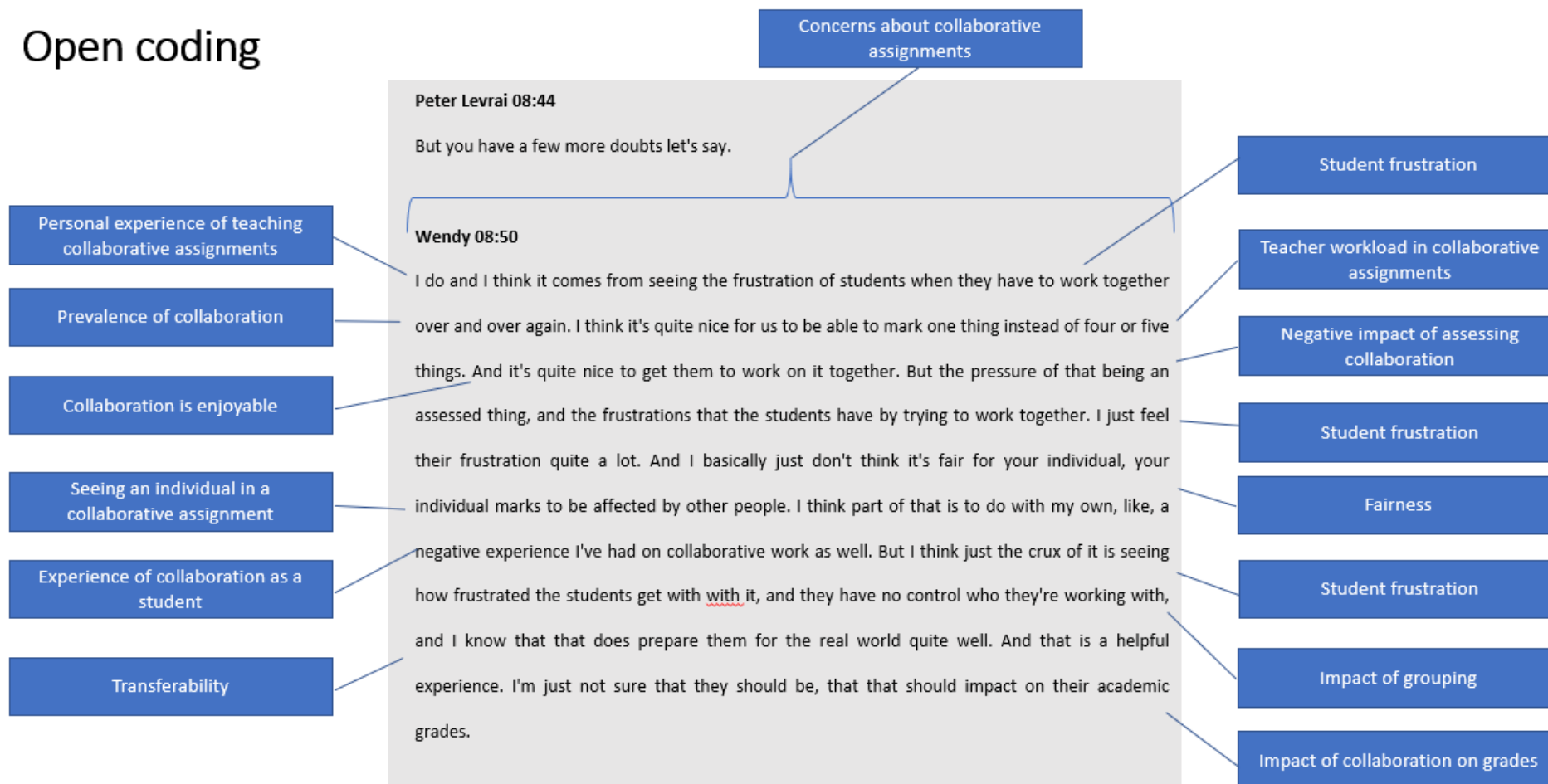
As codes were developed, they were categorised under parent codes and this initial line-by-line coding resulted in 'Open Coding (pilot)' (Appendix 6, p. 3-18), which contained 18 parent codes and over 100 child codes (some of which contained subcodes of their own). An example of open coding is illustrated below (see Figure 2-11), with an extract taken from early in Wendy's interview (Appendix 4, p. 289), during discussion of her attitude to collaborative assignments and some of her doubts surrounding them. The same extract will be used to illustrate focused and concept coding later in this chapter. Figure 2-11 illustrates the density of the initial open coding, with the same text coded multiple times on a phrase and/or sentence level. It also illustrates the richness of the interview data, where a single response encapsulates so many different elements.

Something important to note about myself is that I tend to over-focus on technical elements and, while coding, started exploring different options and functionalities of NVivo12 (Lumivero, 2024a). While this has value in being able to maximise use of the tool it was important to remember that it was just a tool that cannot do the thinking for you (Charmaz & Belgrave, 2012) and that "the true essence of qualitative analysis is based in investigator-insight" (Morse, 2010, p. 233). The real work of the data analysis is the thinking. This prompted a return to immersion when it came to the transcripts, which aided in the thinking process, with multiple re-readings of the research data building

on the close reading of the data forced by the initial line-by-line coding. The second round of immersion was also useful for taking a holistic view of the participant's responses, without always getting caught up in the technical process of coding. However, coding of the survey and pre-interview task did take place, resulting in 'Open Coding (Round 1)' (Appendix 6, p. 19-37), which was an extension of 'Open Coding (pilot)' and contains 20 parent codes and 130 child codes (plus subcodes).

This process served to increase researcher theoretical sensitivity and bring to light nascent concepts, which proliferated during this process. As Borkan (1999) and Green et al. (2007) note, immersion is a lengthy process, it takes time to really sink into data. As such, the ongoing COVID-19 pandemic had a benefit of slowing down the data collection process, affording more time for immersion and crystallisation of emerging concepts. These nascent concepts would form the basis of the next cycle of coding when the study moved into focused investigation (see Figure 1-10 in 1.5.2.4 above).

Figure 2-11 Illustration of open coding



2.8.4 Focussed exploration

The open coding process and immersion had provided considerable insight into the data, with patterns and concept emerging. After Round 2 data collection, but before focused coding of the data collected, some nascent concepts had emerged, which would form the basis of the focused coding and enabled the adoption of hypothesis coding (Saldaña, 2016, p. 171) to further explore the findings from the early data analysis. The concepts are briefly introduced in Table 2.11 below across three main domains, those of assessing collaborative assignments, teachers, and assessment teacher tribes. With the emergence of these concepts it was possible to focus the open codes, which were grouped and categorised, with an example illustrated in Figure 2-12. This process helped both reduce the number of codes but more importantly drew attention to areas of specific importance that multiple open codes spoke to. In the example in Figure 2-12 there is the genesis of what would become the theoretical code of ‘practical orientation’.

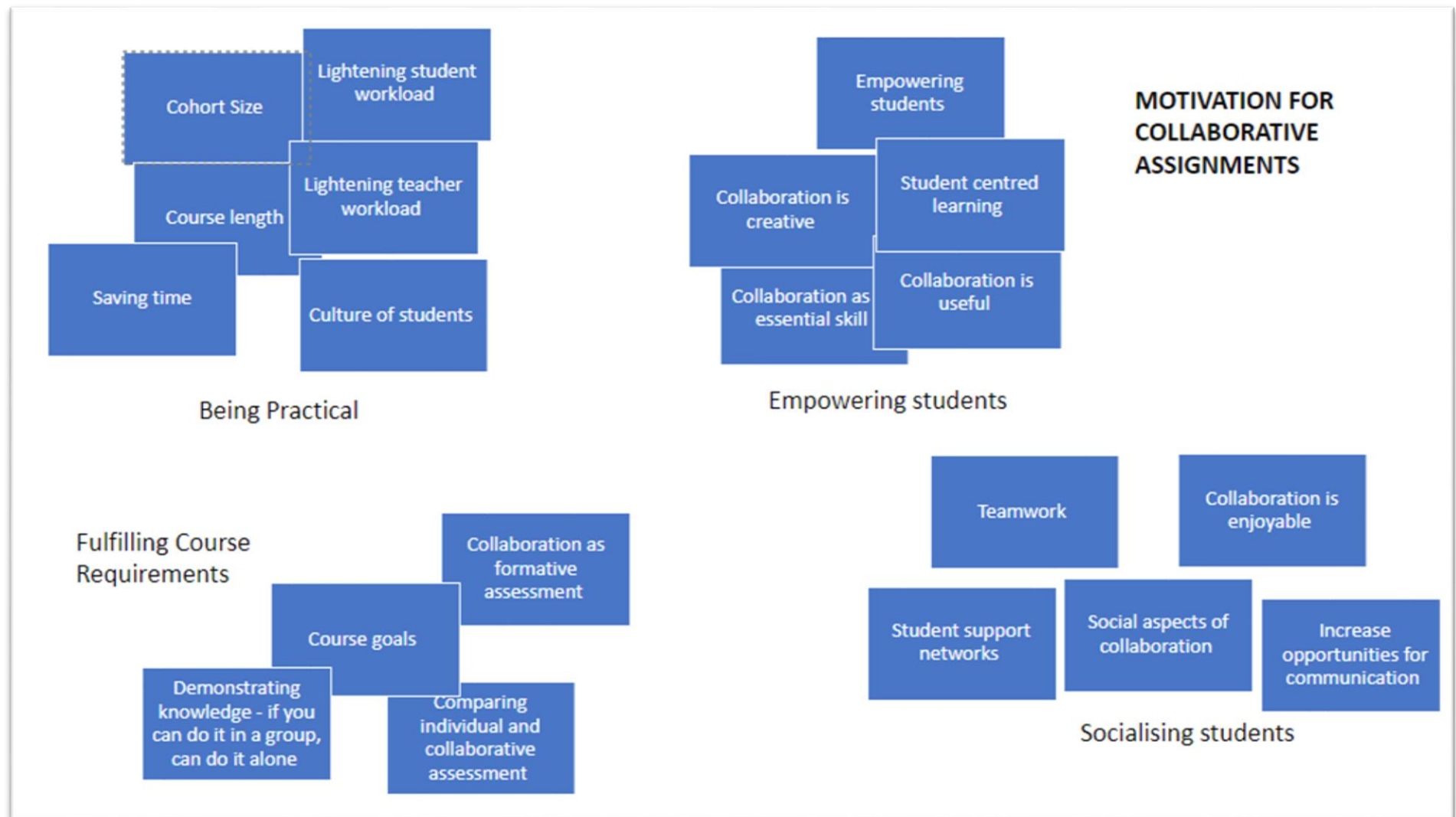
Table 2.11 *Nascent concepts from early data analysis*

Domain	Concept	Brief description
Assessing collaborative assignments	Multi-lens assessment	Collaboration cannot be viewed through a single assessment tools. Numerous tools may be needed to bring it into focus (see Figure 2-14)
	1+1+1+1+1=1 collaborative assessment	Many assessment tools leading to one assignment grade
	Assessment cost / benefit	Assessment takes time, energy and effort – a preliminary decision in assessment design is determining if the planned assessment tools are worth the cost, particularly in the assessment of a collaborative assignment where multiple assessment tools may be in play
	Assessment fatigue	For teacher or students if too many assessment components to an assignment
	Assessment overload / saturation	Too many assessment tools for on assignment

	Audit culture	Used in relation to research (Spooner, 2017) but also potentially relevant to assessment in terms of an assessment needing to pass an exam board or external review so novel forms of assessment are avoided
	Feedback transferability	Feedback to a group on a collaborative task transferring to an group member's individual work or assignments
	Product assessment of collaboration	The quality of a collaboration is apparent in the quality of the product
	Product glare	The quality of the product blinds the assessor to the collaboration
	The compromise of collaborative assessment	Assessing collaboration will always be a compromise, and be compromised. It is a complex, intangible construct which cannot be fully or definitively assessed
	Sweat assessment	Teachers value visible effort in a collaboration
	Task dilution	Students are given a main task to do. Each element of collaboration assessment takes time/energy/effort away from the main task
	Surface level differentiation	In presentations differentiation comes from surface level elements – individual presentation skills and language competence. This is not available in a written task
	Unreliable narrator reflection	Teachers question whether student reflection can/should be assessed and if it is an accurate portrayal of what actually happened
	Goodhart's Law in assessment	If collaboration is codified in criteria students may behave toward that criteria without learning anything about collaboration or collaborating naturally
	Direct peer assessment	Summative assessment carried out directly by peers
	Teacher filtered peer-assessment	The teacher provides the grade, informed by peer-assessment
	Direct self-assessment	Summative assessment carried out directly by the self
	Teacher filtered self-assessment	The teacher provides the grade, informed by self-assessment
Teachers	The reluctant overlord	Teachers favour teacher assessment rather than peer- or self-assessment but feel guilty about it

	Reward orientation	Teachers want opportunities to reward students, not punish them so lenses for differentiation as there so strong students can be rewarded
	Native collaborator	Someone who naturally and instinctively favours collaboration and enjoys collaboration
	Experienced novice	An experienced teacher with no experience of a novel assignment type
	Superficial sameness	Teachers who may have a similarity (e.g. preferred way of assessing a collaborative assignment) but there could be deep ideological differences between them as to why they have that difference.
	Engagement teacher tribes	Categories of teacher who have superficial sameness in terms of levels of engagement in collaborative tasks
	Assessment teacher tribes	Categories of teachers who have superficial sameness in approaches for assessing collaborative assignments, albeit their motivations for doing so may differ.
Assessment teacher tribes	The real-worlder	It is real world to only assess the product of a collaboration, not the collaboration itself. Potentially two types of real-worlder. Full, who believes in product assessment. Pragmatic, who would like to assess collaboration but does not know how
	The would-if-I-coulder	Knows there is a value to assessing collaboration but doubts the process of doing it effectively/accurately/fairly
	The all-seeing-eyer	A teacher confident they can assess a collaboration and an individual in a collaboration effectively (this is a theoretical construct, not met in the data so far)
	The Never-everer	A teacher who believes collaboration cannot/should not be assessed
	The Knowing-unknowner	Teachers deeply engaged in the student collaboration but still do not feel they have enough information or access to assess collaboration
	The I-in-teamer	Teacher who favours individual complementary assignments as an opportunity for differentiation through individual tasks.

Figure 2-12 Grouping & categorising open codes

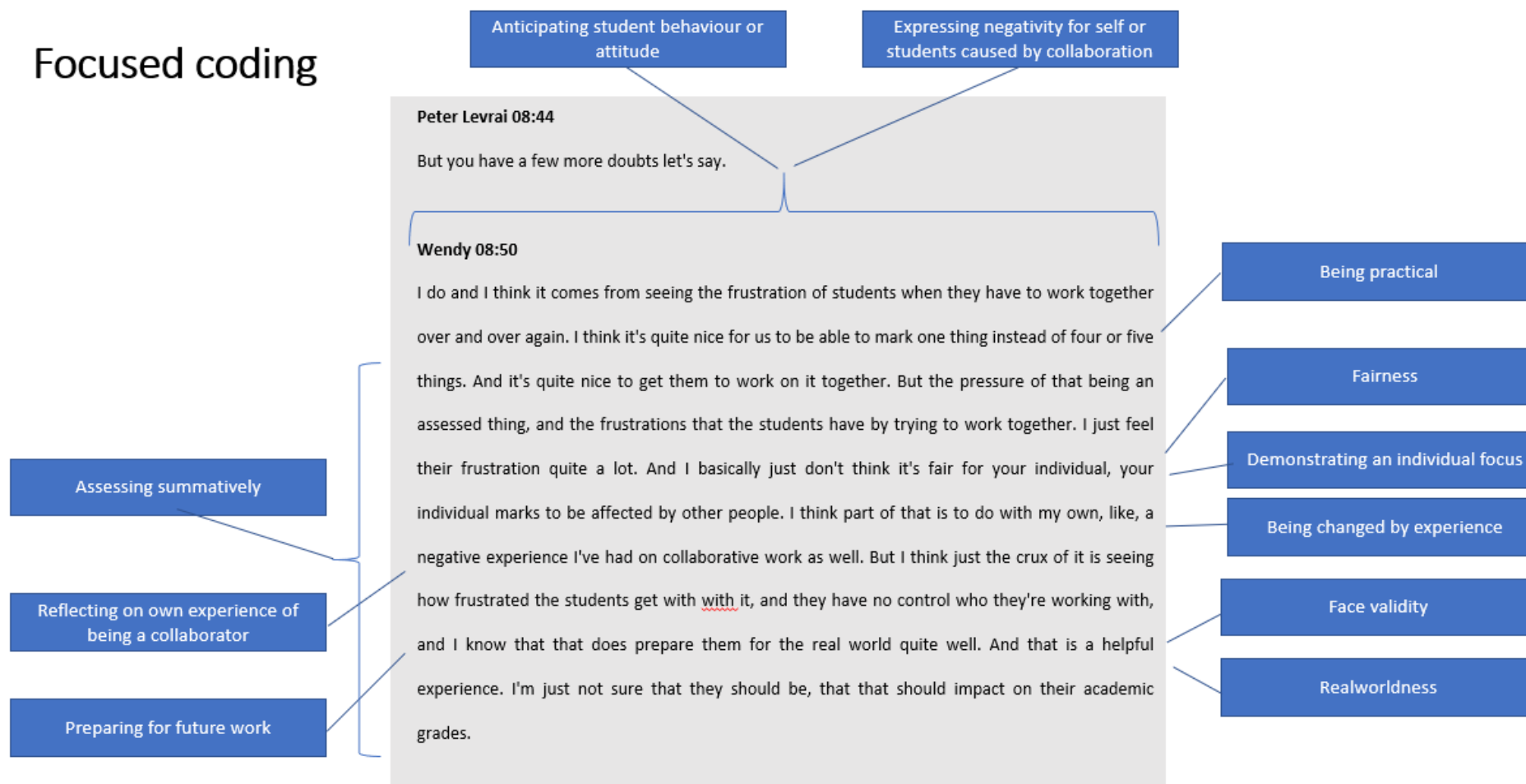


The focused coding codebook (Appendix 4, p.38-36) is still substantial, with 120 codes in total. This is considerably more than intended but in the process of conducting focused coding, some open codes which had been collapsed together needed to be parsed again for the sake of clarity. Figure 2-13 below illustrates focused coding in practice. This example shows how the focused coding began to reflect the concerns brought out through open coding, with practicality emerging as an important code, as was the concept of 'Realworldness' and its child code, 'Preparing for future work'. It also demonstrates how some early codes remained relevant, with fairness a constant important factor.

It assisted in the thinking process to keep the codes specific and not over generalised. In order to understand 'Being Practical', for example, it was helpful to have specific codes pointing to specific practical concerns. Although unwieldy, the process of focussed coding meant constant re-engagement with the data and close interrogation of what it meant as the focused codebook was applied systematically to the data already open coded as well as the new data. It also allowed for recognition of codes that, while interesting, were outside the scope of this study. The concept of the 'native collaborator' (see Table 2.11) for example, is worthy of further investigation but would require quite a different data set to the one collected for this dissertation.

Figure 2-13 Illustration of focused coding

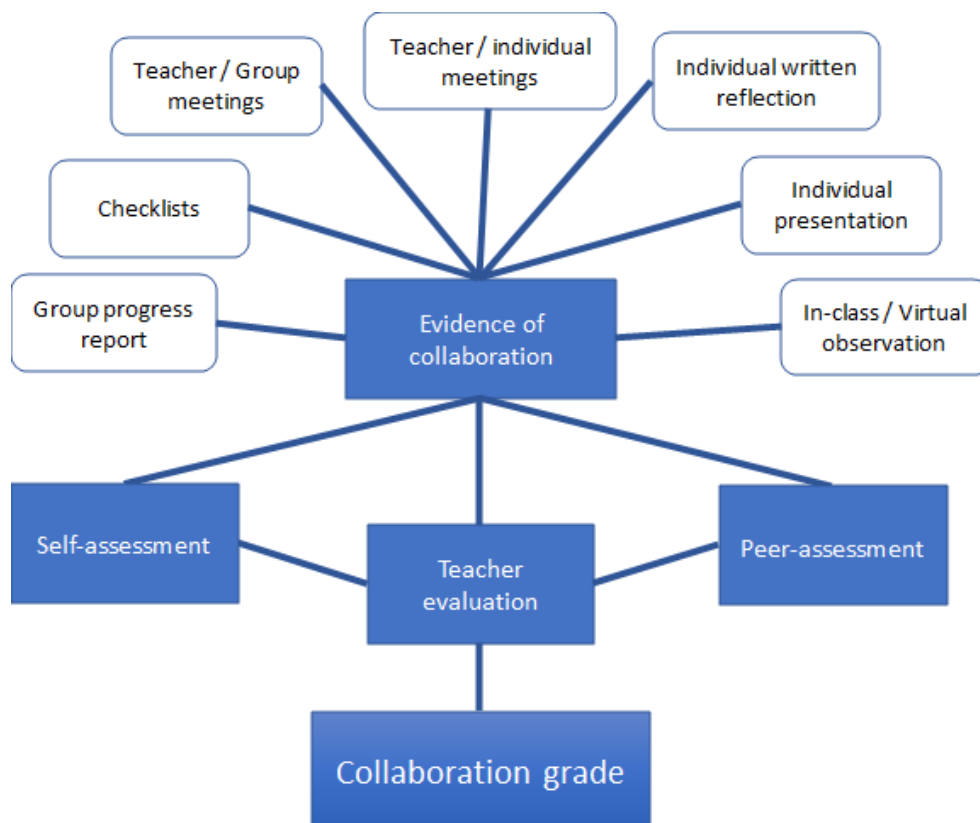
Focused coding



2.8.5 Theoretical coding and diagramming

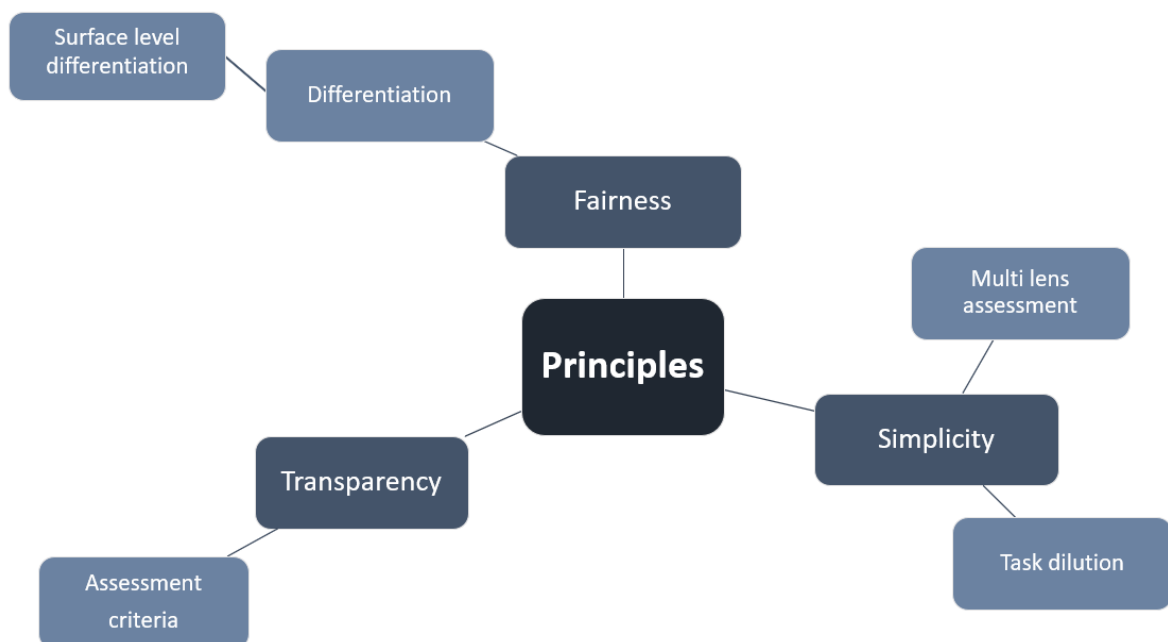
Through focused coding, memoing and diagramming, the theoretical codebook was developed. Theoretical coding in this context follows the Saldaña (2021) conception of bringing concepts together in a way that “suggests a theoretical explanation for the phenomenon (p.314). As such the theoretical codes are emergent from the focused coding, rather than a pre-determined set of theoretical codes (see Charmaz (2014) for discussion of Glaser’s pre-determined theoretical coding families). The diagramming was particularly useful in trying to articulate the relationships between ideas and concepts and, indeed, Birks and Mills argue “diagramming is *the* creative tool to use when operationalising the logic of abduction” (2023, p. 183, emphasis in original). For example, an early diagram (see Figure 2-14) illustrated how potential lenses could provide evidence of collaboration, which could be assessed through teacher-mediated self- and peer-assessment.

Figure 2-14 Illustration of multi-lens assessment of collaboration (theoretical diagramming)



This particular example was less geared towards theorisation and more toward the desire to develop a practical element in a framework for assessing collaboration, something a practitioner could pick up and use in developing their own assessment scheme for collaborative assignments. However, as analysis went on, diagrams became more theoretical, with Figure 2-15 showing the three key principles that concerned participants the most and Figure 2-16 illustrating how these principles and the guiding orientations both inform EAP practitioner collaborative assessment identity.

Figure 2-15 *Key principles shaping attitude to collaborative assessment (theoretical diagramming)*



Perhaps the most challenging theory to diagram was that of the defensive orientation (see Figure 2-17), given the multiple elements in play. Figure 2-18, which maps the components of the three guiding orientations, was incorporated this and helped illustrate the complexity of the defensive orientation in relation to the other two orientations.

Figure 2-16 Principles & Orientations informing a collaborative assessment identity (theoretical diagramming)

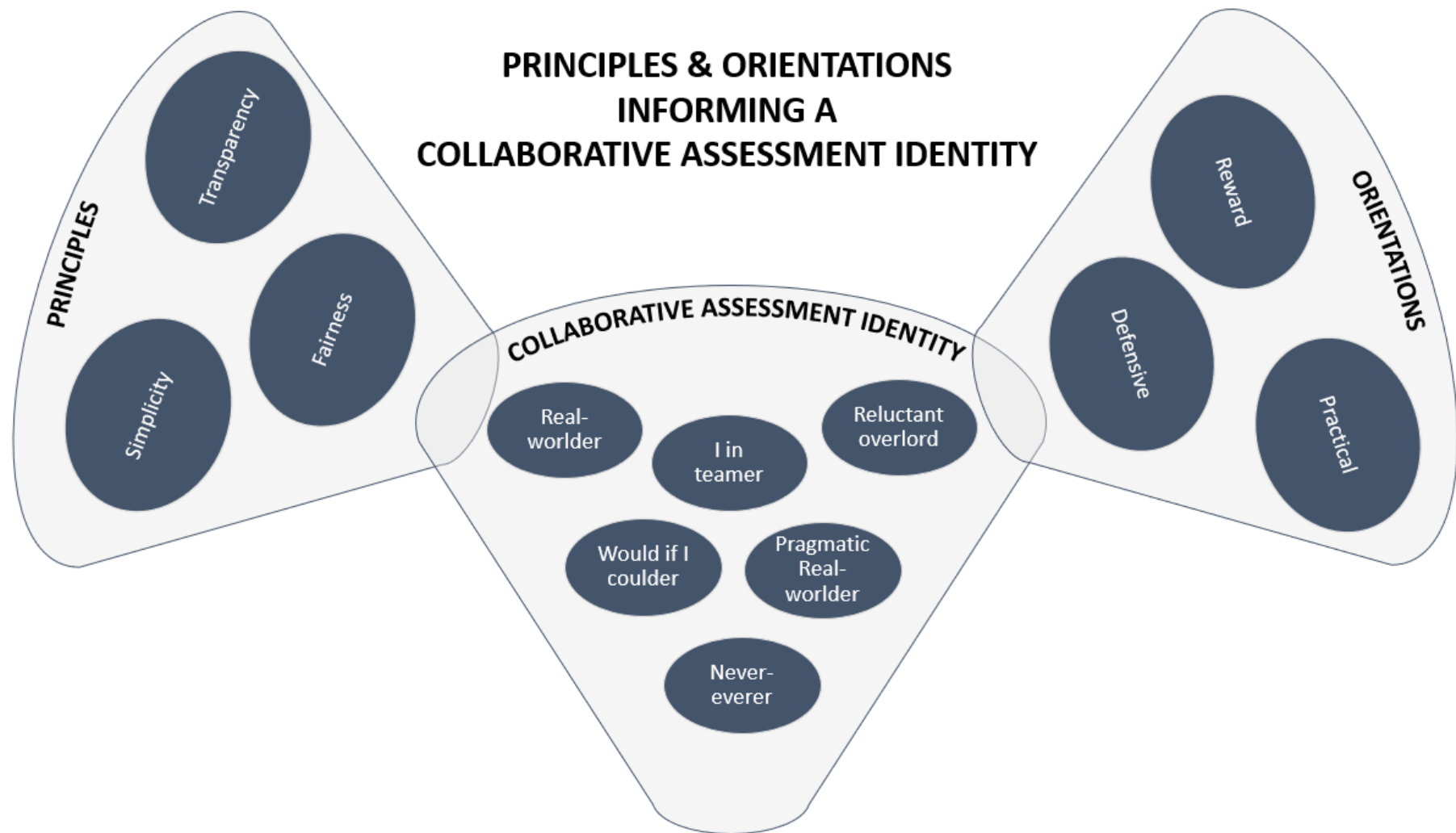
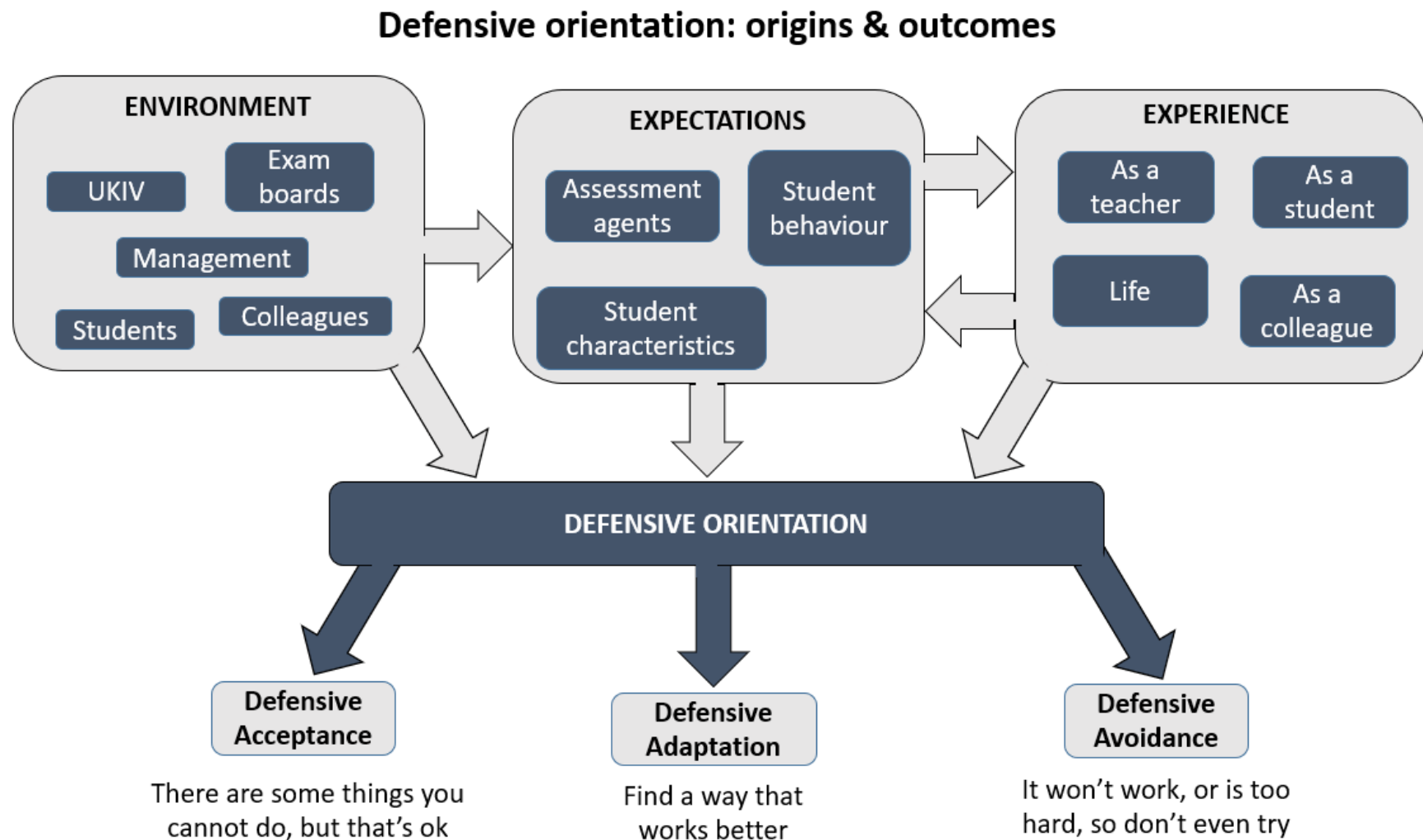
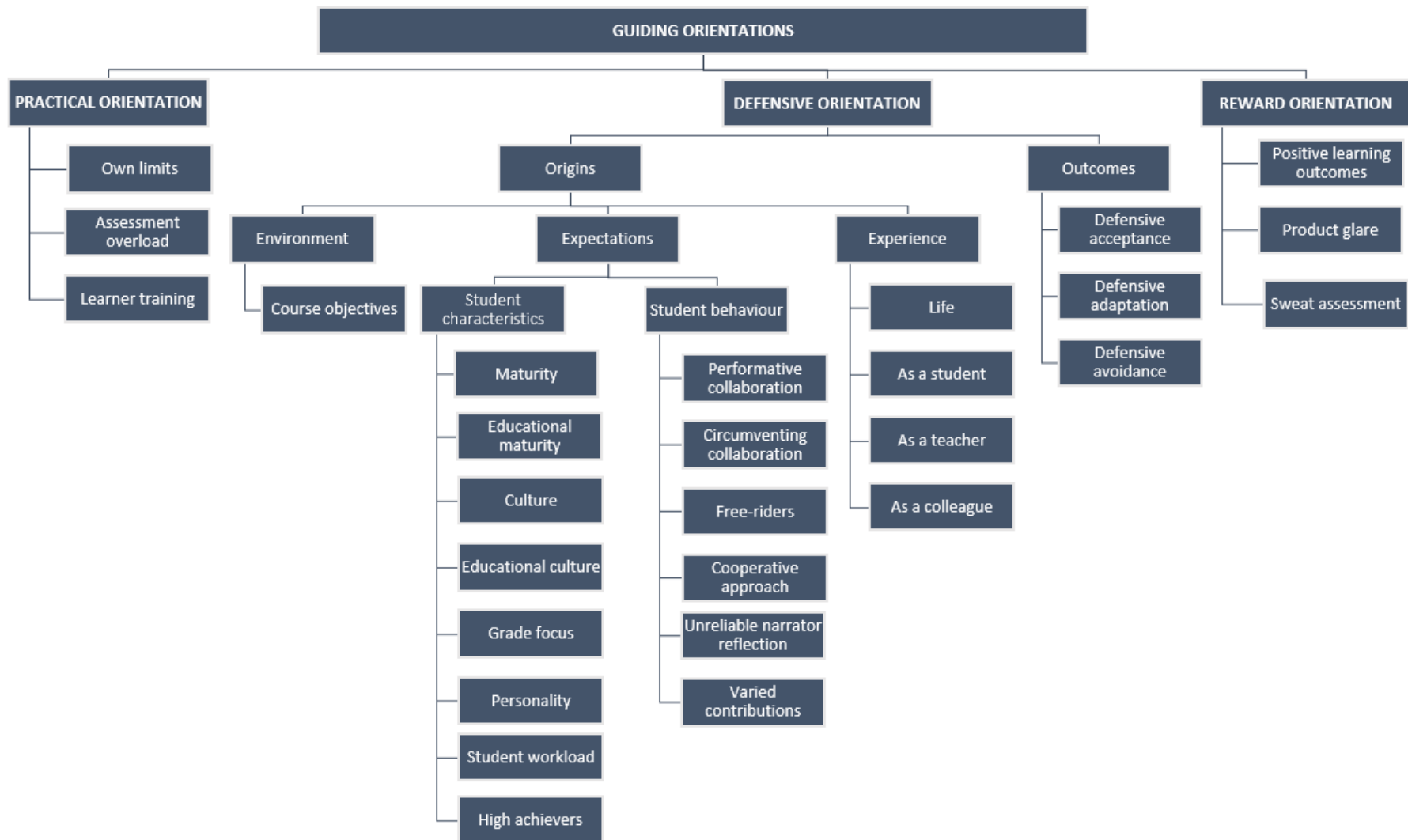


Figure 2-17 Defensive orientation: origins and outcomes (theoretical diagramming)



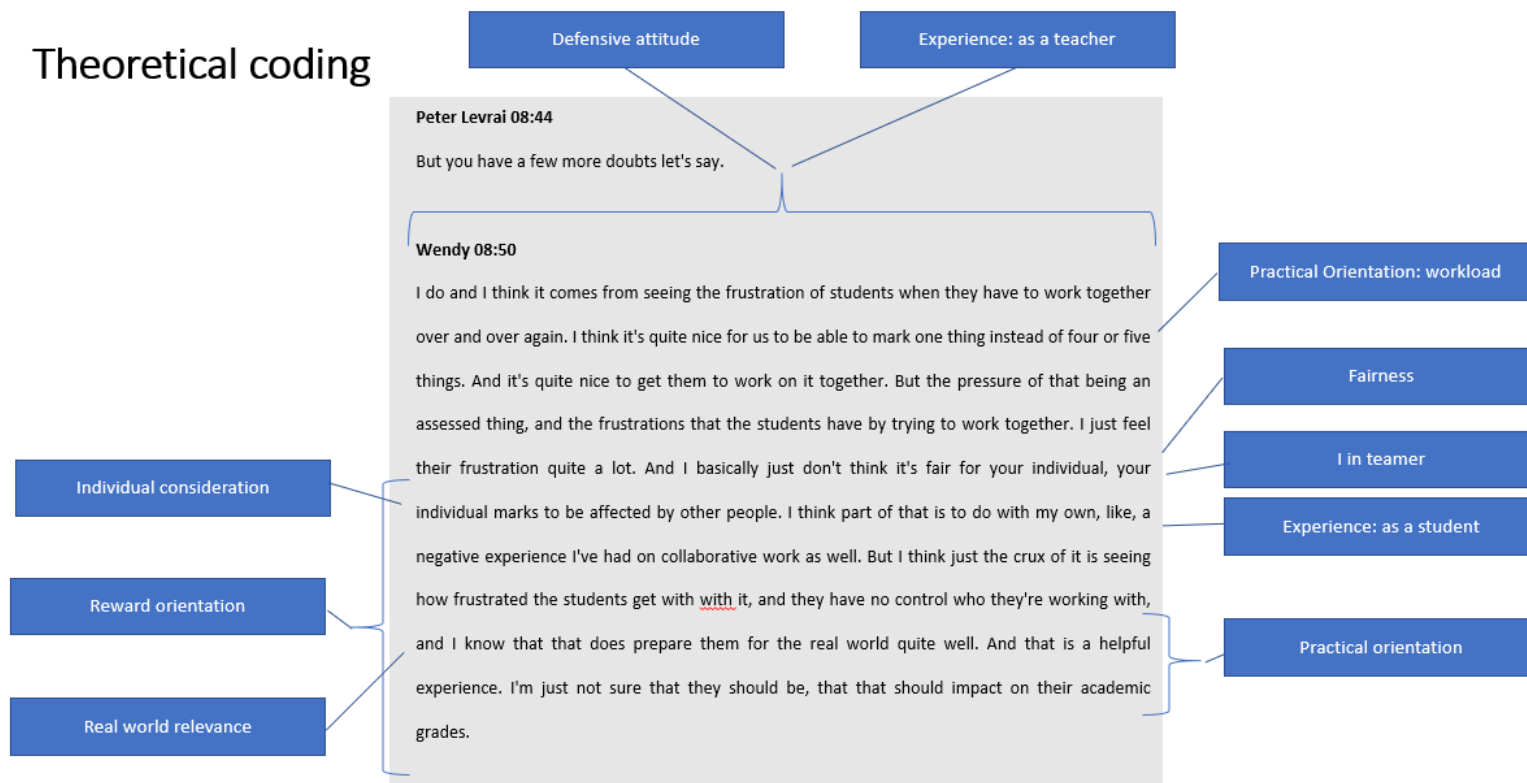
Exploring EAP practitioner collaborative assessment identity

Figure 2-18 Guiding orientations (theoretical diagramming)



The process of diagramming was vital to developing the theoretical codebook (Appendix 6, p. 47-53) and Figure 2-19 illustrates concept coding in practice, using the same extract from Wendy's interview. This remains an extract that carries multiple codes, as it speaks to some of the central concepts of this dissertation, indicating a convergence of defensive, practical and reward orientations, while also pointing to an I-in-teamer assessment identity, with a clear impact from the fairness principle.

Figure 2-19 Illustration of concept coding



2.8.6 A return to immersion

During the writing of the Findings and Discussions sections data was approached through the coding extracts, which were written into different sections of the dissertation. Diagramming was revisited and as ideas refined, so too were the diagrams. Once the coding extracts had been exploited and woven into the text there was a return to immersion, going back and re-reading printed copies of the full interviews. This was to ensure that the full interviews chimed with the findings of this dissertation, to re-contextualise the coding extracts for the researcher, and to identify any further points which resonated and could be worthy of inclusion. Reading off paper also changed the reading experience.

2.9 Going beyond sufficiency

At this stage, no more data collection was planned. The rich data set from the pre-interview task and interviews had already provided the theoretical output of this dissertation: the articulation of EAP practitioner collaborative assessment identity, which is bound by the principles of fairness, transparency and simplicity and is guided by the orientations of reward, practicality and defensive. This was clear enough from the data to meet sufficiency, meaning it was not anticipated that further data collection would add significantly to the understanding of collaborative assessment identity.

Through the writing, and in line with the intention of having a practical output of this dissertation, the idea of a collaborative assessment algorithm developed. This was a series of considerations and decision trees to guide practitioners through the process of developing a justified assessment scheme for collaborative assignments in their context, essentially encapsulating the assessment framework initially envisaged as an outcome of this work. The data gained through the pre-interview task and interviews proved sufficient to develop the algorithm and so, from this perspective too, the data collection for this dissertation was complete.

The value of the work generated through this dissertation was made evident when early findings were presented at the Norwegian Forum for English for Academic Purposes (NFEAP)

conference in June 2024 in Oslo. The presentation, which outlined this research study, collaborative assessment identity and the preliminary considerations for designing a collaborative assignment, was very well received, with participants commenting on how much it resonated. Dissemination of the findings of this dissertation was both rewarding and motivating. Beyond conference presentations it was felt that, in order to be truly practical, a starter guide to assessing collaborative assignments (see Appendix 9: First draft starter guide to assessing collaborative assessments) could be developed to aid practitioners in the development of assessment schemes for collaborative assignments. Once this had been drafted it became apparent there was scope for another round of data collection, having practitioners evaluate the practical output of this dissertation which would, in turn, help in the evaluation of the work of this dissertation overall.

2.9.1 Actualizing the collaborative assessment algorithm: Starter guide design

The starter guide is a practitioner facing text, aimed to help with navigating the assessment of a collaborative assignment. It is also essential to note it is still a work in progress, but operates as proof of concept that such a guide could have value to the EAP community. In terms of coverage it mirrors the discussion of this dissertation, albeit in a more practical way. The guide opens with a preamble, rationalising the use of collaborative assignments in EAP and then provides the vignette from the opening of this dissertation (see An opening vignette, p. 3) as it well illustrates the challenges in assessing written collaborative assignments in EAP. The guide then discusses the case for cooperative assignments, particularly in the case where it is necessary to attest to individual student's language abilities in granular way. It was important when outlining cooperative assignments to clarify the two ways students could cooperate (either on a content level, with each student taking responsibility for a different part of the content, or a task level, with students engaging in different tasks that lead to the development of the final work) and highlight the challenge in designing a cooperative assignment and ensuring students a) have equitable workload and b) are all meeting the learning objectives. The guide then outlines preparatory knowledge areas it would be good for practitioners to be aware of

going in to a collaborative assignment, incorporating aspects like the distinction between group-work, cooperation and collaboration, dynamic assessment, reflection, and self- and peer- assessment. To aid with this the guide ends with a bibliography and suggested reading, drawing on some of the literature of this dissertation.

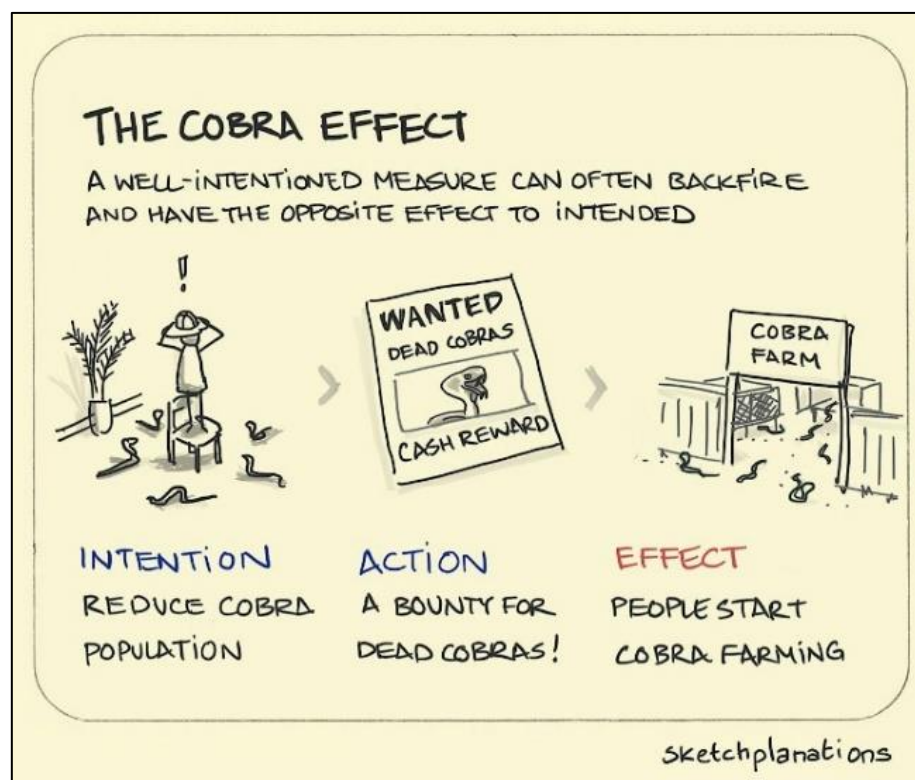
Getting into the assessment of collaborative assignments more fully, the collaborative assessment archetypes are identified and defined. The decision was made not to fully describe collaborative assessment identity in the first draft of the guide as it could become too theoretical and would need considerable space to explain clearly, which could be counterproductive in a starter guide. However, it was felt the archetypes were more digestible and typified the prevalent attitudes amongst practitioners towards collaborative assessment, which would highlight the different starting positions teachers could be coming from. These are an earlier version of the archetypes, which have been further developed since the guide was drafted. Time is then spent on identifying potential assessment lenses and providing an evaluation of the rationales for and against each. The aim in the guide is not or provide an answer, but to ensure practitioners have the knowledge and awareness to ask good questions and make informed choices themselves.

Once the assessment lenses are clarified preliminary considerations which help determine if a collaborative assignment is suitable at all are introduced, before moving into the decision trees of how to assess a collaborative assignment, depending on the importance of collaboration as a course outcome. Attention is also drawn to potential pitfalls that could be encountered during collaborative assignments such as asymmetric collaboration, assessment overload and performative collaboration. The guide closes with some examples of the kinds of questions and prompts that could be used for assessment purposes, to provide a jumping off point for practitioners to develop their own assessment lenses and collaborative assessment schemes.

When writing the guide, it was paramount to keep it a practitioner facing text, ensuing the focus was on practical advice and guidance. Although based on theoretical work, the aim was to develop a guide that could be picked up, understood and followed, without recourse to considerable reading.

However, the recommended reading was provided if users did want to look deeper. As a design choice, graphics from Sketchplanations were included to illustrate some relevant concepts, which could be used for teacher/teacher discussions or in class discussions with students. The example below (Figure 2-20) shows the cobra effect, which draws to mind the idea of assessment criteria having unexpected consequence and prompting performative collaboration at the cost of genuine collaboration.

Figure 2-20 Example of a Sketchplanation used in the starter guide illustrating the cobra effect (Hey, 2019)



2.9.2 Evaluating the starter guide

In order to determine the potential of the guide, it was initially disseminated to two new participants selected through expressions of interest at the presentation giving an overview of the findings of this dissertation at the NFEAP conference in June 2024 and a further participant who was contacted in relation to another upcoming conference where these findings will be shared. The new participants were emailed to see if they would be willing to participate and then sent the starter guide

and the information sheet for participants (see appendix 10) if they consented. Participants were asked to evaluate the guide on three measures:

- **Need** – if there was a need for this kind of guide to assessing collaborative assignments?
- **Coverage** – what do they thought think about the range of the contents and whether anything was unnecessary or should be added
- **Usefulness** – whether this guide be useful to EAP practitioners engaged in assessing collaborative assignments

with opportunity to add any further comments as desired. There were two respondents from the NFEAP conference, designated Anna and Claudette, with the other new participant designated Elsa.

It was decided all data collection would be carried out by email rather than interview. This was a practical consideration as the data collection was planned over summer, when scheduling interviews may be complex. Email correspondence would allow participants to take the time they needed to evaluate the guide and formulate their response. Open questions should also generate richer data than any kind of quantitative survey tool.

Once their email responses were received the comments were reviewed in line with what Saldaña terms sentiment analysis, looking for “opinions, emotions and attitudes,” (2021, p.159), which falls within affective coding. This feedback led to further editing and refinement of the guide until it was shared with three of the original research participants for their feedback and review, providing an evaluation of the practical outcome of this dissertation. Those approached were in positions of responsibility for courses or teaching centres, so the guide could be considered from a programme rather than individual teacher level. Wendy, as she held the strongest stance against collaborative assignments, was also invited to this final round of data collection. Unfortunately, however, she had left EAP and retrained in a different field so she declined to participate in the review.

The limited scope of this round of data collection guarded against the guide being widely disseminated prematurely. There will be scope for publication of the guide at a future date, when it is fully expanded beyond assessment to cover conception, preparation and facilitation of collaboration assignments, in addition to refining the current assessment material. This final data collection provided real world evaluation of the work of this dissertation.

2.10 Chapter conclusion

This chapter has discussed the methodology employed in this study. It described the eclectic theoretical framework of the study, which has a strong practical element of trying to determine the most effective and fairest way to assess summative collaborative assignments. Considerable time was spent discussing the changes the study went through and the impact of the COVID-19 pandemic, which caused a major pivot. However, consideration of that evolution helps explain the shift from a postpositivist to sociocultural mind set. The chapter then addressed practical methodological issues like consent and technological aids.

The chapter then turned to data collection, providing an overview of the participants and the sampling method used, which was geared toward finding data to help understand the concerns of EAP practitioners in assessing summative collaborative assignments, and approaches to assessment which would be considered favourably. The various research tools employed were discussed, as was their iterative nature of the tools as they evolved in light of early data analysis.

Once data was collected, it underwent rounds of analysis, moving from open to focused to theoretical coding. Time was also devoted to immersion, to let the data speak for itself rather than be parsed by codes. Throughout this process, memoing took place to help understanding of what was being found. This led to the development of nascent concepts, some of which flourished into the final theories presented in this dissertation, and some of which remain potential areas of future investigation. Diagramming also played an essential role in elucidating the relationships between ideas and were a central part of theorising.

Once the dissertation neared completion one further round of data collection took place. This involved participants reviewing the starter guide to assessing collaborative assignments, which was written based on the findings of this dissertation, and intended to help them through the decision making process to develop a clearly rationalised assessment scheme for collaborative assessments. The next chapter will turn to the findings of the study, drawing on the rich data to propose both a theory about collaborative assessment identity and a framework for assessing collaborative assignments.



CHAPTER 3

FINDINGS

3 FINDINGS

3.1 Chapter overview

This chapter is going to discuss the main findings of the research, drawing predominantly from the pre-interview task and interviews. The initial survey was reviewed and helped in the selection of candidates to invite for interview but the main data to inform these findings and later theory generation came from the qualitative data gained from the pre-interview task and participant interviews.

The discussion of findings will begin with contextualisation, with the participants' conception of their role as EAP practitioners and their understanding of collaboration. As the interview evolved this understanding of collaboration led into consideration of whether collaboration was a stable skill. The contextualisation section ends with discussion of why collaborative assignments are used in EAP. This sets the stage before focussing on the assessment of collaborative assignments specifically.

This will begin with examination of the four models of assessment from the pre-interview task, to give a broad overview of the attitudes and concerns of the participants toward collaborative assignments. The findings will then move on to examine the key principles and guiding orientations which steer practitioners through assessing collaborative assignments.

The three key principles are simplicity, transparency, and fairness and these principles work in concert with the three guiding orientations: practical, defensive and reward-driven. Of these orientations, the defensive orientation is most complex, being informed by environment (working in a HE institution), and the interinfluence of expectations and experience. This defensive orientation also results in three behaviours: defensive avoidance, defensive adaptation or defensive acceptance.

To locate this dissertation in the wider study into collaboration in EAP, the pertinent findings of the companion dissertation will be outlined, particularly as they pertain to identity and collaboration and the chapter closes with reviews of the draft starter guide for assessing collaborative assignments,

which is aligned with the call from the companion study for a practitioner-led pedagogy of collaboration.

It is interesting to note from the outset that none of the participants expressed confidence in the way they currently assess collaborative assignments and are very much aware of the flaws in various assessment lenses that could be used to assess collaboration. Even participants who are highly engaged during the collaborative process, following students' collaborative workspaces and checking in with groups regularly, feel unequipped or unable to assess the collaboration aspect of a group assignment with confidence. However, participation in this research was valued for providing the opportunity to discuss collaborative assessment, which Karen thought "has also helped sharpen my understanding of myself as a, as a practitioner" (Interview Transcripts, p. 197). As Barry put it

So it's very interesting, thank you, because it gives me some thoughts and some ideas as to how we might be able to make some proposals to module leaders in some of the final year modules that use collaborative approaches and how they might be able to use them more intelligently and create a more effective outcome. (Interview Transcripts, p. 19)

Olga also felt a benefit from participating in the research, explaining "I really learned something, you know, I learned a lot from this" (Interview Transcripts, p. 262). This indicates there is a desire for and appreciation of the opportunity to discuss collaborative assessment with colleagues, which will be returned to in later discussion.

3.2 EAP and collaboration

Before examining collaborative assignments in depth, it is necessary to contextualise the assessment of collaborative assignments in the participant's wider conceptions of EAP, collaboration and collaboration in EAP specifically. As Schmitt and Hamp-Lyons recognised "EAP teachers have much to

contribute to the theoretical understanding of the construct of EAP, how EAP knowledge develops and how development of that knowledge can be measured” (2015, p. 4).

3.2.1 Participants’ understanding of EAP

As outlined in Section 1.3, there has been considerable work done in the literature in terms of defining EAP and its role within HE institutions. Broadly speaking, the understanding of EAP in the literature is shared by the participants in this study. Kyle encapsulated this when saying EAP has, “a transitional or transformative element where, you know, we bring in critical thinking, learner autonomy, collaborative learning comes into that as well, that I think is all encompassed by EAP” (Interview Transcripts, p. 199). Jerry concurred, explaining,

I think EAP really still focuses on the literacies and the writing and the language. So I would like to think... but, but again, the longer I stay in, I don't want to call it EAP, I would call it Foundation or Pathways, I think, you know, what we're really doing is trying to produce more rounded, developed students in many, many ways. (Secondary Interview Data, p. 149)

Laura would agree with the concept of producing more rounded students, but would expand this well beyond the academy, feeling “we see these academic conventions, academic honesty, reliability and showing evidence, whatever you say, these things they will take with themselves to their workplaces. It's just a lifelong skill” (Interview Transcripts, p. 219).

This broad range of coverage means EAP is seen as having a preparatory function, with a responsibility to “prepare students to basically cope at this university” as Fernand put it (Interview Transcripts, p. 129) and “even for their future careers” according to Ava (Interview Transcripts, p. 42). This was echoed by Lorenzo who felt

many countries in the EU, the, the primary schools and K through 12, do a fairly good job of preparing students with a general understanding of English. And so at the university level, it seems to be less about day-to-day type of survival English and more about career professional work. (Interview Transcripts, p. 235)

While this function of preparing students for their other studies could mean EAP is seen as 'lesser' than content programmes, this was not a feeling shared by the participants, who felt EAP has a valuable role to play. As Laura explained, "EAP is the essential part of Higher Education, I would say, because it teaches them not the conventions of simply academic English, but the conventions of the potential possible future workplaces" (Interview Transcripts, p. 219). Bertha mirrored discussion in the literature (see 1.3) when reflecting on the status of EAP within HE, saying

some people view it as being something on the periphery, and others are trying to get more integrated and viewed in a, like, viewed as being equal to other disciplines. But I don't think it's quite there yet, but I can see that that's the path that has gone down. So I think scholarship is important. Yeah, showing that it's not just like a service provider role. (Interview Transcripts, p. 54)

There was broad agreement with the concept of EAP playing an enculturation role, "to help my students get a handle on this new community of practice they find themselves in" (Interview Transcripts, p. 180), in line with Lea and Streets identification of EAP as serving the purpose of academic socialisation (2006). This expanded role of EAP was recognised by Ava, who said, "academic English is no longer simply a grammar class or an English learning class, it's also about critical thinking, about ideas" (Secondary Interview Data, p. 7), speaking to EAP's role in academic discourse socialisation (Duff, 2010). This was a sentiment echoed by Jerry who said EAP was, "a way of equipping students with the academic literacies, language and research skills required to be successful when in

higher education” (Interview Transcripts, p. 175). This sense of empowering students through EAP was also an important element for Gonzalo, who said;

the language that is moving forward as the kind of lingua franca. And so it has a solid place in that because that seems to be the language that a lot of international universities are using, to, to help communication, to help students be able to communicate outside of their own school and country. It seems to be the language of most of the large international journals. So researchers, academics, students all need to be able to understand and utilise it to be able to learn from the widest variety of materials that they can get. (Interview Transcripts, p. 142)

In terms of language, all four skills are covered, as Fernand explained.

the focus is more on academic English, maybe a little bit less in speaking classes, it's just to try and get them to speak. The oral communication is also about their listening as well, trying to improve their listening. But I think as far as writing classes go, it's just heavily focused on them being able to cope, that means being able to read more complex texts than they've been used to in the past. Understanding how to find information, the right information quickly so a lot of it's you know, reading, skimming, scanning that kind of thing. Then reading for detail, then being able to you know, understand what they read, combined certain texts, the information from various texts, synthesise that to write something coherent. (Interview Transcripts, p. 129)

As some of the participants had experience teaching EAP in different contexts, there was also an appreciation that the nature of EAP is, to some extent, context dependent, with the language level of students an important factor. Where students may be entering the institution with lower language levels, the focus is on language development in academic contexts rather than higher order academic skills. Dolly and Gonzalo noted that when the initial language competence of the incoming students is

higher, the focus shifts more towards the development of academic skills. For Kyle, the shift in focus from language to academic skills is not just language level related, but also academic level related, with students with a higher knowledge of academic conventions able to deal with more academic skill work (Interview Transcripts, p. 200). However, Van commented that a lot of what he covered with lower level Chinese speakers in one institution was what he also had to cover in a research writing course for international students, including native English speakers, in another institution, suggesting meaningful academic skills can be covered at lower language levels. As he described the course content,

most of it, it was common, I mean, academic style, they were writing contractions or that kind of stuff. It was just ... I mean the cohesion, obviously, when you go into things like the broader academic literacy stuff about source use, that kind of stuff. It was pretty common across the board. (Interview Transcripts, p. 267)

This belief in the value of what is taught to any student, regardless of language competence, was echoed by Cristobal, who said, "And so, I think even native speakers would benefit from more training and the language that's used in academia" (Interview Transcripts, p. 142). The general value of the types of things taught in EAP classes was explained by Van;

Right now, I'm supervising dissertations and the, you know, MSc TESOL dissertations and maybe 50%, at least 50% of the things I comment on, were things that I would comment on as an EAP teacher. You know, you haven't explained this well enough. You need a source for this, or you haven't evaluated here, or is it language or is it all you know, who cares? But it's ... so and I don't see it as distinct. It's ... any faculty person would, in any discipline, would be giving that kind of EAP instruction and scaffolding and teaching would benefit from it, I think. So I see it as extremely valuable. (Interview Transcripts, p. 266)

That is not to suggest there was not tension related to the concept of EAP. In relation to the balance between academic skills and language competence in EAP courses, Dolly lamented ongoing changes in her institution where “we've gone through so many curriculum revisions, it's hard to keep track” (Interview Transcripts, p. 96) with the changing demographics of incoming students (typically at a lower language level than previously) changing the focus of the EAP courses more toward language development, albeit still with an EAP flavour. Wendy was something of an outlier, feeling there needed academic skills were overly focused on in EAP and there needed to be more language development but even she conceded, “a lot of EAP is skills and I think, my opinion is that they need language and there should be more language in EAP. But that the reality is, a lot of it is teaching skills” (Interview Transcripts, p. 283). Within the UK context, there was recognition of the role of EAP as a “selling point to entice, to entice international students” (Interview Transcripts, p. 54), echoing the neoliberalisation concerns of Ding and Bruce (2017) and Hastings (2019). However, the overall concept of EAP was that it serves an essential role as part of a student's academic career, enabling them to succeed in the rest of their studies by helping them with their “ways of working” (Interview Transcripts, p. 114) as well as language development.

There is also a recognition that teaching practices in EAP may be stronger than those in faculty courses. Jerry lamented attending high level meetings in his institution and “getting tired of hearing things that I know we've done in EAP, and EFL for years,” with “things that we just do, second nature, that doesn't happen in the faculty” (Secondary Interview Data, p. 161). This helps locate EAP as an area that can help students, and help students in a way that faculty cannot, or do not. The respondents shared a sense of responsibility to help equip students for their future studies and professional lives, with this responsibility extending to the development of students' collaborative skills.

3.2.2 Participants' understanding of collaboration

In relation to collaborative skills, the interview moved on to ask participants what good collaboration looked like, what they hoped to see when students were collaborating. Karen highlighted the essential role of collaboration, explaining

I'm a great believer that collaboration improves everybody's learning processes. When we learn to work together. We don't, I mean, you know, at the very top level as humans, we don't exist in a vacuum. You know, we exist through and for each other, so collaboration is absolutely vital.
(Interview Transcripts, p. 184)

Within the context of a collaborative assignment, Andrea saw collaboration as a combination of content and task, meaning an end product that follows the expected academic conventions and a piece of work which is organised. It was unclear how this would be different from assessing an individual piece of work and is demonstrable of the limitations to the four pilot interviews, which were rigid rather than allowing for natural dialogic exploration. However, cohesion was a factor also identified by Barry as an indicator that students had worked together and Laura too emphasised organisation and academic rigour.

Ava expanded beyond the product, stressing the engagement of the students throughout the process of product development, emphasising "I'm more concerned about the process of how this team together crafted a piece of work together, as opposed to just the finished product of one essay" (Interview Transcripts, p. 43), which is supported by Cristobal, who also saw students working together throughout as key. Both participants were echoing Williams' (2017) urging to consider the process of collaboration. Bertha went further into what the process of collaborating could look like and the full response is quoted in full below as it encapsulates the breadth of what collaboration is seen to involve.

I suppose you're hoping that, like I said, that the students have had an opportunity to use the, like, use their language in a variety of different scenarios erm, that's your like, discussing their ideas, maybe they're writing something, their research and something that they're giving feedback to each other about their, yeah becoming better at seeing how things can be improved and working on that themselves, that it's not like going to the tutor all the time, that they're developing those skills amongst themselves. What other learning outcomes? Yeah, all those all the other skills involved, like research skills, whatever skill it is that you want them to develop, like speaking, if it's a presentation, that they've developed some sort of skills in putting together presentation, researching a topic, and that they've used some, like, language that they will be able to use in the future as well. I think also, maybe another important learning outcome is that they're developing skills and language that will be useful for them after, like during their course, whatever course you're going onto, and afterwards as well. Erm, like for future jobs, if you think of all the times that we have to collaborate with people and work on things as a group, and what else, I was gonna say something else, yes, that it builds confidence and helps the students to maybe reflect on their own, their own approach as a member of a group and helps them to reflect as well on how they interact with others. I think that's an important part of it, being able to reflect. (Interview Transcripts, p. 56).

The full complexity of a collaboration is captured here, with language development, academic skill development, learner autonomy, evaluative judgement, social skills, preparedness for the future, understanding of themselves, and reflection, all coming together. Kyle concurred in seeing autonomy as a key element of collaboration, as students are learning to operate “without the teacher or the lecturer overseeing everything that they are doing” (Interview Transcripts, p. 201).

Social skills were also mentioned by Lorenzo, who saw the development of interpersonal skills as an important aspect of collaboration, and Olga, who felt collaborative assignments gave the students opportunities to meet new people and make new friends. Dolly added the idea of good

rapport to social skills, with the group enjoying the process of working together where “everybody chips in more or less equally,” (Interview Transcripts, p. 98). This raises the issue of parity as a feature of good collaboration, which Gonzalo would also agree with, seeing the logistical aspect of collaboration as a means of guaranteeing equitable workload. As he explained,

I guess it comes down to having clear roles and clear, clearly defined kind of areas that people are going to work in is pretty critical to making sure that the work is seen as fair amongst all the participants. (Interview Transcripts, p. 144)

Karen concurred with the need for clear planning from the outset, explaining “first and foremost, make sure that we have a clear understanding of what the expectations are” (Interview Transcripts, p. 182).

Collaboration having so many dimensions has opportunities and limitations when it comes to assessment, as there are many things which *could* be evaluated but little scope to evaluate *everything*. Given the richness of collaboration, the assessment of it can never be complete. Further complicating matters is a line of enquiry that emerged in the later interviews, questioning if collaboration is a stable skill, meaning whether a good collaborator is always a good collaborator or whether each collaboration has a unique alchemy. Should collaboration be stable, the assessment of it has some predictive value for the future. Should collaboration be unstable, the assessment of collaboration can only speak to a single performance, which may have little relevance to future collaborations, as Boud and Bearman (2022) and Pitt (2000) would recognise. This is similar to Vygotsky’s (1978) concern that IQ tests just provided a static snapshot rather than a true picture of a student’s intelligence. Arguably, the best test of a collaborative assignment is the students’ performance in the next collaboration, or the next individual assignment requiring the skills and knowledge developed through the collaborative assignment.

When discussing the stability of collaboration, grouping was a concern for Dolly, explaining “sometimes like you just get stuck in like not a great group, right? And so, like, so if you are unlucky, and then your, your, your group members aren't holding up their end of the bargain” (Interview Transcripts, p. 106). Laura had an experience that exemplified this, with a strong student who took a domineering role in one group, effectively operating as a poor collaborator, but was a more effective collaborator in a different group, so she recognised that people change depending who they are working with (Interview Transcripts, p. 232). This would chime with Cristobal, who felt “effective collaboration is depends on the people involved and one person could be good in one group and bad in another group” (Interview Transcripts, p. 81). He went on to suggest that “it reflects the personality of the people. I mean, do they want to work with others? Will they listen to others?” (p. 81) suggesting there are collaborative tendencies that may remain stable but that will be “affected a great deal more by the other people in the group and by the teacher in the way it's managed for sure,” (p. 81).

Gonzalo also took a firm stance that collaboration is not stable, explaining

I would say every one is different. I know, I know that I collaborate with some people extremely well. And we work very well together. And I know some people who I have difficulty working with, and it's difficult to collaborate with them. I don't think that my skill set has changed when I'm interacting with these people. I think it's, I'm working with a new variable there. And also, the projects that we're working on are different as well. So ... I know what my skill sets are, I know how I collaborate, I know how I go about doing things. And that works great with some, some of my colleagues, and then maybe it doesn't work with a couple of them, you know, for whatever reasons. (Interview Transcripts, p. 152)

Fay would concur with the idea each collaboration is unique, saying

I think it depends who you're collaborating with and what it is about. I know that I work very well with certain people. And I know I don't with others, that I would not be a good person to have on, you know, this other person's team because I just, you know, something the way in which we do things or in the way we think is not compatible. (Interview Transcripts, p. 125)

but she went on to note that

you can learn certain skills that you can then use and take on, but they won't always work to the same extent, in different groups. But the more aware you are of things, the better you will be able to use those skills in different contexts. (Interview Transcripts, p. 125)

which supports the idea of collaborative tendencies which could be developed. However, the alchemy of collaboration, the unique confluence of task and group members, adds further ambiguity to collaborative assessment.

One important thing to note, however, is that few of the participants had any formal training about collaborative learning, collaborative assessment, or direct experience of collaborative assignments as students themselves. As Dolly said regarding learning about collaboration, "that's definitely not a class that you can take in a TESOL programme" (Interview Transcripts, p. 112) and when reflecting on his training, Gonzalo said "I don't remember any specific training I had in any of my courses. Where it said, "Okay, this is how you collaborate." I don't ever remember anything like that" (Interview Transcripts, p. 143). This was also the case for Karen, explaining her knowledge of collaboration "it has sort of, I guess, emerged, and yeah, quite quite organically in that sense. I can't say that it's anything that I learned being taught explicitly" (Interview Transcripts, p. 181). The participant's knowledge of collaboration, which sits in line with the conception of collaboration in the literature, has developed more instinctively. This lack of formal education is brought particularly to light when considering the assessment of collaborative assignments specifically, with Jerry recognising

it is an area in which he and his team would benefit from professional development as “that was the bit that I found a gap in my knowledge really. I you know, I can get tell you what I think but I don't think I know what's best” (Secondary Interview Data, p. 156). It is also important to note, however, that there are few solutions to be found in the literature, even for someone conversant with it.

3.2.3 Collaboration in EAP

One of the further contextualising questions in the interview was why collaborative assignments were used in EAP. Karen made the connection between collaboration, learning and knowledge, explaining “I think it goes to the very heart of how we believe we learn, you know, and how we believe we create create knowledge” (Interview Transcripts, p. 183). Andrea also valued the learning that could take place, explaining collaboration “can be really fruitful and helpful with skills development, and language proficiency development” (Interview Transcripts, p. 10). This is something Lorenzo would agree with because he too saw learning benefit in terms of “both the skill technical skills, focusing on language, effective communication, but also on human interaction” (Interview Transcripts, p. 237). Karen went on to say collaboration was used “because it mimics and it helps prepare students for how they can be expected to work in their real life” (Interview Transcripts, p. 183). Kyle expanded on this, capturing the argument of many participants saying

Because they are more widely used in higher education now, I think, than when I was a student. So I would say it would be fairly unlikely for a student who was studying in the UK or, or a UK International University to go through their academic career without doing some sort of collaborative assignment somewhere along the line. I don't, I don't see that that would happen. So that's the first thing. Why, why are they used in in degree programmes? Because it's an important skill in terms of developing employability, I think. It's like a skill that has value outside the degree programme, it's something that employers are going to ask about. (Interview Transcripts, p. 201)

This employability theme was echoed by Van, who also highlighted opportunities for peer learning, explaining

One is the opportunity for peer to peer scaffolding. I mean, it's EAP so obviously this ... they're getting used to ... there's a usually a bridging element to the course, so they can rely on their other peers and you're formalising that rather than letting them informally just look for a generous classmate. That's the first reason peer to peer scaffolding. And then the second reason is EAP often likes to get generic skills, generic learning outcomes and you know, a generic learning outcome that everybody values is teamwork. (Interview Transcripts, p. 267-268)

Fay also spoke of the link between collaboration in EAP and in the real world, explaining

in the world of professions, people don't work on their own on anything nowadays, everything is collaborative, and they need to be trained in how to work with other people, how to communicate, how to come to a compromise, how to propose things, how to not offend. (Interview Transcripts, p. 113)

Olga added a reflective element since collaboration “helps them to be looking into their own work from different perspectives” (Interview Transcripts, p. 252). Bertha felt collaboration provided opportunities for genuine communication between students and also added the logistical factor of collaborative assignments being “a bit less labour intensive for the tutor in terms of saving time,” (Interview Transcripts, p. 55) as there are fewer assignments to grade. This logistical benefit was also for part of the rationale for Jerry, who admitted “I'd like to say that that's not the primary function of collaborative assignments but I do think from a managerial perspective, let's be honest with ourselves, it is one of the reasons” (Interview Transcripts, p. 175). Workload was also a theme for Fernand, although he admitted in terms of lightening workload “I don't think it really ever works that way”

(Interview Transcripts, p. 130). Tanya also felt there was a workload issue since short course length meant

there's not enough hours in the course for every student to write like a five or six page technical report and it's too much. So the idea is, well, you can show me that you know how to do it by working with others and, and working together to produce a technical report (Interview Transcripts, p. 33)

which would chime with Laura, who felt there was a workload benefit for both students (who did not have to individually produce long pieces of work) and teachers (who did not have to mark many long individual pieces of work), especially as a teacher could be dealing with 90 students (Interview Transcripts, p. 220-221). Ava also felt cohort size was a factor, since “we are facing 96 students per semester. So it's not very possible to monitor 60, 96 projects” (Interview Transcripts, p. 42). However, that was only one component of various motivators she had to use collaborative assignments, with the others being

skill development, cooperation, collaboration is a major part that they need. So, one of them is, not only in their other classes where they will be facing collaborative assignments. Even in their future careers, they might be a very important skill. Another thing is for research and academic work, it's always good to have another pair of eyes and always have, to have someone to consider, reflect with and it's always ... and also in a group different people have different skills so ideally they learn from each other. (Interview Transcripts, p. 42).

For Gonzalo, the students learning from each other was key. As he rationalised it

students are still struggling, they're struggling with two different things. They're struggling, as we talked about earlier, they're struggling with the skills themselves, understanding what they're supposed to be doing. But they're also struggling with the language. And so having, having peer support, to, to work through an assignment, or to learn how to do an assignment ... Collaborating on that type of a task can really, I think, help the students *understand* the components that go into an essay. (Interview Transcripts, p. 146, emphasis in original)

which is something that chimes with Duff (2010) as it speaks to the role collaboration can play in students' academic discourse socialisation. Kyle also saw collaborative assignments as a way of students developing necessary skills in a supportive context, explaining

they're sharing the responsibility for it, there's not the individual pressure on all of them, but they're actually being walked through all stages of the process, which is really what they needed at that stage, they were never gonna produce a great essay, I don't think. It's more about teaching them about the process of it and at the same point, as well, was developing the collaborative skills. (Interview Transcripts, p. 203)

Gonzalo went on to discuss how students bring their different strengths to bear and coach each other, which would also resonate with Humberto, who thought collaborative assignments were used for both the opportunity for students to learn from each other, and also fulfilling the university intention to develop collaboration in students as an important graduate and work-ready skill (Interview Transcripts, p. 160). Rather than looking forward, Wendy made the connection between collaboration and the dominance of the communicative approach in language teaching, suggesting

collaborative writing it's kind of a culmination of communicative language approaches, really. So the belief that they can learn from working with each other, and learn from pooling their writing skills and helping each other and that that will benefit them. And that it's better than working on your own and having all your input coming from the teacher and everything being very teacher centred, or teacher directed, that it kind of pushes them into into working, in student centred sort of ways. And that that benefits their, their spoken English, their written English, their ability to work in teams and in groups. (Interview Transcripts, p. 283)

but she also recognised, as had Kyle, that there is more collaboration in the faculties, resulting in more need for collaboration in EAP. Karen and Gonzalo discussed the relative importance of collaboration in EAP and Gonzalo's statement is quoted in full as it locates collaboration in the broader scheme of an EAP course, speaking to its relative importance in terms of the other things EAP courses have to cover.

I think that's the problem that we all run into teaching. You know, it may be part of our programme learning objectives, you know, or course learning objectives. But they're never the main focus of it, you know. And so, as I was saying before, it's that, as an EAP teacher, I think that the biggest focus for us isn't collaboration, you know, or teamwork. We're trying to teach the students how to to critically think, to solve problems and things like that. Yes, teamwork, collaboration, are components, but, but also we're teaching them how to write an essay, for example, or to give a presentation or whatever. And those things are, in the grand scale of things, we're trying to teach all of these skills and all of this knowledge, and teamwork is one, one component out of maybe 10, things are going to go into the project. So yeah, it's a challenge, I think to to give it the weight that it probably should have in the courses, and also to kind of, well just we don't have the time to do it in the courses often you know, because it's not such a big focus. (Interview Transcripts, p. 155)

In sum, there were parallels between the rationales for collaboration in the literature and from the participants, again underscoring the awareness and professional competence of the interview subjects. Their responses drew on sociocultural theory (with students scaffolding each other's learning), student-centred learning (with the development of autonomy from the teacher), and broader recognition of collaboration as a global skill (with the connection to transferability and employability), albeit with some concern of the commercialisation of HE and the need to get “bums on seats” (Interview Transcripts, p.130).

3.3 Participant evaluation of models of assessment

The attitudes participants had toward collaborative assignments were well illustrated through the pre-interview task, in which they evaluated four models of assessment for collaborative assignments, two of which looked solely at student work and two of which attempted to evaluate the collaboration itself, in addition to the collaborative work produced.

This section will review the evaluation of each model in turn, before outlining some key takeaways of what participants would like to see in assignment schemes for collaborative assignments. Interview comments from the pilot interviews will also be considered, where the participants reviewed the models as a spontaneous talk-aloud task rather than a pre-interview task. In this section, only dialogue directly connected to a particular model of assessment will be examined, with broader assessment issues arising being discussed in the section that follows.

One component of the pre-interview task was participants assigning a star rating out of 5 to each model. This was not intended as a robust statistical measure but more as a general overall impression rating. It also helped determine the structure of this section where the rationale of the highest and lowest rating of each model will be discussed, before a more general examination of positive and negative features identified in each model by the participants. Table 3.1 below illustrates the participants' ratings for each model, with Model B, incorporating individual complementary assignments into a collaborative assignment, finding the most favour.

Table 3.1 Participants' star ratings for the four models of assessment from the pre-interview task

Participant	Model A	Model B	Model C	Model D
Ava	4	5	3	3
Fernand	1	3	3	4
Humberto	3	4	4	5
Karen	1	4	2	1
Jerry	3	5	3	2
Olga	1	5	1	3
Lorenzo	5	4	4	2
Wendy	2	4	3	1
Bertha	2	4	4	4
Van	4	5	3	2
Cristobal	4	4	1	2
Dolly	2	4	3	3
Fay	5	5	3	3
Gonzalo	3	4	4	2
Kyle	3	4	4	5
Laura	3	4	3	3
Mean	2,88	4,25	3	2,81
Median	3	4	3	3
Mode	3	4	3	3,2
<i>Andrea, Barry, Dorian & Tanya participated in the pilot interviews and discussed the models of assessment in a spontaneous talk-aloud rather than completing the pre-interview task so are not represented in the data above.</i>				

3.3.1 Model A: Group grade

In Model A, all the students in the group receive the same grade, based on evaluation of the final essay. Model A provoked star ratings across the board (see Table 3.2), with a mean star rating of 2.88, indicating an overall slightly negative reaction amongst the participants.

Table 3.2 Star rating for Model A

★	★★	★★★	★★★★	★★★★★
3	3	5	3	2

The model was most highly rated by Lorenzo and Fay, who gave it five stars. During the interview, it became clear Lorenzo had misunderstood Model A, thinking it allowed scope for assessment of the individual, but still felt it was a justified approach to assessment, considering it reflected what happens in the real world. Fay also viewed it positively, with the proviso the teacher is 'present' during the process of essay development in the students' shared document. She also felt that "they're all responsible at the end of the day for producing something that is good" (Interview Transcripts, p. 119) and through the Model A approach she would not have "to police the process" (Interview Transcripts, p. 119). While she recognised this could result in unequal contributions from the students, she also made the real-world connection, citing Pareto's Law, that 20% of people do 80% of the work so "Even though I *know* it's not an equal distribution of work. I know. It's fine" (Interview Transcripts, p. 119).

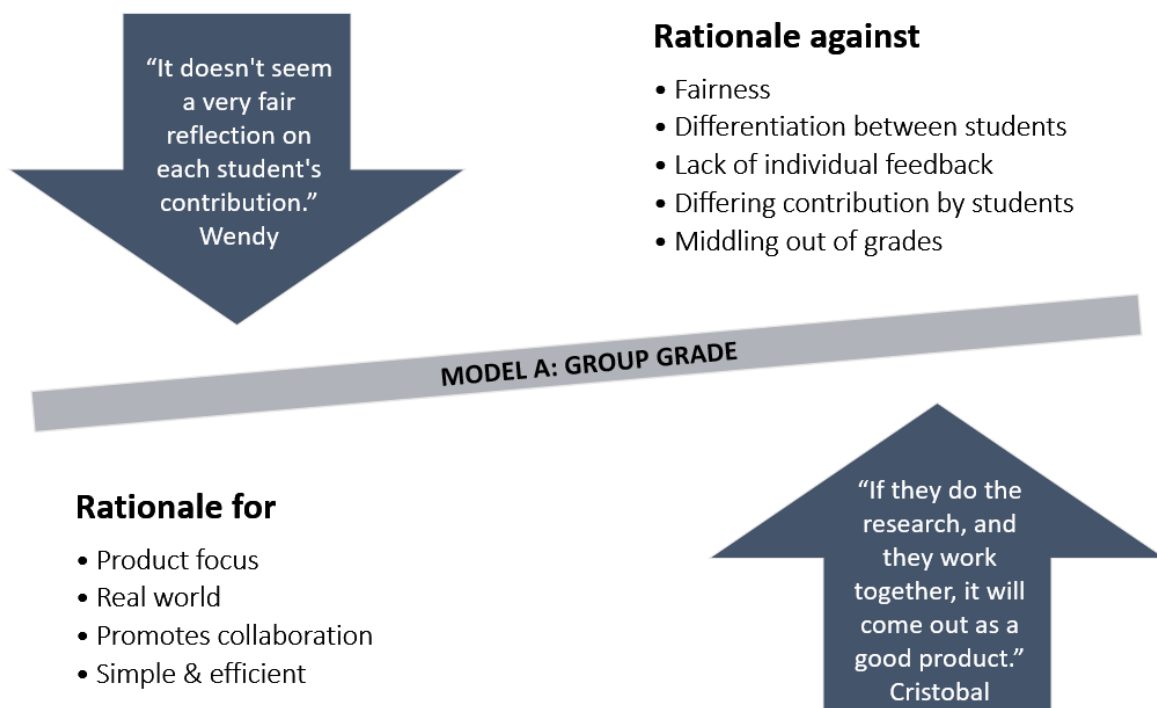
In contrast, Fernand, Karen and Olga rated Model A with the minimum 1 star, with Fernand's pre-interview task comment encapsulating the main issues voiced by others.

This method of assessing the writing project would not allow for a fair assessment of an individual student's ability, nor their contribution to the assignment. It would either reflect the best student's ability, motivation and contribution, or some kind of average mark, which might be unfair to some student(s) and/or flatter others. (Pre-interview Tasks, p. 26)

During the interview he recognised it could be an acceptable model if the aim of the assignment was collaboration and students knew they had to work together to get a good result. However, he still thought there had to be scope for some differentiation between students and an opportunity to "reward them for effort" (Interview Transcripts, p. 135), which was a theme that came up throughout his interview. Karen's main issue with Model A was its lack of fairness because students would be getting the same grade despite the teacher not knowing what each student had done, leading to a potential lack of equity. Through further discussion, considering a context where the teacher was

present during the process and could see what was going on, she became more comfortable with the idea as long as she had a sense of how things had developed and who had done what to know there was a level of equity in the group. Olga echoed Fernand's concern about being able to assess the students in terms of an "individual's contribution and their level or improvement of writing" (Pre-interview Tasks, p. 124). Figure 3-1 below illustrates the main arguments for and against the group grade assessment scheme.

Figure 3-1 Summary of evaluation of Model A: Group Grade



3.3.1.1 Model A: Positive reactions

Through the pre-interview task and interviews, one of the features of Model A that appealed was that it was output driven, meaning the focus was on the work the students produced. As Cristobal suggested,

essentially, if they, if they do the research, and they work together, and they write it, it will come out as a good product. So, you know, ideally, the product should be the easiest way of and most effective and most accurate way of assessing their work (Interview Transcripts, p. 85)

recalling Meijer et al.'s (2020) evaluation of group assessment fostering positive interdependence, leading to students helping each other to make the best work possible. This assessment approach could be motivating for students as they would have to work together to ensure the quality of the final artefact. As Lorenzo put it in his pre-interview task, "I like this model because the students must do their best to ensure quality throughout the whole process and essay" (Pre-interview Task, p. 53). He expanded on this in the interview, suggesting "it would really encourage efficient collaboration, I would hope, at least where each student is hoping to do their best work for themselves as well as for the group" (Interview Transcripts, p. 244). He went on to liken it to a work context and how a team would be judged by the work it produced rather than how they produced it. This realworldness argument, was echoed by Lorenzo and Van also felt such an assessment scheme would encourage collaboration as the students "really can't just say, "Oh, that person is not doing well. I'm just gonna leave them to their own devices," they kind of have to proactively get involved and hopefully that works out positively" (Interview Transcripts, p. 274). Another appeal of the model was its efficiency (Pre-Interview Tasks, p. 5), with Karen and Cristobal considering it "easy" and Bertha recognising that "would be the least time-consuming model for marking and providing feedback" (Pre-interview Tasks, p. 12), although she went on to unfavourably weigh that against the negative considerations discussed in the next section.

3.3.1.2 Model A: Negative reactions

While acknowledging the efficiency of Model A, Bertha also felt it could be the most problematic of the models due to unequal contributions by group members and lack of individual feedback. This highlights what was the most widely commented issue: the potential unfairness of the assessment

scheme, as it did not allow for differentiation. As Tanya put it, “it's straightforward, it's easy to understand, but I do think there's some things lacking here as far as, you know, nuance between students” (Interview Transcripts, p. 35). The lack of fairness in Model A was something that was commented on by eight of the participants. Dolly and Kyle raised the issue of free-riders getting the same grade as students who had made full contributions, with Wendy also showing concern for the situation where

you're very likely to have one or two students to do all the work. And one student who really doesn't pull their weight at all, but then that student will get the benefit of whoever's the strongest person in the group” (Interview Transcripts, p. 290)

which illustrates the kind of behavioural misalignment Meijer et al. (2020) warned against if adopting this kind of assessment approach, which does not foster individual accountability. Karen also had concerns about equity and questioned “how much do I really know what each person did?” (Interview Transcripts, p. 191).

It was accepted that “it's very hard for the students to all contribute equally” (Interview Transcripts, p. 106), particularly in a complex task like essay writing, which incorporates “creating, evolving ideas & iteration” (Pre-interview Tasks, p. 47). Despite this reality, fairness was still a principle concern. As Gonzalo put it,

Don't, don't get me wrong, I don't think we'll ever get it so everyone's doing *exactly* one quarter of the work. That will never happen. And there are going to be students that are going to do a little bit more because that they want to do more, they want to make it as good as perfect as possible. And there will just be some that think, "Okay, I've done enough. I've stuck to the letter of what I was supposed to do. And that's all I'm doing." But I think we have an obligation to try and make it as fair as possible. (Interview Transcripts, p. 150, emphasis in original)

The most widely cited concern was the lack of differentiation, commented on by 10 of the participants. Part of this is connected to the fairness concern, with students feeling “aggrieved at getting the same mark as everyone else if they feel they worked harder or performed better” (Pre-interview Tasks, p. 40). The potential for negative emotions was also an issue for Bertha recognising it would “likely be the most problematic in terms of students feeling disappointed or resentful” (Pre-interview Tasks, p. 12) while the impact on motivation was a concern for Kyle since “the free-rider can think “Well, hold on, I don't actually have to do anything here. I can just let these guys do it. And I'll get the same grade anyway.” So it might encourage that sort of behaviour” (Interview Transcripts, p. 207). Connected to this, Laura also saw the potential for demotivation in laid back students in the face of a group grade since “at the end of the day, she knows that she will receive the same grade with the high achievers” (Interview Transcripts, p. 224).

The lack of insight into the work of the individual also meant the assignment could not have a diagnostic function as “I think having the grade only based on one group work is hard to see individual's contribution and their level or improvement of writing” (Pre-interview Tasks, p. 124). Van, who saw students over two semesters, with the collaborative writing assignment taking place in the first semester, also observed that

they got through the group, a group assignment, and then it was the second semester and you realise, “Okay, this person just basically never does homework.” Or they just don't know how to do a citation or you know, and so, oh, they must just have done nothing in that first semester. And so it would be better to kind of identify that earlier so that you could address it earlier. (Interview Transcripts, p. 275)

Bertha also voiced the concern that students could not receive individual feedback on their work which would allow them to improve their future writing. She questioned how useful group feedback would be to an individual, ceding they could get some benefit from it but that it may be

difficult for students to transfer group feedback to their own individual work (Interview Transcripts, p. 62). Even Lorenzo, who was broadly supportive of the group grade approach, felt “there's a part of me that sort of wants some individual leeway” (Interview Transcripts, p. 244).

There was also the issue that a group grade would have an impact on student results and “might be unfair to some student(s) and/or flatter others” (Pre-interview Tasks, p. 26). Kyle discussed the middling out impact of collaborative assignments, where grades tend to cluster more to the middle, pulling less able students up and more able students down (Interview Transcripts, p. 204) and Van was worried about students who did not have the individual ability to pass a course making it through due collaborative assignments. These concerns are in line with those of Berry (2007) and Meijer et al. (2020) regarding the impact of collaboration on students’ grades.

3.3.1.3 *Suggestions for implementing Model A*

Although there were concerns and issues with Model A, participants saw ways in which a group grade approach could be potentially be implemented, as summarised in Table 3.3, below.

Table 3.3 *Suggestions to implement Model A*

- | |
|--|
| <ul style="list-style-type: none">• Dependent on the course learning outcomes• Appropriacy may be impacted by cohort• Requires focus on the process and ongoing teacher involvement and feedback• Low stakes component of a wider assessment scheme which allows for assessment of the individual |
|--|

It was felt that the learning outcomes would go a long way to determining the appropriacy of the group grade approach to collaborative assignments. Gonzalo accepted that “in the real world, that the outcome is the most important thing” but went on to say in the academic context that is not the case because if collaboration is a learning outcome “we need to be able to, to actually show that that's happening” (Interview Transcripts, p. 153). For Olga, if collaboration was a learning outcome and the

course was skills focused rather than language focused, then the adoption of group grading was more acceptable (Interview Transcripts, p. 256). This was particularly the case if there were opportunity for the teacher to follow the process, providing feedback throughout. The process focus, according to Ava, Karen, Fay and Cristobal, increased the perceived fairness of the group grade approach as it gave the students the opportunity to develop a quality artefact and they would learn through the process and feedback. It is important to note, however, that Ava would “really, really hate the model” (Interview Transcripts, p. 34) if assessment were just focused on the final product.

Lorenzo highlighted the importance of the grading scheme being transparent from the outset saying, “I think it's fair if the students know going in into the assignment, how it's going to work” (Interview Transcripts, p. 244). The cohort of students collaborative assignments were being used with could also have an impact on the suitability of adopting a group grade assessment scheme. Fernand considered “their age, what stage of the university system are they in” (Interview Transcripts, p. 135) as relevant factors, with Lorenzo adding “maybe the type of course is important to consider, and the the future, the the field of study” (Interview Transcripts, p. 244). Thinking outside of the immediate assignment, Van suggested using collaborative assignments as one component in a wider curriculum so as not to base high stakes decisions on the outcome (Interview Transcripts, p. 274).

3.3.1.4 General summary of Model A evaluation

Within the EAP context, the group grade approach is problematic. While there are arguments in favour of such an approach, realworldness principle amongst them, the group grade approach lacks differentiation between students and is also does little to discourage asymmetric collaboration. The prime focus on the product also works against EAP’s developmental function, since part of the aim of such an assignment would be to help students collaborate more effectively in the future. Furthermore, the quality of the final product does not demonstrate the quality of the collaboration, as Boud and Bearman (2022) would attest. A dysfunctional collaboration where one student dominated could result in a satisfactory end product while a poor piece of work could be the product of a group

misinterpreting the task, despite collaborating effectively. A student may experience a smooth collaboration and work in a group that produces a good quality product but they might gain no real insight into collaboration or themselves as a collaborator. In contrast, another student may have a more challenging and demanding experience, leading to a flawed piece of final work, but they might learn a lot more about themselves and collaboration along the way. As such, in courses where assessment has a diagnostic or developmental function, or is used to evaluate individual students' discrete language or academic skills, the group grade approach is not appropriate. Where it could have a role is when collaboration is a secondary learning outcome and the assignment is part of a broader scheme of assessment.

3.3.2 Model B: Group grade + complementary individual assignments

In Model B, all the students in the group receive the same grade, based on evaluation of the final essay, along with individual grades for complementary assignments, meaning assignments that help with the main collaborative task. In the case of an essay, this could be a summary of one of the sources used or reflection on the essay development process. Model B was the most warmly received by the participants, with an average star rating of 4.25 and no negative responses (see Table 3.4).

Table 3.4 *Star rating for Model B*

★	★★	★★★	★★★★	★★★★★
0	0	1	10	5

The model was most highly rated by Jerry, Olga, Ava, Van and Fay. Olga appreciated the opportunity to see individual work and felt the assessment scheme was easy to justify to students. Ava felt the group grade aspect would encourage collaboration as “students have motivation to to work on the collaborative assignment together to offer a really polished product and work with one another

because they have an element of working together and sharing the same grade” but that the individual element “merits those who perhaps work a little bit harder” (Interview transcripts, p. 49).

Fay also felt there was a value to the individual element as it “can ensure that everyone is doing some work” (p. 120). Van appreciated the individual work for a slightly different reason, as it could help identify students who may need additional support. The lowest rating of three stars was given by Fernand. He still expressed concern around the group grade aspect and the potential of that to artificially inflate some students’ results but did appreciate the opportunity for giving individual grades and, in an echo of Ava, rewarding students for how hard they worked. He also felt this was similar to the assessment approach he was familiar with and it was manageable for the tutors.

3.3.2.1 *Model B: Positive reactions*

One of the benefits of Model B was that it was “clear to follow” (Interview Transcripts, p. 257) and “more comprehensive and would therefore address the shortcomings of Model A” (Pre-interview Tasks, p. 34). As Cristobal noted, “I like the way it brings in other skills, you know, like, it actually, it makes the course bigger, because it brings in other skills, like writing a summary and, and a reflection” (Interview Transcripts, p. 87). The group grade element of Model B was valued as it helps promote collaboration amongst students with Bertha suggesting, “the collaborative task that they’re all getting the same grade for. So then that’s, that is encouraging the students to work together on that, rather than putting bits of things of their own together” (Interview Transcripts, p. 64). The addition of the individual element was seen as something that worked effectively with the group element as it could increase student motivation (Pre-interview Tasks, p. 13) and incentivize students (Pre-interview Tasks, p. 27).

The main appeal of Model B lay in the individual component, providing the opportunity for differentiation between students and the opportunity to see individual work that students produced. As Humberto put it, “students have to feel that they are getting value for money. And this has more credibility perhaps to the students, because they’re seeing that their own efforts are getting a bit more

emphasis" (Interview Transcripts, p. 167). This resonates with Karen's point that students "want to be *noticed* for what they've contributed" (Interview Transcripts, p. 192, emphasis in original). A number of participants appreciated that the individual component provided "room for each student to kind of show you know, here's also what I can do" (Interview Transcripts, p. 35-36). As Dolly put it, students "can still work harder to sort of, you know, get a higher grade" (Interview Transcripts, p. 106), which fulfils the need Laura saw for, "room for students where they feel comfortable and where they can show maybe their real performance" (Interview Transcripts, p. 226). For Tanya, this performance in the individual component could give the teacher insight into the collaborative product since the teacher gains

the discernment that you need that when you look at the collaborative assignment, you can kind of tell Okay, you know, you know clearly this is, not that you can tell individual voices, but you'd be able to say I can see a little more of Student A here maybe than Student C. (Interview Transcripts, p. 36)

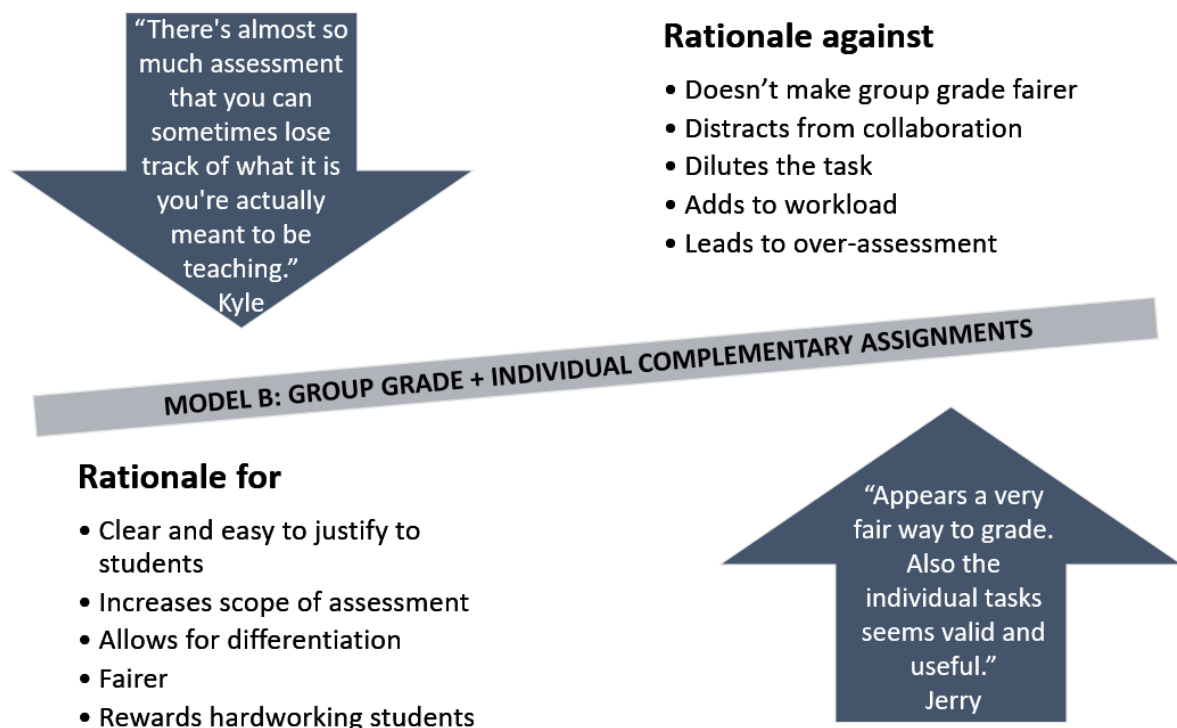
It is important to note, however, that she went on to caution against evaluating collaboration like this, feeling there had to be more insight, which could come through self- or peer-assessment which allows the teacher to "kind of a peek inside that black box a little bit better" (Interview Transcripts, p. 37).

Wendy valued the assessment being based on something concrete, the individual work produced by students, rather than "just a teacher's 'sense' of how much effort they've made" (Pre-interview Tasks, p. 68). As she expanded in her interview,

I think this has a lot more likelihood of being a fair reflection on what the students, it gives the teacher a better reflection of what the students can do or how much they've, how much work they've put in, probably you'll, you'll be able to see it from from the source summary and from the reflection or from whatever their individual work is. And then it's a lot less subjective, I think. (Interview Transcripts, p. 290)

Figure 3-2 below illustrates the main arguments for and against the group grade assessment scheme.

Figure 3-2 Summary of evaluation of Model B: Group grade + individual complementary assignments



3.3.2.2 Model B: Negative reactions

Although Model B was generally well received, it was not without criticism. While Kyle acknowledged there was still a simplicity to Model B, he felt that it was “more difficult to manage from an operational point of view” (Interview Transcripts, p. 208) in comparison with Model A. He also had a concern about over-assessment, explaining “at a certain point, there's almost so much assessment

that you can sometimes lose track of what it is you're actually meant to be teaching" (Interview Transcripts, p. 208). Lorenzo also questioned the number of assessments "if you're giving them one grade for multiple tasks" (Interview Transcripts, p. 245) and Andrea disliked bringing in other tasks which assess "something different other than the assignment itself" (Interview Transcripts, p. 7). Related to this, Lorenzo also shared a concern expressed by Bertha and Kyle, considering excessive tutor workload with additional items to mark. In addition to the grading, Bertha also voiced concern about having enough time to prepare students for the individual tasks (Interview Transcripts, p. 64). Fernand also worried about workload, from both the tutor and student perspective, reflecting on experience with a similar assessment scheme where,

we were getting overloaded, I think they'd completely lost the plot here with the amount of forms, you know forms and stuff they were getting students to submit to, and they were getting burnt out and fed up with it, and we were kinda getting fed up with it. (Interview Transcripts, p. 136)

Bertha raised a different concern, questioning whether "is it maybe kind of contradictory signals that you're asking people to work on something together and then in the second breath saying, but you'll get marked individually on bits of it?" (Interview Transcripts, p. 59), a concern shared by Meijer et al. (2020) and Barry, who felt "as a method to foster collaborative frameworks, I don't see a huge degree of benefit there" (Interview Transcripts, p. 15).

3.3.2.3 Suggestions for implementing Model B

Model B was the most familiar to participants but implementation would still come with some caveats, as summarised in Table 3.5.

Table 3.5 *Suggestions to implement Model B*

- All assignment components must lead into learning aims
- Ensure individual tasks are supportive of the main assignment
- Weight assessment toward the collaborative output
- Make sure there is time for relevant learner training and the assessment load

The implementation of Model B would again be dependent on the learning aims for the course. If, for example, reflection were not a learning outcome, then reflection would not be an appropriate individual element in assessment. So too must the individual elements be genuinely supportive of the main collaborative assignment, meaning they help the students produce the final collaborative work or the process of collaborating itself. The supporting role of the individual elements should also be reflected in the grading scheme. The proposed Model B had a 50/50 weighting between the collaborative component and the individual component(s). This was felt to be imbalanced by Karen, Kyle and Barry, all of whom would like to see more weight given to the collaborative component given that it was a collaborative assignment. As Barry explained, “complementary assignments are sort of the subordinate part of that, normally a 30%-40%” (Interview Transcripts, p. 15), which is in line with the weighting favoured by Kyle and Karen as well. It would also be necessary from the outset to ensure there is time enough in the course for any required learner training to ensure they can complete the individual components successfully.

3.3.2.4 *General summary of Model B evaluation*

Model B was the most warmly received of the four models. While it was appreciated that there were issues with it, principally around task dilution and the potential for over-assessment, where it gained most favour was in the opportunity to differentiate between students. It was acknowledged that the individual assignments did not really give any real insight into the collaborative process, nor

did it mitigate the problems of the group grade approach, but it was felt that through the individual assignments a teacher could see what each student is capable of. The individual assignments ensure that the students have to be doing something for the assignment, which does not negate the free-rider phenomena but the participants seemed more comfortable nonetheless. Having access to students’ individual work provided the opportunity to reward high performing students and identify students who may be in need of further assistance.

3.3.3 Mode C: Group grade + individual collaboration grade

In Model C, all the students in the group receive the same grade, based on evaluation of the final essay. They each also receive an individual grade for their collaboration. Model C was the only model not to receive a five-star rating (see Table 3.6), but the average star rating was 3.00, making it the second-best rated model behind Model B.

Table 3.6 Star rating for Model C

★	★★	★★★	★★★★	★★★★★
2	1	8	5	0

While no participants awarded Model C five stars, five of the participants gave it a four-star rating, albeit with considerable qualification. Lorenzo liked there was an element of individual consideration in the assessment scheme but also considered it unrealistic, as there is not that kind of individual consideration in collaborations in the workplace. Bertha advocated for the evaluation of the collaboration to be done through peer-assessment rather than teacher-assessment, questioning whether the teacher should “assess the collaboration if they weren't actively involved in the collaboration” (Interview Transcripts, p. 64-65). Her previous experience with peer-assessment had shown that it could be effective and was a better measure of collaboration than teacher-assessment, which could be quite arbitrary. Humberto, Kyle and Gonzalo agreed that Model C was potentially a

good model for assessing collaboration, although they also questioned *how* collaboration could be assessed. As Humberto put it, “if you can do it, then it's good. But I'm just not sure how you would do it” (Interview Transcripts p. 167).

Both Cristobal and Olga gave it the lowest possible one-star rating. In his pre-interview task, Cristobal claimed, “it would be too difficult to assess collaboration - the teacher could not do it and the students could be too biased!” (Pre-interview Tasks, p. 78). He expanded in his interview, discussing the difficulty of knowing what someone had contributed to a collaboration and how this was complicated by the various ways in which people could contribute, explaining,

It's all different elements, isn't it? Some could have done more research. One person might be better at reading and research and they've contributed it. Like another one might be ideas, like you say, someone else might have got them all going and said, "Look, Let's meet at three o'clock" and organised it. (Interview Transcripts, p. 88)

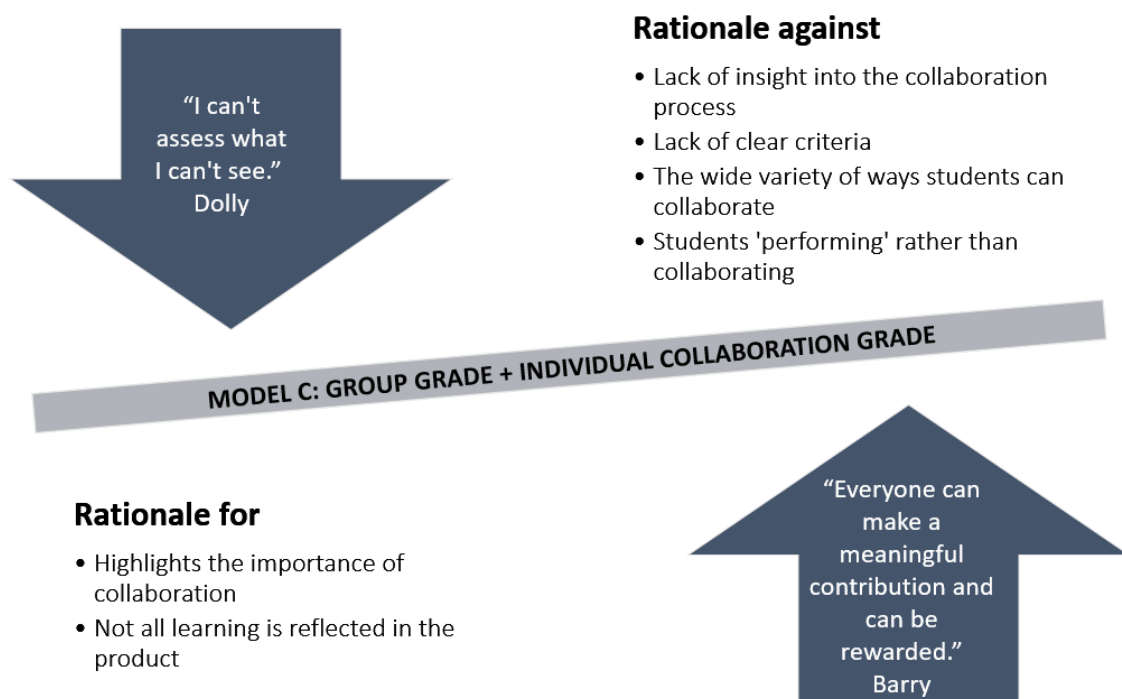
Even though in terms of facilitation he was probably the most deeply embedded of the participants in the student groups, taking an active role in their shared work documents, he still felt there was too much he did not see or did not know to be able to reliably assess collaboration. He also expressed concern that any mechanisms to try to track contribution could just become an exercise in box ticking with

the teachers doing all this admin having to collect all these different things to confirm whether students have done it or not, you know, and it's just now Oh, my God, you know, it's all this extra work, which just brings us back to this simple there's the product, mark, the damn product, mark the video, is it good?” (Interview Transcripts, p. 89)

For Olga, the struggle was with defining collaboration for assessment purposes, drawing attention to the varied ways students could collaborate, questioning “do you define, you know, in terms of how much percentage of the writing they do, would that count as a collaboration or the background research, should that count as collaboration?” (Interview Transcripts, p. 257).

Figure 3-3 below illustrates the main arguments for and against the group grade plus individual collaboration grade assessment scheme.

Figure 3-3 Summary of evaluation of Model C: Group grade + individual collaboration grade



3.3.3.1 Model C: Positive reactions

One of the positives about Model C was grading the collaboration, in the sense that grading collaboration “gives a message to students that, you know, your collaboration matters” (Interview Transcripts, p. 226). As Dolly put it, “I still want to reward collaboration, because I'm a believer in collaboration” (Interview Transcripts, p. 108-109). Humberto also felt that collaboration should be recognised and “being able to help a team be a team and do the collaboration. That must be given

some credit” (Interview Transcripts, p. 168, emphasis in original). However, he added the proviso that it was learning outcome dependent and while assessing collaboration in a communication course could be valid, it would be less so in a dedicated writing course. Cristobal also made the connection back to learning outcomes and felt that if collaboration was an outcome then it should be assessed. An interesting development during the discussion was that he was unsure on how to assess collaboration but liked the concept of “assessing awareness of collaboration, or, you know, rather than collaborative skills” (Interview Transcripts, p. 90). This arose through discussing the possible learning that could take place during a problematic collaboration wherein, “their product was so bad, but at least they can be aware of where it went wrong, and why it went wrong and therefore they're learning from that” (Interview Transcripts, p. 91).

The role of technology was also discussed in relation to giving the tutor insight into the collaboration process. Fernand discussed using OneNote, thinking “you could actually probably get a good feel for what was their level of collaboration” (Interview Transcripts, p. 136). As will be seen in the next section, the lack of clear criteria for assessing collaboration was a major stumbling block, but one which Karen suggested could be overcome by engaging students in the process of developing evaluation criteria for it, finding out “what's good performance for *them* in a team?” (Interview Transcripts, p. 193, emphasis in original).

3.3.3.2 *Model C: Negative reactions*

The biggest issue with Model C was the question of how to assess collaboration, which was raised by all participants, and this was encapsulated in Wendy's pre-interview task, where she wrote,

Assessing collaboration is quite subjective. How can the teacher know how much each person put into the groupwork? So much of it happens outside the classroom, and manifests itself in ways the teacher can't access (i.e. replying to messages, being open to other's ideas, easy to work with etc). (Pre-interview Tasks, p. 68)

Knowing what to look for and what to assess was a major concern. When discussing student contributions, the participants considered activities related to the task itself, in terms of research, writing and editing, but also broader issues like organising and motivating the team. As Ava explained,

for an individual to contribute to a group, there are just simply so many different ways. So it might got this stud-, this person contributes to ideas, to research, to language, to simply getting everybody on time, or be the person that really balances out the dynamic of the group. (Interview Transcripts, p. 50)

Kyle also reflected the issue that students in different cultural contexts may collaborate in “a different way that we weren't necessarily recognising” (Interview Transcripts, p. 209), which in turn brings in the question of intercultural competence and how students with different backgrounds could navigate collaboration in ways unfamiliar to each other or the teacher. Our recognition of what good collaboration looks like was also discussed by Tanya, who said,

I fundamentally believe that our students are unknowable (...) I see a lot of room for just kind of assuming a lot of things that that may or may not be true, based on all kinds of things, I just think it's problematic. I mean, you know, I like Student A a lot, Student A comes to class on time Student A dresses neatly and pays attention in ways that are more visible to me than Student C. I mean, that alone could account for me getting Student A a 90% when Student C could be doing all the same things is also on time but maybe sits in the back presents in a different way that doesn't match my personal idea of what a paying attention student look like, looks like, you know what I mean?” (Interview Transcripts, p. 36-37)

The lack of visibility into collaboration was a common theme, with Dolly succinctly explaining, “I can't assess what I can't see” (Interview Transcripts, p. 108). Part of this is because a lot of

collaboration could happen outside the classroom, or through communication channels that the tutor is not privy to. Fay was very aware that she

may not see everything is going on in all groups all the time as to give a fair assessment of what was going on. As I said, unless I had cameras going on and all the groups all the time, and took the time to watch all these films and analyse them and then give grades then, which again, I don't want to do. (Interview Transcripts, p. 121)

Laura, in contrast, felt more positively about utilising technology, not to necessarily get insight into collaboration but more as a means so students would know such insight was possible. As she explained

I just have this add-on tool Revision History Analytics, just I follow who is doing what, who is contributing to what extent. I let the students know that everything is digitally followed. Not to play the Big Brother, but just to see if each and every student is taking his or her own responsibility in the best way possible. (Interview Transcripts, p.222)

Considering the role of technology, Karen was not so sure about understanding a collaboration in online environments, feeling it was easier in the face-to-face context where “I can see, I can observe the levels of interaction between people. I can assess their, their body language” (Interview Transcripts, p. 193), a point of view echoed by Bertha, who also felt classroom observation gave richer information. Dolly also questioned the insights gained through technology, noting, “I feel like there are like, sufficient enough gaps where I feel, yeah, the online collaborative document is only one piece of evidence” (Interview Transcripts, p. 108). The example she gave was of the research process, which could be time consuming but “none of that time is credited or documented” (Interview Transcripts, p. 108), meaning it is invisible in the collaborative workspace. Cristobal agreed that online collaborative

documents only provided limited insight as students could be “communicating with each other on WhatsApp, or, or Zoom or phone calls” (Interview transcripts, p.88).

There was discussion of ways of making the collaboration more visible, through channels like tutor-group meetings, diarying or memoing, or use of checklists. The immediate concern through utilising these kinds of tools went back to Cristobal’s point about box ticking, with Fernand also explaining the negative impact of being overloaded with admin when having to track student work and Dolly speaking of the tediousness of having to go through students’ documenting of what they had done. Van felt that asking for documentation could make the process artificial, comparing the reality of collaborating with the assessment of it.

Because if you want to collaborate, you know, we know what collaboration is, don't we? We, we look at each other, whether it's online or face to face, and we smile, and we say, and we make a decision. That's what collaboration is. So we just do it. We don't write about it, I mean, most people, they don't write about it. And also, you don't have somebody standing, sitting there giving you a score for it. So both of those things, make it artificial. (Interview Transcripts, p. 277)

He expanded further on the issue of producing things just because they have to be produced, thinking of the potential washback where, “You get your students producing all these behaviours, not because it's, they're producing it because they want to get the grades for collaboration, not because they think it *is* valuable collaboration” (Interview Transcripts, p. 277, emphasis in original). This sense of students doing things just because they were expected was also discussed by Dolly, who talked about students who “perform for you right in class while you are observing them collaborating. But then, you know, it's another it's another situation altogether when they're not in class” (Interview Transcripts, p. 107).

There were also concerns about three potential avenues for looking into the collaboration: reflection, diarying/memoing and peer-assessment. While reflection was felt to be important and valuable for students, its use of assessment purposes was questioned. Lorenzo neatly summarised the one of the issues, considering, "Are you grading the student on what they wrote they did or what they actually did?" (Interview Transcripts, p. 247). Wendy expanded on that concern, saying

I think what it shows you is what the student thinks has happened. It doesn't show you what's actually happened. And and again, then, so let's imagine you have student A and Student B, Student C, Student A and B didn't get on, they had a falling out. Um, one of them writes about it in their reflection. One of them doesn't, and, and then what do you, are you pitching them against each other to see if then Student C's account corroborates Student A's a count of what happens? I don't know how you can, how you can really know what's gone on. And, and I'm not sure it's ethical to mark one person's collaboration on another person's account of their collaboration. Because that might not really be the way it was, or there might be other things getting in the way of how they talk about it. (Interview Transcripts, p. 291)

Van was also dismissive of reflection, as it was not something he personally found useful, explaining,

But I could be in a team where I collaborate really well, or a team where I collaborate really badly. And in neither of them would I write a reflective diary. So to require me to do that it's, your requiring, of me, it's something artificial. There may be, in the, over the semester, if I spend six hours writing those reflective diaries, maybe in those six hours, maybe this 30 minutes or 20 minutes, where it's actually where I start to think, Oh, yeah, and the collaboration kind of deepens for me or something. But most of the time, I'll just be thinking, "Urgh, I have to write this stuff down," you know, so it's just a waste of my time" (Interview Transcripts, p. 277)

Bertha also questioned the use of peer-assessment since, "it could be quite arbitrary that someone just says, "Oh yeah, the other two did a good job, so we'll give them both X points" if they're doing a peer feedback thing" (Interview Transcripts, p. 64). Students' capacity to peer-assess was also discussed by Dolly, who felt it could be problematic, even with learner training.

I feel like that is too high stakes. The students will talk to each other, and they will not mark another student, they will not give a low mark. So even with that case that I told you were like three people were like, this student did not the fourth student didn't do his or her work. And then their grade for that student was 85. Because they just feel, they feel bad, right? To to give a low mark. And then we'll be like, "Well, you know, he's just a student, you know, he'll try, hopefully, we hope they will try harder next time. 85!". So, it's, it's, it takes training, right to like, I feel like the students, they, even if I ran like a norming session for them, and gave them descriptions of like, oh, if this student, you know, only shows up to like three or five meetings, then they can only get like, three points for participatory, or whatever. I think even if I train them on that they will in the end still feel bad. And, and would not, would not give, like an accurate assessment of the collaboration grade. (Interview Transcripts, p. 107)

Another concern was that of appropriacy of devoting time and resources to the assessment of collaboration. Ava recognised the amount of work that may go into assessing collaboration and questioned whether it was appropriate for an EAP course to devote so much time to collaboration when, ultimately, EAP delivers language courses. This was a point of concern also raised by Cristobal who wondered "should that be a different a different type, of course, for a student, you know? Is that something that needs to be in an EAP course? Or is that more of a programme level issue?" (Interview Transcripts p. 151) when discussing all that would be needed to support assessing collaboration. The resources used in assessing collaboration was also a concern for Van, although for an alternative reason. He held that the quality of the collaboration would be reflected in the quality of the final work

so any attempt to assess collaboration would be double scoring and unnecessary duplication (Interview Transcripts, p. 276).

3.3.3.3 *Suggestions for implementing Model C*

Model C is an approach that the participants seemed to want to work, even though the assessment of collaboration itself is problematic. In order to implement Model C, certain steps could be taken, as outlined in Table 3.7 below.

Table 3.7 *Suggestions to implement Model C*

- Ensure collaboration is a key learning outcome of the course
- Have a clear conception of how you expect students to collaborate
- Make those expectations clear, with explicit discussion of collaboration at the outset
- Have assessment criteria, which are clear to the students (or which are developed in discussion with the students)
- Award a minor proportion of the grade to collaboration.

Successful implementation of a collaboration grade would be dependent on having a clear conception of what collaboration is from the outset and ensuring that is transparent to students. A marking scheme would have to be developed to evaluate performance. Some participants favoured this being on concrete actions or behaviours, with Ava explaining how she had used a peer evaluation sheet where students' can comment on their group mates on, "for example, the student always participate in meetings, or the student is never late for any of the deadlines. So they're evaluating something that that is everybody can see" (Interview Transcripts, p. 51). Careful consideration would also have to go into the potential washback from criteria so that they would promote positive learning outcomes. Ultimately, the evaluation of collaboration is going to be a compromise, as Dolly explained,

I can't assess what I can't see. So the only thing I can control is to sort of reduce the percentage of the collaboration grade. But I still want to reward collaboration, because I'm a believer in collaboration. But I, I understand that, you know, students, what they care about the most is their grade. Right? And so I have to make a compromise with myself and with them. (Interview Transcripts, p. 108-109)

3.3.3.4 *General summary of Model C evaluation*

There was a clear mismatch between the rating of Model C and the discussion around it. There was a recognition that if collaboration is a learning outcome or if the teacher wanted students to value the collaboration, it should be assessed. Conceptually, participants valued the idea of assessing collaboration but ran into problems when it came to envisaging how that assessment could take place, as illustrated by Humberto who acknowledged that assessing collaboration, “ideally would be a good way of doing it, a good way of assessing, I'm just not sure about the practicality” (Interview Transcripts, p. 168). This issue of practicality also resonated with Karen, who said, “For me, at least B and C, were the best in terms of equity in terms of fairness. But also, it's about what's practical. And what's possible in terms of what we could do” (Interview Transcripts, p. 193). For Lorenzo, the practicality lay in how applicable assessing collaboration would be outside of the academy as he thought assessing collaboration “maybe more fair to the students and to consider their individual circumstances. But at the same time, it's it's a little bit less applicable and realistic to their future professional careers” (Interview Transcripts, p.246). This pull between fairness and practicality will be discussed further when exploring the key principles guiding teachers through the assessment of collaborative assignments.

It is indicative of how complex assessing collaboration is that across all the interviews with experienced, well-qualified EAP professionals, some of whom strongly advocated for the use of collaboration in EAP programmes, none of the participants had confidence in how they would go about assessing a collaborative assignment. This was due in part to the varied ways in which students

could contribute to a collaboration, compounded by how opaque some of those ways of collaborating could be to the teacher. The majority of participants felt clear rubrics, criteria or metrics would be needed to justify an assessment of collaboration, albeit with some concern as to how effective those measurement tools may be. The need for a fair and practical way to assess collaboration was highlighted by Gonzalo, who ended the interview saying, “I think if you come up with a way to assess collaboration effectively, you could make millions” (Interview Transcripts, p. 155).

3.3.4 Model D: Group grade x weighted collaboration grade

In Model D, students receive a grade determined by the grade for the final essay, adjusted according to their weighted collaboration grade, meaning a student who had collaborated most effectively would get a higher grade than a student who had collaborated less effectively. Model D also received star ratings across the board, with the lowest average star rating of 2.81 (see Table 3.8).

Table 3.8 *Star rating for Model D*

★	★★	★★★	★★★★	★★★★★
2	5	5	2	2

Humberto and Kyle awarded Model D the highest five stars. During the interview, Humberto pre-faced his rationale by saying,

Yeah. If you can do it, if you could pull it off, then yeah, it's, it's excellent. But again, I have my doubts about how practical it is. But, you know, I'm open to suggestions. You know, if you can point me to something that would clearly enable me to do this without bursting some blood vessels, then I'm fine with it. (Interview Transcripts p. 169)

He went on to explain that he favoured the model because “It's meaningful to the students, they can see that *their* effort has been rewarded” (Interview Transcripts, p. 167). Kyle's justification took a different angle, appreciating how the weighted collaboration grade took out any direct individual element to the assessment, which could compromise the group element. As he put it, “Well, it's it nicely kind of solves the problem of the individual and group thing whereby if we overemphasise the individual contribution, to what extent does it actually, is it actually a group project, right?” (Interview Transcripts, p. 210).

In contrast, Karen and Wendy rated this model with the minimum one star. For Karen it was the complexity of the model, joking, “This one gave me a headache. I stopped, I'm thinking I'm just a qualitative researcher” (Interview Transcripts, p. 194). When expanding on this she focused on how the relative contribution of students could be weighted, questioning

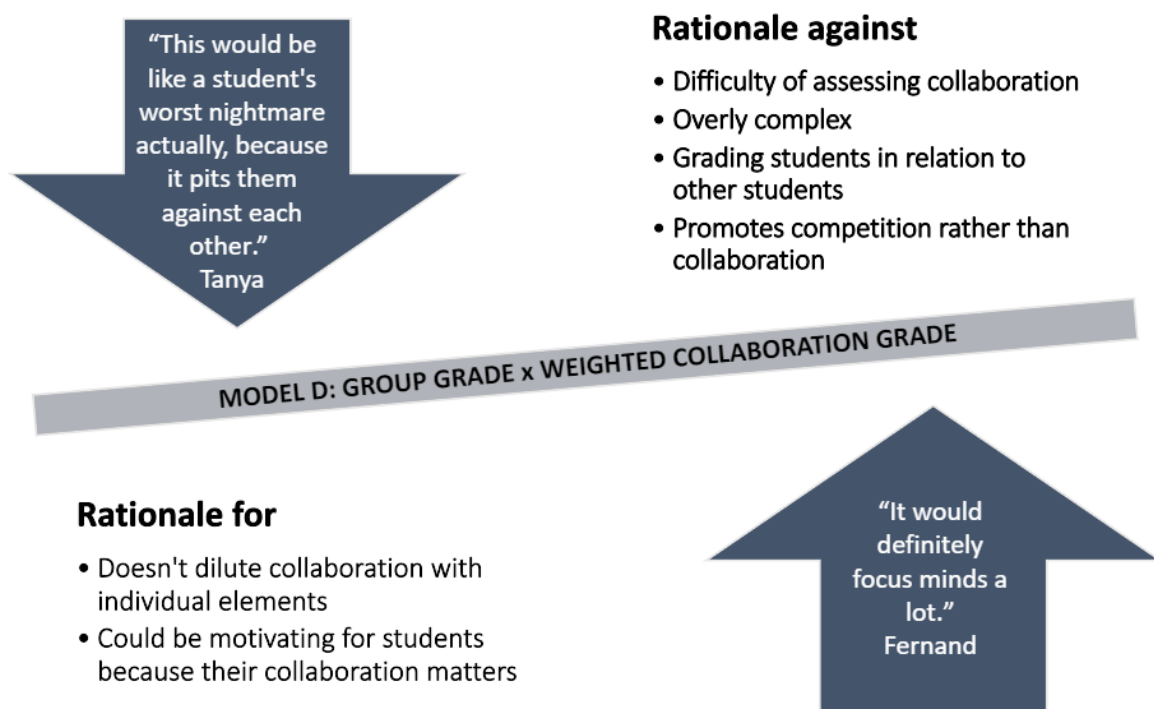
So again, what is above average? What's average? what's below average? See, I need to see those things first. And again, how can we clearly know without observation, without clear tracking of some kind, that Student A has made that above average, and Student B has just made an average and Student C has made a below average contribution?” (Interview Transcripts, p. 194)

Wendy agreed that, “It's the same as C really - how does the teacher know who's really average/above/below? I think the weighting aspect makes it appear overly complicated” (Pre-Interview Tasks, p. 69). She also saw it having an impact on how students could work together, considering a situation where,

Student C says, “I worked harder than Student B, I should have got the middle grade,” and it just pits them against each other, rather than, surely with something that's collaborative, it should be encouraging them to work together, not in competition. (Interview Transcripts, p. 293)

Figure 3-4 below illustrates the main arguments for and against the group grade x weighted contribution grade.

Figure 3-4 Summary of evaluation of Model D: Group grade x weighted collaboration grade



3.3.4.1 Model D: Positive reactions

Fernand rated Model D four stars, one higher than Model C, and his contention was that Model D was fairer and more motivating for students. Where a student could potential free-ride through Model C and pass the assignment on the quality of the work of the other students in the final essay, in Model D “you can't lie back and hope this passes because if I see lying back back, I'm going to award

your marks, I'll fail if it comes down to it, basically" (Interview Transcripts, p. 137). He did concede it could also generate competition as students tried to get the highest mark for themselves and proposed that,

It needs to be maybe sold to them that everybody, collaboration is not about everybody being great and doing the same thing. It's about everybody using the best of their ability to get the good, get as good a result as possible at the end of it. So it might take a little bit of explanation. (Interview Transcripts, p. 137)

Laura also mentioned the possibility of the approach generating competition, but thought it could be positive competition, with everyone doing their best work to help the group get the best grade. She talked at length about high achievers taking over collaboration and who "break the group dynamics" (Interview Transcripts, p. 227) and it seems the argument was that knowing the final grade would not just be based on the quality of the work but on how well they collaborated would encourage high achievers to collaborate more effectively. This was something Dorian also considered as a possible outcome, explaining it would instil a sense of responsibility to try to do away with the issues of some people contributing more than others" (Interview Transcripts, p. 26). However, this was qualified with the other consideration that Model D,

is kind of a shaming technique, isn't it really, in the sense that everything's then adjusted according, according to how much other people have contributed and I think, you know, on one on one level it could be a little daunting for some students. (Interview Transcripts, p. 26)

3.3.4.2 *Model D: Negative reactions*

Model D carried all of the concerns about Model C and the complexity of assessing collaboration, with further issues. Although Kyle rated Mode D highly, his support was qualified, with consideration of the complexity of implementing the model and of justifying it. As he said,

It's just the sort of complexity of working out what the weighted contribution would be, and how that's going to be reflected in spreadsheets and exam boards. Students sort of appeals and stuff ... You can almost see the student kind of questions kind of coming back if they were presenting work. (Interview Transcripts, p. 210)

The complexity of the model was a common concern, both in terms of determining students' relative collaboration levels and of applying that to the group grade. The way the group grade would be allocated was a major concern because, as Tanya explained, "if this person gets more then that person, by definition, gets less" (Interview Transcripts, p. 37). Barry also recognised this and felt it ran counter to collaboration itself explaining, "an above average contribution pushes a below average contribution out and I don't see that as being real world or a fair representation of collaboration. I think collaboration can push the entire group higher" (Interview Transcripts, p. 16). The sense of this model putting students in competition with each other was also shared by Wendy, Laura, and Van, with Laura explaining, "This model may push some high achievers to be more ambitious and this may have a negative impact on the group dynamics" (Pre-interview Transcripts, p. 119). Lorenzo illustrated this when imagining

you have maybe a student who is very outgoing and, you know, contacts you and says, "Oh, I worked so hard on this, I tried to work hard, but the other students didn't want to work with me," you could sort of you can sort of push the other students under the bus, you know, to try to increase your own score. (Interview Transcripts, p. 247)

For Ava, the competition could come out in the situation where a student could get unlucky and be in a highly motivated group where, “just because they have more competitive peers, they're really at a lower rate, which is not fair” (Interview Transcripts, p. 51). The impact of the group a student was in was also a concern for Van, who thought,

It seems to have the same disadvantages as Model C, but the additional threat to validity of the score being norm-referenced *within the group*. If student A and student X have the same collaborative ability, but A is put with two grouches and X is put with two great teamworkers, then X will get a lower grade than A but through no fault of her own. (Pre-interview Tasks, p. 62)

3.3.4.3 *Suggestions for implementing Model D*

There was not a great appetite to see Model D in action, although two important considerations stood out (see Table 3.9).

Table 3.9 *Suggestions to implement Model D*

- Have clear criteria to determine what constitutes “Below average”, “Average”, and “Above average”
- Explain the assessment scheme to students carefully, ensuring they understand it should engender collaboration rather than competition

Bertha felt Model D had potential, in the same way as Model C, if peer grading were implemented and there was time for learner training to support it, explaining “Sts would need a set of parameters to determine what the collaboration involves, beyond just the research and writing,” (Pre-interview Tasks, p. 14). She expanded on this in her interview, but also voiced concern over the impact it could have on students, saying,

So I suppose it needs to have some sort of like, kind of like a checklist or something that they can rate themselves against of what a good group member is, what a good group member does, what, like, what's, what is an average, they could define in those terms, what's above average, what's average, what's below average, maybe having some kind of like checklist, and maybe the terminology of like "below average" I don't know, maybe if a student got that maybe they would, like, how would they feel by being labelled. I don't know, I'm just always concerned about what the students are left feeling after these kind of tasks. (Interview Transcripts, p. 65)

3.3.4.4 *General summary of Model D evaluation*

Model D was the least popular model, with the two biggest issues determining the quality of a student's collaboration and considering student performance in relation to each other. It seemed as if the participants felt they would need an even clearer view of the collaboration than in Model C if they had to determine the relative contribution of group members. The general consensus also seemed to be that Model D could actively undermine collaboration by pitting the students against each other. As Van put it, Model D

has all the problems of C that I mentioned, but even worse, because they're explicitly competing within the groups. It's not even, it's not even normed across the class, It's normed within, within the group. So I'm literally competing against these people I'm supposed to be collaborating with. It's, I mean, it's not gonna work out. Either they actually collaborate and forget about the grade or they sneakily try to do fake collaboration, "I'm gonna out fake collaborate you," and they fake collaborate you and so on. (Interview Transcripts, p. 278)

3.4 Participants' proposals of a model of assessment

In the early iterations of the interview, after discussion of the four models of assessment, participants were asked which assessment lenses they would use to assess a collaborative assignment. This was with a view to asking participants to give a weighting to each component to construct an assessment scheme for collaborative assignments. The assumption was that they would produce a scheme broadly similar to the one outlined in the motivation study (see Appendix 8), incorporating assessing the final work, individual complementary assignments, and the collaboration itself. However, during the pilot interviews it became clear that expecting participants to spontaneously provide a well-considered assessment scheme with assessment lenses and weighting decisions was unrealistic. This is not least because it quickly became clear that for the participants any assessment scheme would be dependent on the learning aims of the course in question, meaning there is no single answer to the question. Consequently, the focus of that question area shifted more toward whether collaboration should be assessed.

When discussing how to assess collaborative assignments, only two participants did not favour assessing the collaboration. Tanya felt students in her context could be relied upon to collaborate so there was no need to assess it directly. Rather than collaboration, she would assess a collaborative assignment through the final work and individual complementary assignments. Van was the most resolutely against assessing collaboration as he believed that is already done implicitly through assessing the work the group produces. For Van, the product was the only assessment lens needed. For all the other participants, the feeling was collaboration should be assessed in collaborative assignments if it is a learning outcome. Further to this, ten of the participants favoured an assessment scheme comprising the three components of the final work, the collaboration itself and individual complementary assignments, with that individual assignment likely some kind of reflection. There were caveats (that the lenses selected and the weightings were outcome dependent (Andrea, Dorian, and Fernand), that assessing collaboration was desirable if possible (Humberto and Lorenzo), that collaboration and individual complementary assignments would necessarily be the minor part of the

assessment (Olga and Cristobal), that our scope for successfully assessing collaboration is limited (Humberto)) but this does indicate both a need for collaboration to be assessed and an acceptance of multi-lens assessment for collaborative assignments.

3.5 Key principles and guiding orientations

It is apparent from the evaluation of the four models of assessment, and wider discussion about assessment of collaborative assignments which followed, that they pose particular challenges wherein traditional conceptions of assessment may need to be modified, as illustrated by consideration of general principles of assessment below.

It is important to note that the concerns of EAP practitioners, albeit their being conversant with common principles of assessment, are not automatically those identified by Hamp-Lyons of “reliability, validity, practicality and (sometimes) consequences” (2009, p. 256). Only one participant, Andrea, mentioned reliability directly. Humberto spoke directly to validity, when arguing that that evaluation of a collaborative assessment had to be focused on the collaborative product as, “in order to have validity, I believe that the technique has to match the outcome,” (Interview Transcripts, p. 164), meaning he saw greater validity in students getting the same grade rather than trying to “deconstruct it at the individual level,” (Interview Transcripts, p. 164). Laura also referenced validity when discussing the possibility of assessing awareness of collaboration rather than collaboration itself, although she cautioned it should be “should be practised with smaller tasks with smaller gradings,” (Interview Transcripts, p. 232). During the course of the interview when discussing assessing collaboration Wendy realised “the more I think it's not a valid assessment tool,” (Interview Transcripts, p. 297). Barry and Humberto discussed the face validity of different approaches to assessment, both in terms of how they could be perceived by students and by wider stakeholders. ‘Consequences’ and ‘Practicality’, however, did resonate, and will be relevant when considering the impact of environment at practitioner beliefs and the practical orientation held by participants.

Rather than the traditional principles of assessment, EAP practitioners had a different focus when it came to collaborative assignments. As Kyle explained,

I think the nub of this is to find a way of being able to differentiate, focus on the collaborative process, but do it in such a way that you're not overburdening everyone with all the different assessment elements, and then losing sight of what the actual task is. (Interview Transcripts, p. 208)

This encapsulates an array of concerns voiced by EAP practitioners: needing to differentiate between students, needing to evaluate the process of collaboration, ensuring the multifaceted assessment is practically manageable for both teachers and students, and ensuring the main aims of the task are met, be that development of academic voice or competence in a particular academic genre. Through the interviews it became clear that there were three key principles and three guiding orientations, which were more prevalent than traditional large-scale assessment considerations and shaped practitioner beliefs about assessing collaborative assignments (see Table 3.10).

Table 3.10 *Key principles and guiding orientations for EAP practitioners assessing collaborative assignments*

Key Principles	Guiding Orientations
Simplicity Transparency Fairness	Reward Practical Defensive

In this context, ‘principle’ draws on the following definition: “a fundamental, primary, or general law or truth from which others are derived” (Dictionary.com, 2024). These principles inform beliefs and attitudes toward assessing collaborative assignments and, from these principles, orientations emerge. ‘Orientations’ follows the Merriam-Webster (2024) definition of “a usually general or lasting direction of thought, inclination, or interest” meaning orientations speak more towards action and behaviour. The nature of each principle and orientation will be discussed below.

3.5.1 Simplicity

One thing that was clear from the discussion of assessing collaboration is a desire for assessment to be simple. Looking retrospectively at the assessment scheme from the motivation study (see Appendix 8) it would seem to fall at the first simplicity hurdle since it had six discrete elements incorporating Moodle forums, individual complementary assignments, peer feedback, the processes of collaboration and the process of essay development on top of assessing the final essay itself. Certainly, for Andrea, the multiple components of the assessment would have been problematic as this approach “measures something different” (Interview Transcripts, p. 8) rather than the piece of collaborative work produced and that it also in turn “furthers the task” (Interview Transcripts, p. 9). This was not always seen as a problem, with Dolly, for example, seeing individual complementary assignments as an opportunity for learning, commenting

I like the way it brings in other skills, you know, like, it actually, it makes the course bigger, because it brings in other skills, like writing a summary and, and a reflection. So I just think it actually, you know, increases the scope of the, of the assignment as such, of the course and allows for individual effort and individual quantity to come through and be part of it as well (Interview Transcripts, p. 119)

and Fay considering it a “fabulous idea” (Interview Transcripts, p. 120) as a way of ensuring everyone is doing some work and involved in the assignment.

Barry considered trying to evaluate relative levels of contribution “a very complex model of assessment because of the very many factors that one would need to take into account” (Interview Transcripts, p. 44) and he admitted, “you can see how simple I like my models of assessment to be” (Interview Transcripts, p. 44). Ava shared this preference for simplicity, in part because any assessment scheme has to be explained and justified to both students and the institution, and also because assessing collaboration would take time and resources which it may not be appropriate for

an EAP course to devote to collaboration. Karen echoed concern about the amount of work involved in evaluating collaboration, thinking about putting Model D into practice and realising “Oh, this is too tricky for me” (Interview Transcripts, p. 194). Even Dolly, who appreciated task expansion in an assignment, thought “collaboration is much more of an abstract concept ... there are so many different sub skills” (Interview Transcripts, p. 163), which makes teaching and assessing collaboration overly complex.

For Van the issue of simplicity went beyond collaborative assignments to course content as a whole, as he lamented EAP courses “all become this kind of mush of critical thinking, blah, blah, that kind of stuff” (Secondary Interview Data, p. 273) rather than focusing the core content of more relevance and benefit to students. This overloading of courses was also a concern for Kyle, who said

It's a perennial issue in EAP in particular, isn't it? That the assessment becomes A) too much of it, and it overburdens both the people who work, well everyone, the course coordinators, the teaching, the people are who teaching it and the students themselves, and, and the students. And then at a certain point, there's almost so much assessment that you can sometimes lose track of what it is you're actually meant to be teaching. (Interview Transcripts, p. 208)

3.5.1.1 Product glare

The urge for simplicity is part of the motivation for focussing assessment on the product rather than the collaboration, with Bertha recognising an assessment scheme she had set up for a group presentation lacked any mechanism for looking beyond the final product: the presentation. Rather than looking into the development of the presentation and the collaboration, the focus was on the final presentation and individual presentation skills of the students because “that was the most simple way” (Interview Transcripts, p. 69). Kyle shared this practical view, considering a product focus “easy to operationalize” (Interview Transcripts, p. 208), although he preferred an assessment scheme that went beyond the product, despite it being more difficult. Cristobal also felt there was a simplicity in

looking at the product because “if they do the research, and they work together, and they write it, it will come out as a good product” (Interview Transcripts, p. 116-117). This point about good collaboration leading to a good final product was shared by others, suggesting product glare could be an issue, where a good product masks what goes on during the process of development. Ava contended that

the reality is, if a group cooperate with each other, they are in more communication. They consult more, they spend more time together. We can tell from their essays that it's more coherent, their presentation has a better flow. Or even they rehearsed together, it's a better presentation, eventually. So I don't feel it's necessarily that we have to grade collaboration and who does what, specifically, to understand whether students are doing collaboration or not. (Interview Transcripts, p. 53)

Humberto concurred, explaining, “And you know that collaboration has been effective because the deliverable is effective” (Interview Transcripts, p. 170), which is something Kyle would also endorse. He discussed the quality of the product reflecting the quality of the process, suggesting “we would assume a better end product means a more successful collaborative process if we set it up right.” (Interview Transcripts, p. 209). The latter clause in that comment is a very important caveat that will be returned to later. One of the consequences of product glare was highlighted by Laura, who felt it could make it more difficult for teachers to assess individual contribution to an assignment as

if the instructor will read a strong final product, he or she will not care about the individual contribution. And he or she will just focus on "Oh, that's focused on the paper, that's strong work." And she will mark it higher, being excited with the strong paper that she has just read, or he has just read. (Interview Transcripts, p. 227)

However, participants also conceded that a quality final product may mask as problematic collaboration. It also opened up a line of enquiry about the learning that can take place through a problematic collaboration. Such a collaboration may result in a poor product, but the difficulties could lead to considerable learning about collaboration that is not reflected in that product. While product-focussed assessment is the simplest approach to collaborative assignments, simplicity is not the only principle in play.

3.5.2 Transparency

Allied with simplicity is the need for an assessment scheme to be transparent and clear to the students from the outset. As Ava explained

I think we have been very, very clear with how the grades are composed of at the very beginning, it's on the course outline. It's on the PowerPoints, usually on the first day or the second day, they would know exactly how their grades are composed of. So the students are very clear about their evaluation. (Secondary Interview Data, p. 14)

Her approach is by no means unique, with Olga and Lorenzo also agreeing on the need to be clear with students at the start of courses, with Lorenzo providing a rubric so students “know what I was looking for” (Secondary Interview Data, p. 193). Laura explained in more detail how her centre works with criteria with students,

We have separate criteria. For example, for the preliminary research, we have one criteria, and their performance is assessed based on that criteria. And all the criteria is ... all the criteria, it's shared on LMS is transparent. We go over the criteria with the students. We got samples, real students samples, we analyse student samples, together with the criteria so they they are fully aware of the expectations. So we got the criterion for each stage. And they practice, they know what the criteria has, and just they analyse it, they work on it. And then we mark each stage using different bits, parts of the criteria at hand. And the presentation, problem/solution paper presentation, also it has another criteria on its own, which marks the group dynamics. And the other half marks the students individual performance. (Interview Transcripts, p. 222)

That does not mean all participants took a criteria-driven assessment approach. While Andrea had criteria to assess the written work, assessing collaboration was done “out of experience” (Interview Transcripts, p. 4). However, she did go on to admit that “if there could be a clear set of descriptors, of criteria described you know, that the teacher could refer to, then that would be a very good idea,” (Interview Transcripts, p. 10). Humberto, with first year students particularly, did not go in to too much detail with “assessment scales and criteria. I'm just saying to them, go and do this, and blow our socks off when you come back” (Interview Transcripts, p. 162). Even for advocates of transparent assessment criteria, collaboration posed particular problems. As Ava noted,

I think it should be assessed, because it's important, but it's not as much assessed because it's hard. Because, to me, being a teacher, the number one thing that I think I value is fairness and transparency. So I want to explain very clearly to my students at the end of the semester, and throughout actually, they can, they can know why they got such a grade. And if I grade based on collaboration, I find it very difficult to explain to my students why you are getting an F when your teammate is getting an A. So to me, I don't, I don't find it comfortable to do that. (Interview Transcripts, p. 52)

Fay took a broader view, considering transparency “key to this collaborative work. It has to be very clear and transparent for everyone. So that you will always know what is allowed, what is not, what is correct” (Secondary Interview Data, p.88). Van also expanded transparency beyond assessment criteria, to making the justification for tasks, activities and assignments, clear to students, explaining

The first stage of teaching on some approaches to pedagogy is, you know, setting the goal, setting the objective and tend to be quite proactive about that thing. Yeah, this is what you're gonna do. And these are the reasons why it's good and linking it to their future of study, future career, all that kind of stuff. (Secondary Interview Data, p. 271)

Transparency about assessment could also overcome some perceived deficit in an assessment scheme, with Lorenzo explaining in relation to the Model A assessment scheme that “there's a part of me that sort of wants some individual leeway. But I think it's fair if the students know going in into the assignment, how it's going to work, then I think a teacher is justified in this type of model” (Interview Transcripts, p. 244). Ava concurred, as she felt more confident in assessing a collaborative assignment after she came to terms with the fact that she could not do it perfectly but she could be transparent with students about the expectations and requirements of the assignment and hold consistently to the standards she had set (Interview Transcripts, p. 47).

Jerry took a broader view of transparency, thinking not just about a teacher and a group, but about how a collaborative assignment is conducted across a wider programme, involving multiple teachers. In his view,

one of the worst things you can do in a collaborative assignment as an EAP tutor is that different tutors are giving different information. This is frustrating for students and then of course, they will talk so they start comparing each other. So the information needs to be consistent. (Interview Transcripts, p. 176)

However, in the way that 'process' could be potentially confusing as it could be used to refer to both process writing and the process of collaboration (as discussed in 1.4.6.5) 'transparency' had another dimension where participants were referring to something quite different. Participants did not just use 'transparency' to consider how clear and understandable the assessment was to the students, but also how visible a student collaboration could be to teachers. Dorian contrasted assessing collaboration and the final product, feeling "collaboration is very difficult to observe, final product you can see" (Interview Transcripts, p. 28). This reflects the aspect of assessing collaboration Ava had difficulty with, the lack of transparency of collaboration, and she struggled with coming "to terms with some of the things that I cannot see," (Interview Transcripts, p. 47). Wendy made a similar point, suggesting assessing collaboration feels "a bit subjective or how can, you know, how can you know what goes on in their collaboration?" (Interview Transcripts, p. 298). Humberto contended that good collaboration itself was invisible, drawing a parallel between good collaboration and good acting, explaining, "Good acting is acting you can't see. You don't know, you sort of get, you have this feeling that "Oh, that person was a good actor, wasn't, weren't they?" (Interview Transcripts, p. 159). This difficulty of witnessing and therefore evaluating collaboration is a central concern in assessing collaboration, and was well expressed by Fay when saying

I may not see everything is going on in all groups all the time as to give a fair assessment of what was going on. As I said, unless I had cameras going on and all the groups all the time, and took the time to watch all these films and analyse them and then give grades then, which again, I don't want to do. (Interview Transcripts, p. 121)

There was discussion in the interviews as to whether online collaborative workspaces could help mitigate the concern about not knowing who had contributed what to an assignment. The potential of utilizing online workspaces was recognised, with Dolly discussing the opportunity to look through the version history of a document to see which students had been working on which parts and Fay talking about being present in the students' document throughout the process to "judge how the interaction and collaboration is working, as well as how the writing is being treated/developed" (Pre-interview Task, p. 84). However, even strong advocates for the use of collaborative workspaces, such as Cristobal and Fay, question how effective online spaces are for giving genuine insight in to collaboration. Both Cristobal and Gonzalo highlighted the issue of shared documents only showing who typed what, but not capturing all the collaboration which could take place outside a shared document, through discussions on Whatsapp, on the phone, or over coffee. Fay was concerned about the level of monitoring going into following online documents, which would change her role into that of "a police inspector rather than a guide" (Interview Transcripts, p. 117).

3.5.3 Fairness

Perhaps the dominant principle the participants held when it came to collaborative assignments was fairness, with a quote from Ava from the previous section capturing the position of many of the participants. "The number one thing that I think I value is fairness and transparency" (Interview Transcripts, p. 52), which chimes with Dolly's assertion that "I think, as a teacher, what I try to pay attention to is try to be fair" (Secondary Interview Data, p. 65). Indeed, a form of the word "fair" was used by every participant in their interview, with 131 total uses across 20 interviews. Of the three principles, fairness seems to hold the greatest sway, with Gonzalo finding it important that students "know that they're not doing any more work than anybody else for the same, the same grade" (Interview Transcripts, p. 149). This issue of fairness of student workload was also discussed by Kyle, who considered the impact on motivation if unequal workload was not addressed (Interview Transcripts, p. 207).

Fairness links back to both senses of transparency, with Karen explaining that in she is always clear with students about the assessment criteria from the outset because she knows that in some other cases “students are told after they've done the task, what they're going to be assessed on. And I think that's horribly unfair” (Interview Transcripts, 185). She also referred to the transparency of collaboration to teachers when discussing Model A, explaining

I gave this one star simply because I couldn't see it was, it was fair, how much do I know they've really done? And how do I know they did what they said they did if there was no observation (Interview Transcripts, p. 191).

However, she did feel that if observing the process of collaboration was possible, then assessing the collaboration, “would appeal to my sense of fairness” (Interview Transcripts, p. 191).

The challenge of fairly assessing collaboration was also discussed by Ava, who explained

I actually observed the dynamics of groups in class when they're working in class, and also from talking to people, but I actually, I really tried to shy away from grading them based on my perceptions because I don't think that's fair in the sense that some groups are just more open, they talk to me more, I'm more involved in their work. Or some students are just more talkative. They come to me more asking more questions. But I don't think those are objective. (Interview Transcripts, p. 52)

Her concern was shared by Lorenzo, who admitted

In some ways, yes, we should assess the collaboration. But my question is, how can you fairly and effectively put a point value or should you put a point value to that collaboration? I don't know. I don't know, Peter, I just don't know. (Interview Transcripts, p. 249)

Olga's concern (Interview Transcripts, p. 253-254) was not just with the need for assessment to be fair, but also that the assessment had to be justified to the students. This was something she struggled with more in the Finnish teaching context, where there is a high level of teacher autonomy and academic freedom, rather than more structured educational cultures in which assessments would be determined and justified at the programme level, rather than by individual teachers. She also went on to make the connection between assessment approach and course type, suggesting that in a skills course, taking a product-focussed approach and giving students all the same grade based on the quality of the final work would potentially be acceptable "but if we're looking at the essay to see their linguistic aspects their, you know, language skills, then I will find it a little bit unfair or difficult to justify" (Interview Transcripts, p. 256).

Fairness is also a factor in the assessment tools that can be used. Bertha (Interview Transcripts, p. 65) expressed concern about the fairness of peer-assessment being employed and the need for careful scaffolding so students are able to engage in peer-assessment effectively while Cristobal (Interview Transcripts, p. 95) shared a similar concern regarding using reflection and checklists as "I guess you can get a sense. Yes. But that's the difference, is a sense the same as ... is that accurate enough to be an assessment?" (Interview Transcripts, p. 95). While different assessment tools could have issues, Fay felt that utilising peer evaluation was not perfect "but it's much fairer than just having me decide what's going on" (Interview Transcripts, p. 123, emphasis in original).

It was also acknowledged that no assessment scheme was perfect and fairness could be compromised. As Fay explained

They're all responsible at the end of the day for producing something that is good. And I ... because I didn't want to police the process and put someone more, you know, give someone a higher grade for having done more work, I felt that this is this was fair. Even though yes, some people do more work. But again, that's, it's Pareto's Law. I see this at work as well, in our department, you know, 20% of the people to 80% of the work. In any institution, I've been in any organisation where this is true. And so that is I believe, nature. So I was okay with saying that this is fine. Even though I know it's not an equal distribution of work. I know. It's fine. (Interview Transcripts, p. 119, emphasis in original)

While Gonzalo would concur in part, accepting that, “I don't think we'll ever get it so everyone's doing *exactly* one quarter of the work. That will never happen,” but going on to stress that, “I think we have an obligation to try and make it as fair as possible,” (Interview Transcripts, p. 150, emphasis in original).

3.5.3.1 *Differentiation and surface level differentiation*

One of the central issues raised by fairness is students getting the results they should get based on their work. As Wendy explained, “I basically just don't think it's fair for your individual, your individual marks to be affected by other people” (Interview Transcripts, p. 284). While she was supportive of the developmental benefit of students collaborating and facing the challenge of working with people with different approaches and attitudes, she wondered whether collaboration “should that be something that then impacts on a grade? If, if all of that happens, and then they're graded on a piece of writing? Is that fair?” (Interview Transcripts, p. 284). Her concern was crystallised when discussing how students could contribute in a group, considering

you're very likely to have one or two students to do all the work. And one student who really doesn't pull their weight at all, but then that student will get the benefit of whoever's the strongest person in the group. Um, so it doesn't seem a very fair reflection on each student's contribution. (Interview Transcripts, p. 290)

The desire to differentiate between students was one of the factors in Model B getting positive feedback, as the individual complementary assignments are, as Cristobal put it, "testing their individual skills. And in a way, that is, it's a good way of assessing their, what they've learned from the course" and "allows for individual effort and individual quantity to come through" (Interview Transcripts, p. 87).

Fairness also came up as an issue in assessing an oral collaborative assignment rather than a written one, and raised the issue of surface level differentiation. As Wendy explained, during an oral presentation "it might be easier to make an individual judgement on presentation skills, like oral presentation productive skills in a way that you can't do with a piece of writing" but she recognised that the presentation did not give insight into "what goes on in the background or who's, who's the driving force, in terms of the preparation and the workload" (Interview Transcripts, p. 287). Both Van and Fernand felt it was easier to see the ability and knowledge of individual students in a group presentation, with Fernand highlighting the value of the question and answer section. Although he admitted there was no real assessment of the collaboration, or what students had contributed to the development of the presentation, he still felt a group presentation, "was easier to to judge and to judge more fairly" (Interview Transcripts, p. 132) as in an oral task there is a surface level differentiation, where individual elements like oral fluency and pronunciation offer a chance to distinguish between students in a group. Van also felt that it was easier to see collaboration in an oral rather than written assignment (Interview Transcripts, p. 270-271). That is not to say assessing collaborative oral assignments is without challenges. As Olga explained,

It's hard to see the contribution because you can only see when they present, but you don't know how much work they put in. And then you could tell like when, how, whether they rehearsed before or not when they're presenting, and then you could see their performance. But it's very hard to tell, like, how much contribution they made, you know, into the whole presentation. (Interview Transcripts, p. 254)

Whether considering written or oral collaboration, EAP practitioners value fairness, despite the challenges and difficulties in arriving at a fair assessment scheme.

The concerns around fairness mirror the issues identified by Kane (2010) and Tierney (2014). There is a desire for Kane's procedural fairness, in that students are treated the same, and that assessment and the decisions based on it are appropriate and reasonable, as well as Tierney's call for transparency, a constructive environment, equity and equality. The appreciation participants had of the opportunity to discuss collaborative assessment during the interview also speaks to an appreciation of the need for critical reflection on assessment practices.

The three principles of simplicity, transparency and fairness, work in concert, underpinning attitudes and approaches to assessing collaborative assignments. These approaches are further impacted by guiding orientations, which are more practical in nature than the underlying principles.

3.5.4 Reward orientation

EAP practitioners want students to benefit from their courses and the participants value the positive learning outcomes from a collaborative assignment and the learning that takes place is seen as its own reward. As Bertha put it,

I think the students gain a lot from collaborating and it would be a shame if they didn't have, have a chance to do that. And I think, I think there are other benefits to it beyond just assessment, like the team building and the confidence building and all of those other factors that I think they'd be missing out if they didn't experience that. (Interview Transcripts, p. 61)

Wendy described a collaborative assignment as something of a “dry run” (Secondary Interview Data, p. 283) so students could develop the knowledge and skills for future individual writing assignments, which was a common position. Collaborative assignments are, in and of themselves, rewarding and, when it comes to assessment, the participants also tended towards looking for opportunities to reward students.

It is important to note that while there was considerable awareness of the potential for dysfunctional behaviour during collaborative assignments e.g. free-riding or strongest student domination, this awareness did not translate into EAP practitioners wanting to penalise problematic behaviour. As Van put it, his concern with free-riding and borderline students passing who should not “is not because I really want to, you know, weed these people out and fail them” (Interview Transcripts, p. 275) but because other students would have to put in more work to compensate and because it is better to identify lower performing students earlier. In the case of unequal workload, for example, there was not a desire to punish or downgrade students who had not fully contributed, as the loss of learning was considered already punishment enough. Barry discussed changes in their programme to explicitly raise awareness of effective collaboration earlier in students’ academic careers so students would truly benefit from the experience and be prepared for later courses. As he explained

what we have found is that students have gone into say a Level 5 module on research methods where they've, if they're allowed to kind of work in a group but work unchecked, they've bound themselves together in a group of four where one very strong person can be the dominant voice and the others can coast. And that comes back to bite them at Level 6 when they have to have acquired all of those research skills in order to complete their final project and you find that two or three students don't have essentially the basics that they need to complete their final projects. (Interview Transcripts, p. 13)

Cristobal was also aware that students may not truly collaborate during a group assignment but felt "if a student doesn't do it, then they're not learning how to write an essay, and they will suffer in the future" (Interview Transcripts, p. 85), a position Fernand would also hold, expressed in his concern of collaborative assignments being an avenue for lower level students to advance without really having the skills to do so.

Rather than punishment, the participants had a reward orientation, where they wanted to reward the students who had taken on the greater load. Fernand illustrated this when discussing group presentations, explaining

it's also easy to identify who's done the work, especially if you're allowed to ask a few questions after because you can pick, you can identify basically the slacker. You can ask him a question him or her question and you got them basically right away. Okay, so you know, you reward the students that knew what, had done the research. (Interview Transcripts, p. 106)

The opportunity to reward students was one of the appeals of Model B and the complementary individual assignments, with Dorian seeing them as an opportunity for students to "get something a little bit more for themselves if they work well on the individual assignment" (Interview Transcripts, p. 25) and Dolly seeing them as a way to overcome being in a problematic group

so if you are unlucky, and then your, your, your group members aren't holding up their end of the bargain, then you still have a chance, like, you can still work harder to sort of, you know, get a higher grade. You can put in more effort to try to make up that difference. (Interview Transcripts, p. 106)

While Dolly considered effort as something which could be rewarded in the complementary individual assignments, it was something that other participants were also keen on rewarding, resulting in a kind of sweat assessment, where students who try should be rewarded. Fernand was particularly drawn to this, and saw it of a way of helping determine distribution of tasks in groups with mixed abilities. He argued more capable students could take on the more sophisticated tasks, while those who were perhaps less capable in terms of language ability could be assigned lower order tasks. As he put it, “maybe a stronger student is doing the more abstract side, the weaker student is just doing maybe more of the spade work” (Interview Transcripts, p, 132). What mattered, in terms of contribution to the group, was that the students did what they were able to do, to the best of their ability, and all the group members were making comparable effort.

Humberto considered sweat assessment in relation to differentiation, arguing students would appreciate an assessment scheme that allowed for differentiation as “they can see that *their* effort has been rewarded,” (Interview Transcripts, p. 169, emphasis in original). Despite her concerns about assessing collaboration, Wendy conceded the use of individual complementary assignments was beneficial because they enabled a teacher to see “how much work they've put in” (Interview Transcripts, p. 290). When asked why effort was of particular concern in collaborative assignments, she discussed prosocial behaviour and having responsibility to the other students in the group as a students' behaviour could have an impact on the other student's grades so “I think that's why it matters or that's why we want them to make an effort. It's not just you” (Interview Transcripts, p. 299). Cristobal continued this theme, suggesting that effort was central to the conception of

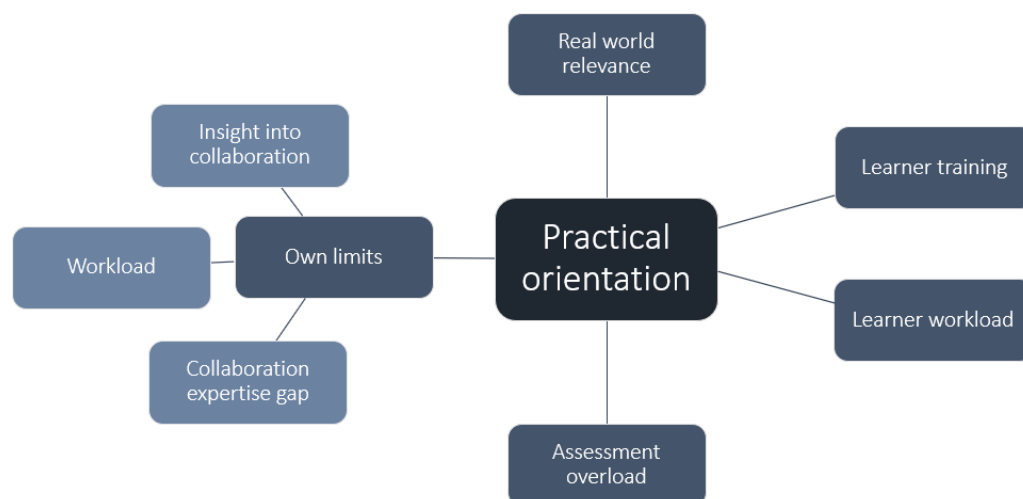
collaboration and would be a necessary assessment criterion since “what is collaboration, it is making an effort to work with other people, isn't it?” (Interview Transcripts, p. 94).

The reward orientation means collaborative assessment schemes need to be geared towards being able to give the students the credit they deserve for the work they've done, requiring a level of visibility into the process.

3.5.5 Practical Orientation

While the participants were keen to find opportunities to reward students for the work they did, they were also very aware of what they as teachers could realistically and practically do. Collaboration was valued for its real-world relevance, but there were concerns about their own expertise gaps, the need for both tutor and student training, awareness of their own limits, workload and assessment overload. The key elements informing the practical orientation are illustrated below (see Figure 3-5).

Figure 3-5 Key factors informing the practical orientation



3.5.5.1 Real world relevance

Collaborative assignments are felt to have a practical value, as they mirror the way students will have to work after graduation. Karen articulated this when explaining

we know that when our students go out into the real world, they're not going to be alone, they're not going to continue doing everything in a vacuum. So it tries to reflect, or at least we try to reflect, how they will be expected to contribute in their workplace to knowledge creation, to knowledge sharing, to the general teamwork principles, you know, that exists within organisations. (Interview Transcripts, p. 183)

and Bertha makes a connection for the students to her own experiences, telling them

things like, like, think about collaboration, about practical things that you have to do at work. To try to bring in em, or you know, even like different skills, like presentation skills, for example, you know, giving examples of - I had to do a 10-minute presentation for every job that I've ever gone for. (Secondary Interview Transcripts, p. 183).

This practical relevance of collaborative assignments means practitioners seemed to accept them as a valid assessment tool, albeit one that posed challenges.

3.5.5.2 Expertise gaps

While collaboration was valued, the facilitation and assessment of it were not areas participants felt fully confident in. Part of this is due to not having experience of such assignments when they were students. As Van reflected, when discussing his attitude to collaborative assignments, “I just wasn't used to it. And because I never did it. You know, it's something you're not familiar with, and you don't understand how it works” (Interview Transcripts, p. 268). Kyle also noted the mismatch between collaboration being an area which is discussed in the literature but is then “not really reflected in terms of, you know, how do you actually do it if you're a teacher?” (Secondary Interview Data, p. 183).

This lack of experience of collaboration as a student is compounded as collaborative assessment was not a feature of participants' teaching training programmes, potentially since, as Dolly observed,

there were so many other things to cover. Lorenzo realised collaboration was something he had learnt about on his own as “it's not something that I received any explicit instruction in as a Bachelors, Masters or PhD student” (Secondary Interview Data, p. 202). Tanya also had no experience of explicit instruction in facilitating collaboration, as ability to do so was just assumed, with teaching collaboration being “just one of those things that everyone is like, 'oh, well, you know how to do this, right?’” (Secondary Interview Data, p. 256). Indeed, evaluating the models of collaborative assessment and assessing them was valued as “a rewarding opportunity” (Interview Transcripts, p. 197) to discuss an important area that was too often ignored, with Bertha acknowledging that “I think it's a topic that's not given enough focus really” (Interview Transcripts, p. 70).

Once involved in teaching collaborative assignments, participants were very aware of gaps in their knowledge due to the complexity of collaboration with “so many things I need to take into account” (Secondary Interview Data, p. 213). Gonzalo, for example, confessed he “would personally like to know how to teach it better” (Secondary Interview Data, p. 130) and expressed a wish “to learn how to make things more collaborative. I would like to know that. If I could figure out a way to make collaboration more fair” (Secondary Interview Data, p. 121). This was echoed by Olga who also felt “I think there's a lot I need to learn about, as a teacher how to conduct it, how to evaluate it” (Secondary Interview Data, p.208). This apprehension about facilitating collaboration can lead to avoidance, with Van electing to leave managing the collaboration up to the students because “I just don't feel like I know enough about it” (Secondary Interview Data, p. 272).

Beyond facilitation, the assessment of collaborative assignments posed the most challenges, with Lorenzo questioning, “how can you fairly and effectively put a point value or should you put a point value to that collaboration? I don't know. I don't know, Peter, I just don't know” (Interview Transcripts, p. 249) and Van wondering “how do you measure this collaboration?” (Interview Transcripts, p. 276). Even in a situation where a teacher was closely monitoring a collaboration Cristobal wondered “how a teacher would know how much they collaborate” (Interview Transcripts, p. 88) as it would never be possible to observe every interaction.

All of this highlights collaboration, and the assessment of it, as an area where EAP practitioners need more support and guidance and further underscores the need for this dissertation.

3.5.5.3 Tutor and student training

Before a student collaboration begins, preparation work is needed, although for some participants, such as Cristobal, actively teaching collaboration skills to students was a novel idea and others, such as Dolly, chose not to spend time on teaching collaboration explicitly, mirroring her own experience of just having to get on with it when she was a student. More widely, however, there was an awareness of the need to prepare students. The need for learner training at the start of a collaboration was discussed by Ava and Karen, with Jerry reflecting on what can go wrong during a collaboration if that learner training does not take place. For Cristobal, learner training is essentially linked to assessment as “you can only test people or assess people on what you've taught them. If you haven't taught them collaborative skills, you've just assumed they've got it. We haven't taught it, you can't really assess it” (Interview Transcripts, p. 82).

The question then becomes when such learner training can take place. Wendy was concerned that “there's too much of everything at the moment, it seems, in the courses. There are too many students. There's, there's a lot of input, there's not enough hours,” (Interview Transcripts, p. 289). When discussing preparing students for self-assessment, which is known to require prior training (Rasooli et al., 2024; Topping, 2003), Dolly discussed the choices that would have to be made about where to allocate resources, explaining

well, how much time do you want to devote in class to like doing this? Yeah, I mean, it's definitely, it's useful and helpful for them. But you, I think as a teacher, I would be faced with some difficult choices, because if you if you wanted to focus on the assessment, you can do like a whole class, right? Just on like, okay, instead of just me showing you the rubric. Right, okay, well, let's actually apply, like different parts of the rubric, let's take this sample, writing and discuss it, and then we can turn that into like two classes. Right. But then, you know, we have to make ... we ... there's just things that we need to cut out. Right? Because, for the lack of time. (Interview Transcripts, p. 110).

Although Bertha felt learner training was necessary, she was also wary of “the time investment that would go into training, like training the students and, and talking through all of those things with the tutors as well” (Interview Transcripts, p. 69). Jerry also saw the need for tutor training around collaborative assignments with “more support and potential training around issues of the collaborative assignments and how to scaffold it and you know, how to make sure that work is being done” (Secondary Interview Data, p. 164).

3.5.5.4 *Awareness of own limits*

As was clear in the discussion of the models of assessment for collaborative assignments, the participants were aware of what they were not able to do. There were limits to how much of a collaboration they could observe, either because collaboration may take place through channels they were not aware of, the collaboration may not look like they expect collaboration to look, or there simply was not the time or bandwidth to go through track changes to try and identify every student contribution. There were further concerns about “scalability” (Interview Transcripts, p. 186) of an assessment approach when the number students/groups increases and the challenge of paying attention to everything it is necessary to assess and so “cognitively I have to be so 'on' all the time,” (Secondary Interview Data, p. 69). While training could mitigate these to an extent, they remain

factors that have to be considered in the design of any collaborative assessment scheme. Ultimately, the practitioners were aware that they cannot know or do everything and there are instances when the ideal of what could be done has to surrender to the reality of what can actually be done.

3.5.5.5 *Workload*

Once collaborative assignments are set up and running, workload remains an issue, particularly in terms of trying to utilise different assessment lenses, such as reviewing tracking changes in online documents, conducting review meetings with groups, conducting peer- and/or self-assessment training with students, or grading individual reflections. Although Bertha and others pointed out the practical value in having fewer overall pieces of work to assess in collaborative assignments, Tanya would qualify this as it is “a lot harder to grade the five” (Interview Transcripts, p. 32). This was further illustrated during the evaluations of the models of assessment and assessing collaboration itself. One of the concerns was the “degree of oversight” (Interview Transcripts, p. 16) needed and the fact that it “sounds like more work for the teacher” (Interview Transcripts, p. 245). As Fay explained, she would not feel comfortable assessing collaboration “unless I had cameras going on and all the groups all the time, and took the time to watch all these films and analyse them and then give grades” (Interview Transcripts, p. 121). Cristobal went further, challenging the value of the putting work into assessing collaboration. As he explained,

it ends up, you're just jumping through hoops, the teachers doing all this admin having to collect all these different things to confirm whether students have done it or not, you know, and it's just now Oh, my God, you know, it's all this extra work, which just brings us back to this simple there's the product, mark, the damn product, mark the video, is it good? (Interview Transcripts, p. 89).

So too can the assessment tools become onerous for students, who are documenting and diarying, potentially performatively rather than productively. The act of looking closely at collaboration can distort it, paralleling a concern of Humberto's regarding teacher intervention in writing when he said, "too much intervention and feedback can actually undermine their writing skills" (Interview Transcripts, p. 163). If students are always thinking about and discussing and getting feedback on how they are collaborating the danger is they will not collaborate naturally. As he put it, "the more I interfere with collaboration, the more ... the less collaborative it becomes" (Interview Transcripts, p. 170). Fernand spoke directly to the workload, talking about a collaborative writing assignment where

they'd completely lost the plot here with the amount of forms, you know forms and stuff they were getting students to submit to, and they were getting burnt out and fed up with it, and we were kinda getting fed up with it. (Interview Transcripts p. 136)

Whatever assessment scheme is in place, and whatever assessment tools are used, they have to be doable amongst the other work of both teachers and students. Bertha recognised this when discussing the possibility of students and teachers co-constructing assessment criteria for assessment, as Karen advocated, explaining

maybe it's something that the students could work with the tutor to come up with, a set of criteria for that. So what's important for the students as well in collaborating, seeing as they're the ones who are going to be doing it? Erm, yes, so I think maybe the students and tutor that could work well that they come up jointly with, with with the themes that should go into these criteria, but then the tutor can, you know, organise it so that it's got the different like levels within, within each theme. But this is like total ideal world thinking here, like this isn't, I don't think this would work in reality. It's like you have to get going, you can't spend like a session discussing this but yeah. (Interview Transcripts, p. 68).

This illustrates the concern that there is too much to cover in EAP courses as it is, leaving no time for 'ideal world' solutions. She expanded on this in terms of preparing teachers for collaboration assignments, acknowledging

it's the time investment that would go into training, like training the students and, and talking through all of those things with the tutors as well. It's not that, it's not that I'm saying that tutors couldn't do that, but I'm just thinking of the time investment you know, especially if you've got people who are part time and who just want to get, come in and get their work done and go home and, you know, don't have time for all of this other, maybe, legwork before, before it all starts. But no, I like the idea but yes there's huge time investment. (Interview Transcripts, p. 69)

3.5.5.6 Assessment overload

Allied to the concern about workload, the multiple assessment tools which may come into play in a collaborative assignment could result in assessment overload, as previously highlighted by Kyle (see 3.5.1). This mirrors Torrance's concern that assessment could come to dominate learning in his conception of AaL (2007). Unlike Kyle, Van did not see assessment overload as a specific EAP phenomena, but rather an institutional issue since

universities spend too much time assessing or rather, they put it, yeah, they they just put put in too much effort to producing too many scores with to fine a level of granularity. And it's just a waste, you know, debating, getting into arguments of whether something's a 63 or 65. (Interview Transcripts, p. 271).

Fernand was concerned, that for incoming undergraduate students, adding collaboration to the mix when they are already trying to adapt to a new academic culture and come to grips with academic English was a challenge too far (Secondary Interview Data, p. 99-100). When considering the weighting

of different assessment tools, Barry was cognisant of the need for “a fair balance of assessment” (Interview Transcripts, p. 17), to try and avoid assessment overload and come out with a fair result.

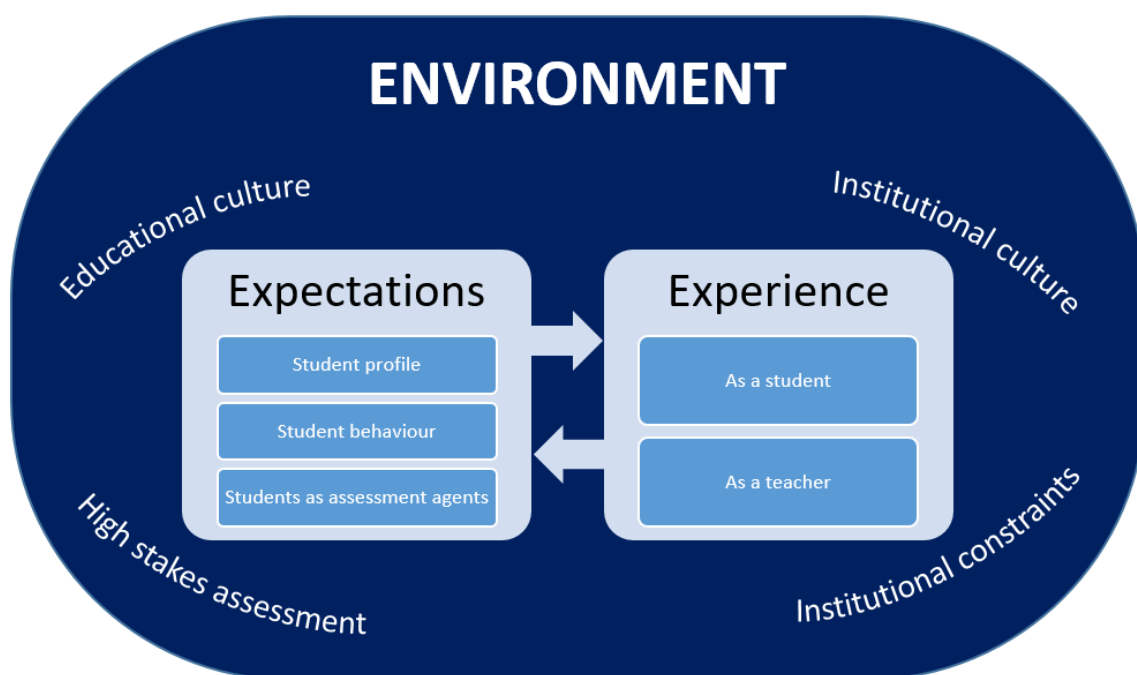
In light of these practical considerations in place, potential collaboration assessment schemes have to be practically actionable whereby the assessment is worth the time and effort put in by both students and teachers. Time has to be available for teacher and student training, with the assessment scheme clear and justified to all involved.

3.5.6 Defensive orientation

Superseding the reward and practical orientations is the defensive orientation of the participants. Defensive is not intended in a derogatory way but speaks more to preparedness, of being aware of what could or might happen and the need to be ready to explain decisions and results since assessment schemes in EAP have to be robust and justifiable to a range of stakeholders. This is in line with Stiggins’ conception of assessment literacy being able to recognise what could go wrong and the ability to prevent such problems before they happen (1995). The defensive orientation is shaped by experience and expectation, which have interinfluence, within the context of the HE environment (see Figure 3-6).

The prominence of environment on shaping practitioner beliefs is not unexpected as “teachers’ conceptions of assessment are ecologically rational in that they reflect the legal, cultural or social priorities placed on assessment for their work environment” (Brown, 2011, p. 65). Consequently, the discussion of the defensive orientation will begin with looking at the impact of environment in which EAP collaborative assignments take place. Then attention will turn to the participants and their own experiences of collaboration as students, their initial expectations when first teaching collaborative assignments, their early experiences as teachers and the established expectations that emerge.

Figure 3-6 Factors shaping the defensive orientation



3.5.6.1 Environment: Working in HE

EAP practitioners work in the university context and are consequently aware that assignments and assessment schemes needed to be defensible to students, colleagues, direct language centre management, wider university management, exam boards and external examiners, and, in the UK context, United Kingdom Visas and Immigration (UKVI). As such, the wider university culture can either foster or stifle assessed student collaboration. This was reflected in Ava's consideration of how much time could be devoted to assessment of collaboration in an EAP class, echoing Gonzalo's recognition that EAP had so much to cover that collaboration was only a minor focus. Ava went on to wonder whether it would be accepted by the board of the University to spend a lot of time on collaboration in an EAP course (Interview Transcripts, p. 50). Olga also contrasted different countries she had worked in and how collaboration could be problematic in the UK system due to exam boards and external examiners but would be more acceptable in Finland, where there was much greater teacher autonomy. Indeed, the institutional culture could cultivate collaboration if there is "a public intention in at the university level that we encourage collaboration between students" (Interview Transcripts,

p. 160). So too could course objectives either lend themselves toward collaborative assignments or put the focus more squarely on individual work and skill attainment.

Building on this theme, institution level decisions could be impactful because, as Dolly bemoaned, “we’ve gone through so many curriculum revisions, it’s hard to keep track” (Interview Transcripts, p. 96), which affects what is valued and taught. Further, institutional decisions to, for example, raise the number of students per intake can result in students with lower language levels being accepted into the university, meaning EAP classes need to deal with more language related issues rather than higher level skills such as collaboration. Even where collaboration was felt to be necessary for students, institutional constraints could make it problematic. As Kyle explained, when talking about the need for more collaboration in courses, “So there’s clearly some need to address that. But due to all of the other pressures of that sort, of course, I can see that it probably isn’t really practical” (Secondary Interview Data, p. 186). The other pressures alluded to were institutional constraints such as the requirements of UKVI and mapping students’ individual language skills onto CEFR. Conversely, institutional decisions can support collaboration, with Fay describing a move in her university towards more collaboration as problem based learning was encouraged across all faculties, with the university “trying to change the way in which academics and lecturers do the teaching, just to make it more effective through more collaborative work” (Interview Transcripts, p. 115).

Moving from the institution to faculty level, Karen discussed how important collaboration was in the courses she taught because in the faculty she taught in “collaborative assignments are very much at the forefront of how they will carry out their projects” (Interview Transcripts, p. 183). The impact of student profile in terms of field of study was also a consideration for Tanya, with some students in Education collaborating all the time, something also noted by Gonzalo, who contrasted that to students of Mathematics who did not encounter any collaborative assignments. Lorenzo also felt faculty could have an impact, explaining students of Economics would potentially have greater need to collaborate, although he also saw a utility for using collaborative assignments with students from any faculty. He made the important point when considering how much of a course could be

assessed collaboratively that “definitely the answer has to consider the course itself. I doubt there is one response that's appropriate for every different type of course” (Interview Transcripts, p. 241).

Educational culture also had an impact on how collaboration could be best facilitated and assessed. As Tanya recognised,

I think it's really context dependent. Because about halfway through the interview, I was like, Oh, wow. Remembering previous places I've taught and then get my answers would be totally different. So I think this is something that can really be localised. (Interview Transcripts, p. 40)

That does not mean she felt collaboration was more or less appropriate in different contexts, just that how it is approached would necessarily be different in different educational contexts. Speaking of students in different places she held, “They all should do it, but the assessment and directions would be very different” (Interview Transcripts, p. 40). Wendy also felt educational culture could have an impact, saying

Culturally, I think if your students have an experience of an academic culture of collaboration and communicative tasks, then they'll probably pick it up and run with it a bit more comfortably and a student from an academic culture, which is very individual. (Interview Transcripts, p.286)

Participants were wary, however, of falling into national culture stereotypes and were much more comfortable talking about educational culture, with Wendy going on to explain

I hate to generalise and say, well, the Chinese have an individual culture and ... because I think increasingly, I'm seeing students in the UK from Chinese schools that have had Communicative Language Teaching, and they've done flipped learning and they've done you know, all sorts of things. (Interview Transcripts, p. 286)

Fernand also wondered if students would have the initiative to work collaboratively with feedback, if coming from an educational background of “simply being told what to do” (Secondary Interview Data, p. 101). Olga also recognised the reluctance students could have to collaborate if they came from a very competitive, exam driven background (Interview Transcripts, p. 253).

Where national culture seemed to have greatest consideration was in relation to peer-assessment and how comfortable students were assessing or giving and receiving feedback from each other, either due to face-saving or having an expectation that assessment is the teacher’s responsibility. This has knock-on effects for how peer-assessment could be utilised as an assessment lens. Educational culture could also have an impact on students’ familiarity and comfort with teacher- or student-centered learning. Students coming from a rote learning education culture could potentially struggle more with collaboration (Interview Transcripts, p. 78) or students could find collaborative writing difficult because if “they haven’t had any, too much in-depth collaborative writing experience, then it poses like additional challenges” (Secondary Interview Data, p. 57). However, even within the same national culture, participants were aware there could be quite different educational cultures in different institutions (Secondary Interview Data, p. 89).

The high stakes of EAP courses are also a factor, with Fernand being aware that the results of an EAP course mean “you’re saying this student is capable of going ahead in this university” (Interview Transcripts, p. 132). Consequently, the middling out of grades that occur in collaborative assignments can be a barrier for practitioners who need to attest to an individual’s language competence. The import of EAP results have an added weight in the UK context, where outside bodies also influenced participant decision making. In relation to pre-sessional or foundation courses, students’ results indicate language level competence, which is required to support students’ visa applications. Consequently, ensuring assignments are acceptable to UKVI is an additional consideration at odds with collaborative assignments, particularly in the conception of collaboration wherein individual voices are indistinguishable.

Environment can have an impact on collaboration in more practical ways, creating further institutional constraints. Fernand reflected on the impact semester length had had on what could be done during a course, with experience of 15-week semesters being reduced to 11 (Secondary Interview Data, p. 106), whereas Karen discussed the frustration of students having many collaborative assignments to do but inadequate physical spaces where they could meet and work in small groups (Interview Transcripts, p. 187).

That is not to suggest that practitioners are fully at the mercy of the environment and institution in which they teach. Rather, as Kyle put it, “It all kind of comes into how innovative are you prepared to be against the backdrop of your, your institutional commitments” (Interview Transcripts, p. 211).

3.5.6.2 The interinfluence of expectation and experience

EAP practitioners bring a lot of expectation of collaborative assignments with them, partially informed by their own personal experience of collaborating and more general beliefs about students and learning. These expectations are then influenced by the experience of teaching collaboration and coming across both positive and negative experiences. It is difficult to parse the relative impacts of expectation and experience as experience informs expectation while expectations can shape experience. There can also be contradiction between the two, as we can see this in the following quote from Bertha, when considering if all a course could be collaborative.

I don't know. I just feel like if I were that student, you would kind of feel like maybe you're getting ... you're getting lost in the crowd, and that you're not maybe getting the feedback from the tutor that you would want. That your course maybe wouldn't feel personalised or applicable to you. Maybe I'm wrong in thinking that. I just haven't, I haven't, I've never had the experience of being involved in a course like that where there was so much collaboration. So I think, like my assumption that the students would be upset and annoyed at having to rate each other, that was completely wrong. Erm, because they did it really well, and they were honest, and so on. So my assumption here that students maybe would maybe feel sort of hard done by, that could be completely wrong as well. But yeah, I don't know why I have that feeling, maybe it's my own thoughts of myself as a student that I wouldn't, I wouldn't have liked to have been assessed in that way. Therefore, maybe I'm putting that assumption on other students, which could be completely wrong. But yeah, I think that's why I think that. (Interview Transcripts, p. 60)

Here we can see her personal experience as a student not wanting to be assessed collaboratively turning into expectation that that is what her students will also think. We can see her experience running counter to her expectation in the case of peer-assessment, which she had thought would be more problematic than it had been. This allows her to acknowledge her assumption about students rejecting collaboration may also be wrong, but she still maintains her position.

When participants were describing some student behaviours it was sometimes unclear if it was based on what they thought students would do, if it was based on consistently observed behaviour, or if it was a generalisation borne out of a single observation. Nonetheless, the defensive orientation is shaped by the interinfluence of expectation and experience, with the participants' earliest experiences of collaborative assignments coming when they were students themselves. When asked about their experiences of collaboration, several participants recalled taking part in sports at school, and collaboration as part of working life was also common. In relation to their own experiences as

students, few recalled participating in assessed collaborative assignments, although they may have worked with other students in non-assessed activities. Perhaps this indicative of the age profile of participants, as Jerry explained,

I honestly don't think I did. I just ... no. I went to university, what, over 20, 25 years ago and my Master's was done by distance. So there wasn't, there wasn't any kind of group work there. We have, you know, discussion boards and things but it wasn't group work as such. (Secondary Interview Data, p. 154)

This lack of experience with taking part in collaborative assignment themselves made it a novel form of assessment for participants when they became teachers and, as Olga explained, “every time when someone suggests a collaboration, it's a little bit like taking you out of your comfort zone” (Secondary Interview Data, p.214).

While Ava was one of the participants who could recall participating in collaborative assignments herself as a student, she was very aware that she went to elite institutions with driven students so did not feel her experience was fully relevant to that of the students she was then teaching (Secondary Interview Data, p. 10). Fay also had a very positive experience of collaboration as a student (Secondary Interview Data, p.79) and Lorenzo had both positive and negative experiences in groupwork, with positivity coming when he could “get into a field that you're personally interested in” (Interview Transcripts, p. 243). Fernand had two experiences he could remember, one of which went well, and one of which did not. The more successful collaboration was a pair work activity where they circumvented collaboration and adopted a more' cooperative approach (Secondary Interview Data, p. 101), while the other was a larger group assignment of six people “where one student was a bit weak and the other one was basically a slacker” (Secondary Interview Data, p. 100) and he struggled, as the team leader, to navigate the situation as he found it “a bit hard to manage and I think that might have negatively influenced me somewhat” (Interview Transcripts, p. 140). Kyle realised he had likely been

the problem partner in a collaboration he had been a part of (Secondary Research Data, p. 176), as did Bertha, who acknowledged “I think when I was a student, I did not see the value at all in collaborating. And I thought it was a complete waste of time. And looking back, I was the weak link in the group,” (Secondary Interview Data, p. 23). Cristobal shared this negative attitude toward collaboration as a student, although he recognised the contradiction in that as he actually advocates for student collaboration, realising

as a younger student, I didn't like collaborating, I'm one of those ones who prefer to get on with it on my own, I don't want to rely on other people. So I can understand, although I love teaching, collaborative teaching, I'm very into group work, and so on. So I kind of go against myself. (Interview Transcripts, p. 79)

However, it is recognised that “that teachers should have multiple and possibly contradictory conceptions of assessment,” (Brown, 2011) so this dichotomy is not unusual. Cristobal’s dislike of collaborating as a student himself was echoed by Tanya, who carried considerable anxiety around collaboration when she was a student as she was quite shy and recognised

I also was kind of a, maybe a perfectionist at times, like, 'oh, well, my group members' work isn't going to be up to my standards I need to do myself'. And then maybe the reality of group members not actually, not doing the work. (Secondary Interview Data, p. 252)

Tanya’s fears were realised by Wendy, who was pro-collaboration (although not pro-assessment of it) but had quite visceral recall of a very negative collaborative experience she had while doing her MA, describing it as “an awful experience. It was awful. It was just such a bad experience for me on so many levels” (Secondary Interview Data, p. 290).

Echoes of some aspects of participants' own experiences of collaborating as students can be seen in their attitudes and approaches to collaborative assessments as teachers. Wendy's negative experience of being graded as part of a problematic collaboration perhaps feeds into her firm stance that collaboration should not be assessed. Fay's positive advocacy toward collaboration may be the result of her own positive experiences of collaboration as a student. The resonance of personal experience as a student was clear when Bertha was explaining her rating for Model B, saying

I think maybe I was thinking about my own experience as a student rather than as a tutor. I was thinking about, I don't know, there was a couple of times at uni that we had to do collaborative tasks. And to be honest, I hated doing that stuff myself. I'd rather have just, you know, in hindsight, I'm thinking now that it was, you know, good practice and everything, but I didn't, I didn't like doing those kinds of tasks. So I think maybe for that reason, just for my own personal erm, feelings on, on those kind of group work projects, that I would have preferred to have gotten an individual mark as well as a group mark. But, I don't know, that's that's the thoughts of me as a 18 or 19 year old, you know, maybe it doesn't reflect fully what I really think now. (Interview Transcripts, p. 63)

Personal experience of collaboration as a student is not the only role experience plays in shaping practitioners' beliefs and attitudes towards collaboration. Experience of teaching and facilitating collaboration is also a factor, providing a very different perspective, as alluded to in the previous quotation. As Ava explained,

when I was a student, I barely think of it from the teacher's perspective. I think as a student, all I thought was, I have this collaborative assignment and just do it, deal with it. Right. But now that I'm a teacher, I think there are a lot of things going into this. (Interview Transcripts, p. 50)

For many participants, their first experience of collaborative assignments was the teaching of them. They had little to no personal experience to draw on, and no guidance in terms of training. As such, they were guided by their initial expectations. Some, such as Van, approached his first experience with collaborative assignments with scepticism, explaining

I guess I just wasn't used to it. And because I never did it. You know, it's something you're not familiar with, and you don't understand how it works. And so yeah, I was just a bit sceptical to start with. So maybe it was a fudge, a fudge you're just getting around it, you're allowing these free-riders who can't do it individually, allowing them to piggyback on on their classmates. (Interview Transcripts, p. 274)

Fernand went beyond scepticism, saying of his first experience with collaborative assignments, “Most people did seem to be negative. There wasn't too many, can't remember too many positive voices about it” (Secondary Interview Data, p. 110). Even teachers with experience of collaboration as students themselves could have rather negative expectations, as was the case with Lorenzo who recalled

You know, I was a student for a long time, as well. And I remember these sort of group projects and things. And sometimes they go well, sometimes not so well. So I think I was maybe a bit apprehensive” (Secondary Interview Data, p. 195).

It is relevant to note that in Fernand's case he was discussing the introduction of a collaborative assignment into a course for the first time and the resistance it encountered from the teaching staff. In contrast, Bertha had more positive expectations, coming into the same institution when collaborative assignments were already established. For her,

it probably made me trust that this was something, em, that worked and something worthwhile because I wasn't overly experienced at that time. So I think it was good for me to have guidance from others more experienced. So I think it made me, yeah, made me trust that that was a good approach, and that it worked, because they'd done it before. (Secondary Interview Data, p. 20)

Perhaps reflecting the way they themselves had received little to no instruction in collaboration and it had been assumed they would just work how to work with other people, so too did some participants initially assume students would be successful at collaboration if left to their own devices. However, once participants gained experience, they became more aware of what might happen during collaborations and became, as Tanya explained, “more confident myself in what I expected” (Secondary Interview Data, p. 248), enabling more effective facilitation as they had a better idea of what they wanted students to do and of what could go wrong.

Ava reflected on struggling with the assessment of collaboration initially, particularly in terms of assessing a teacher/group consultation to discuss the collaboration as she didn't “feel very comfortable grading students self-reported challenges” (Interview Transcripts, p. 46). Lorenzo also encountered a difficulty in assessment. Although collaboration went well the first time he used it, the second time one of the groups had a complete breakdown, to the point where he assessed two of the students on the individual work they'd done towards the assignment so they could finish the course and the other two students “are just somewhere in a black hole” (Interview Transcripts, p. 236).

Not that participants only encountered problems. Bertha discussed possible issues that could arise during collaboration with students at the start of the group assignment and asked them to contact her as a last resort if the suggested solutions were not effective. This demonstrates seeing collaborative assessment as dynamic (Anton, 2012), since the teacher is there to intervene to try and improve performance during the task. Bertha found very few students ever needed to come for her intervention, explaining, “it didn't happen as much as I probably would have expected. I would have imagined that students would come asking for help more often than they did” (Interview Transcripts,

p. 58). However, she also questioned how effectively collaboration had been taught when reflecting on a group presentation assignment which “was something like three individual sort of like, stuck together, rather than one thing that was seamlessly worked on collaboratively” (Interview Transcripts, p. 56).

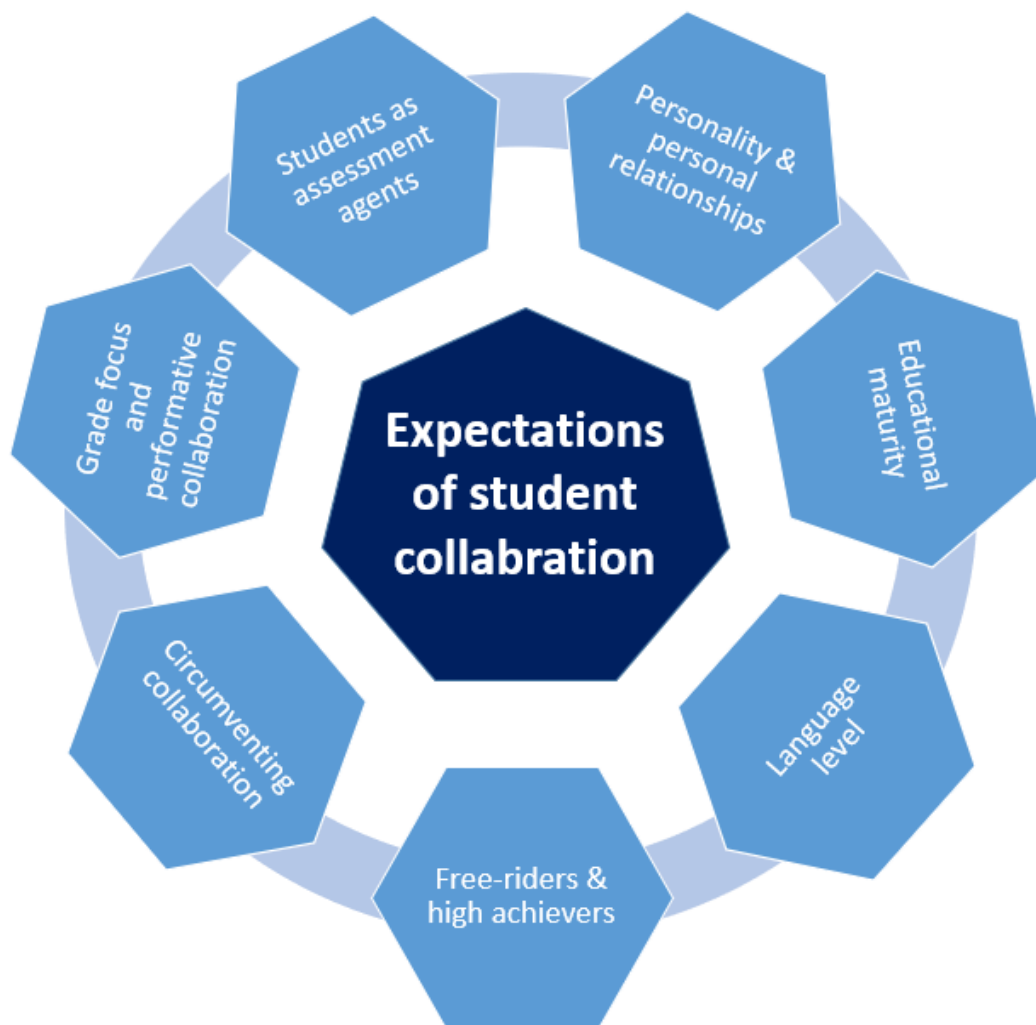
Dolly also appreciated the impact of task type, contrasting two collaborative assignments she had delivered. One with trainee teachers producing an e-book seemed to get more buy-in from students, as they could see the relevance and transferability of what they were producing. The other, with students from different majors working on a project was more difficult as students had more difficulty finding something they were all interested in (Interview Transcripts, p. 103).

The next section will move on from direct experience to wider expectation the participants held around collaborative assignments.

3.5.6.3 Expectations of collaborative assignments

Participants held expectations in three broad areas related to collaborative assignments, from student profile, to student behaviour, to the role of students as assessment agents. Many of these expectations are coloured more toward hampering rather than facilitating collaboration. That does not mean participants expect, for example, all students to freeload or fall out with group mates. Rather, there was awareness that such behaviours could happen, which has an impact on practitioner attitude and decision making, helping to shape the defensive orientation. The range of factors impacting EAP practitioner expectations of collaborative assignments are illustrated in Figure 3-7 below.

Figure 3-7 Impactful expectations of student collaboration



These are not illustrated with a clear delineation between student profile and student behaviour as aspects of student profile (e.g. language level) impact behaviour (e.g. circumventing collaboration) or students having a grade focus (profile) could lead to performative collaboration (behaviour).

3.5.6.3.1 Personality and personal relationships

One area that created expectations for participants was personality and personal relationships. Jerry outlined these when saying

Sometimes it's quite simply that but they don't like or they don't get on with someone in the group. So they've asked to change for that reason. So that that can be one reason. The other reason, again, is that people feel that this student has a much lower level of proficiency in language. So they think that's challenging. Another student is maybe not responding to emails or texts, or just basically not being very well organised. (Secondary Interview Data, p. 157)

Fernand also touched on how friendships may impact openness of groups

You'll get two or three students that know each other. And the problems are, they're possibly quite good and they'll go in the group together. There'll maybe be two of them that not want anybody else in the group, but there has to be, There might be three of them that don't want the fourth person there. (Secondary Interview Data, p. 104)

Karen was similarly concerned about exclusion if students were able to self-select friend groups, so adopted a random grouping system. Fernand was more deliberate with his approach to grouping, explaining

I try and get a balance of personalities, try and get a balance gender wise as well, okay, as much as possible. And, also, you'll see that by this sort of second week, they've already started to form like, like cliques within the class and then maybe think I'll break those up and try and get them to work with, with people maybe they wouldn't normally sit with or work with them when we're just doing like pair work or something within the class, you know, when we're doing the initial writing tasks or whatever. Maybe it's me getting my own back for what happened in third year. I think it's like, you know, if it is preparing people for for for life outside of university, you can't pick who you're going to work with, you get a job and these are your colleagues, get on with it. (Interview Transcripts, p. 137).

Laura was also aware of the impact of personal relationships, suggesting “they will not be collaborating with the people that really get on well with” (Interview Transcripts, p. 231) but then also went on to describe an experience with a friend-group, where one of the members confessed, “We have been good friends, I assume that things would go smoothly. But that was not the case. Friendship is one thing, being professional and working together is one thing” (Interview Transcripts, p. 231). The impact of dynamics of friendships on student collaboration is also unclear in the literature, with Maldonado et al. recognising “Prior research has been split on when the relationship between friendship and group performance is a positive or negative one” (2009, p. 230).

Karen reflected on her personal collaboration style and how she takes a big picture approach and plans a pathway, where as some others “start from the let's do this first and then let's do that second. And I get very frustrated, because I can't see where we're going” (Interview Transcripts, p. 182). Wendy also discussed how people's differing approaches to work could create interpersonal issues, with students experiencing

the frustration of trying to work with a group of people who just have very different working styles, or you're trying to organise something and other people, like one person, just isn't getting involved or isn't replying to their emails, or one person's very domineering. (Interview Transcripts, p. 284)

and she also recognised that one student in a group could act as a “poisoned seed”, turning everyone in the group against the collaborative task (Secondary Interview Data, p. 297). In relation to personality types, Bertha questioned, “if somebody's shy or, you know, somebody who's quite outspoken, what impact does that have?” (Interview Transcripts, p. 66). Cristobal also saw the importance of personality to collaboration explaining “I guess it reflects the personality of the people. I mean, do they want to work with others? Will they listen to others?” (Interview Transcripts, p. 81). Laura broadened out the impact of personality from collaboration to education more widely, opining,

Some of us do well, when we are working on our own. Some of us are natural born team players, they can do really well when they're in teams. So I think we should be addressing different personalities and characters. It's not simply about EAP also, there must be room for students where they feel comfortable and where they can show maybe their real performance. (Interview Transcripts, p. 226).

When Wendy was asked if there were some students who would just not enjoy collaborating, however it was set up or facilitated, she said

Absolutely. Definitely. 100%. Control freaks. You know, you, that's what I mean about in any group, you'll always get an X and a Y and a Z, I think, or like the chances are quite high. And I just think it doesn't work for some people. In the way that exams just don't work for some people. Definitely, some people just, it's never going to be their preferred way of working. And that's just not how they operate. (Interview Transcripts, p. 287)

and this response likely reflects her own negative experience of collaborating as a student.

3.5.6.3.2 Educational maturity

The expectations participants had based on national and educational culture impacting the ease of implementation of collaborative assignments were touched when discussing environment. Educational maturity was also felt to have some impact on how able or effectively students would collaborate. Fernand saw a place for collaboration later in student's time in university, suggesting the fourth year when it would be

setting them up or trying to set this up for them to understand what they need to do, either in the workplace, or postgraduate, where they'll expected to be able to be mature enough and confident enough to go ahead and work with other people and produce a good piece of work. (Interview Transcripts, p. 139).

Bertha would concur, appreciating

but I feel like it may be useful, like, particularly useful for master's students. I'm just thinking sort of practicality wise, you can sort of, I don't mean sell it to them, but like you can help them to see that it could be useful for them for, career prospects and for jobs. So I suppose there's something very tangible there that, you know, whatever job you're going into, you'll likely have to work with other people. (Interview Transcripts, p. 57)

This is in line with Cristobal's personal experience, when he felt he was better able to collaborate with other people during his MA rather than his BA. In contrast, Van felt the postgraduate level was when students needed to focus on "producing individual work faster and sooner and higher volume" (Secondary Interview Data, p. 262), making collaborative assignments more suitable at undergraduate level. However, he did concede there might be scope for collaboration PhD level since students might be moving into co-authorship of journal articles.

As mentioned previously, Tanya felt collaboration could be used with a wide range of student cohorts, but saw that her role could change, with her taking on more of a supportive facilitator role with lower educational maturity students. Kyle agreed with this, commenting

those students in the in the first year out of high school, the teacher's role as a mediator is probably more pronounced than it would be, if we had some Masters students where, you know, we would probably expect them to just be getting on with it. (Secondary Interview Data, p. 186)

Dolly also saw a contrast based on educational maturity, finding it easier to implement collaboration with first year students, finding “I usually get the enthusiasm from, with the year one students. They don't know any better yet, right?” (Secondary Interview Data, p. 62), whereas more mature students seem more jaded with collaborative assignments. She went on to note that collaboration in the faculties was something that seemed more apparent later in students’ journeys through the university. Wendy illustrated conflicting expectations when discussing educational maturity, seeing reasons for and against collaborative assignments at undergraduate and post graduate level, explaining

if they're postgraduate, they they will probably be more confident and they're more likely to have done it before. So they're probably going to do it better. Does that make it more suitable for them? I don't know, they might find it all a bit frustrating and condescending at that point. I suppose the undergraduates stand to learn more from the process. Although that is a massive generalisation, you're just assuming that undergraduates have less, less academic experience and less less life experience and so wouldn't wouldn't be able to do it as well, and so they'd learn a bit more from it. (Interview Transcripts, p. 286)

When discussing using collaborative assignments with different student cohorts, Dorian speculated that while it would be a useful learning experience for students relatively early in their academic careers “it might be better later on in their studies, when hopefully they would be more mature and understand what's needed” (Interview Transcripts, p. 22).

3.5.6.3.3 Language level

In an EAP course, the language level of students also carries with it considerable expectation, particularly as in EAP students may not be streamed according to language level but rather to academic level e.g. incoming first-year students or field of study. Ava voiced concerns around language level of students and how it could impact a collaborative assignment with low-level language level students perhaps not being ready for collaborative work. The concern about lower level language students was part of Wendy's motivation for wanting to see a rebalancing in EAP more towards language development, as the basic language skills needed to be there to start addressing academic and higher order skills. On a more optimistic note, Fay felt collaboration could fruitfully be done with students at A2, with it being a matter of changing the task type to one with a lower cognitive load (Secondary Interview Data, p. 80). Alternatively, Fernand was content for students of differing language level taking on tasks of differing complexity, suggesting lower language level students could do more on the research side, supplying the ammunition for the argument articulated by the higher language level student (Interview Transcripts, p. 137).

Barry also touched on how language level could impact how students collaborate, recognising that during a collaboration by a monolingual group with lower English language level, "there is a degree of trust that they will in some respect self-police" (Interview Transcripts, p. 12) and operate in the target language. The concern of group operating language was also voiced by Bertha when she said of collaboration,

it's trying to create a genuine communicative purpose for a task that the students will, in theory, that they should be using the target language to work together. But that's one issue that I have with it, is if you've got a group that are from the same linguistic background, you're not going to be able to ensure that they're working together in English (Interview Transcripts, p. 55)

The largest concern about lower level language students, voiced by Ava, Jerry, Olga, Wendy and Bertha, was the impact on group dynamics with recognition that “the dynamic is, a lot of times it can like, make or break the group” (Secondary Interview Data, p. 64). They all discussed the issue of higher level students either being held back by, or not wanting to be grouped with, lower level students and the concerns higher level students had about how a collaborative task would be assessed. As Jerry explained, “some of the more proficient stronger students, shall we say, or more confident students might feel a little bit aggrieved if their marks are not what they feel they would have achieved, had they done on individual assignment” (Interview Transcripts, p. 177). One of the factors which could add to that grievance was a widely discussed phenomenon, that of free-loading or free-riding.

3.5.6.3.4 Free-riders and high achievers

A clear expectation of the participants was the likelihood that there could be free-riders in the class, meaning students who contribute little to the collaboration and get the benefit of others doing the work. This phenomenon is well-recognised in the literature (see Gaur & Gupta, 2013; Levin, 2003; Maiden & Perry, 2011; Strijbos, 2016) and also the participants’ lived experiences. As Dorian said,

there are certain kind of systemic issues with particular work because it works on the assumption to a certain extent that everybody will pull their weight. And I mean, I'm sure I don't need to tell you that, that that's obviously not the case. (Interview Transcripts, p. 22)

This ties with Van’s argument against collaborative assignments that “you’re allowing these free-riders who can’t do it individually, allowing them to piggyback on on their classmates” (Interview Transcripts, p. 268). Lorenzo also acknowledged free-riding was a problem but saw it as developmental since the students would have to “negotiate these difficult situations which will arise I think, in their future as well” (Interview Transcripts, p. 236). However, he had also seen a free-rider bring down the performance of the group since their non-contribution “led to sort of a general malaise or depression

in the group in general” because “knowing that at least one student was probably not going to do very much seemed to lead to failure before it even got started” (Interview Transcripts, p. 239). Although concerned about free-riding, Kyle argued students that did the work in a dysfunctional group still learnt from the process and would be able to apply that learning more successfully in later assignments (Interview Transcripts, p. 204), benefitting from the dry run nature of the collaborative task.

When considering what could motivate free-riding, Ava made a link to lower language levels, and “students with lower proficiency tend to be left behind in the group, they tend to do less work. Some of them tend to be free-riders” (Interview Transcripts, p. 43). Kyle had an alternative take on things, making the association to free-riders when discussing Model A, wondering if solely assessing the product could motivate students toward free-riding as they would get the same grade regardless of the individual effort they made. Wendy also made the connection to assessment, and the principle of fairness.

Inevitably, you know, in any group, there's going to be one lazy person, there's going to be one driven person who wants it to be as good as it can possibly be. And there'll be the kind of the, you know, the, I'll just go along with whatever kind of person and it's like a group dynamic thing that happens, and I think it happens in all groups. You know, you've always got the one that doesn't pull their weight, and what do you do with them? And it comes down to the other students to have to try and deal with it. I just, and all of that's fine, in the kind of personal development sense. But again, should that be something that then impacts on a grade? If, if all of that happens, and then they're graded on a piece of writing? Is that fair? (Interview Transcripts, p. 284).

In terms of dealing with the issue, Fay employed multiple collaborative assignments during a course and she felt that natural selection would help address any free-riding. As she put it,

And one of the questions that I'm always asked when I present this idea of collaborative work is, you know, what about free-riders? You know, those you have. And yes, there always will be free-riders, I imagine it's sort of human nature or nature in general. But because it's not a one-off project, but a series of projects, you might have a free-rider in the first one, and then nobody wants to work with him or her the second time over, so then there was a natural process of people either shaping up or being thrown out of this process. (Interview Transcripts, p. 117)

Dolly, in contrast, looked to take pre-emptive action, addressing free-riding with the students early in the course, explaining “you have to anticipate all kinds of problems, and then sort of like lay it all out, like try to quash kind of squash them before they even become problems,” (Interview Transcripts, p. 103), with the latter statement encapsulating an important aspect of the defensive orientation. However, she was also very aware there was little a tutor could do to stop free-riding since “they're more likely to get pressured into it from their friends than they are from you anyway” (Secondary Interview Data, p. 52) and it was something students would just have to learn to deal with. Ava also looked to address the issue early, explaining to students she would be able to help with a potential free-rider during the process but it would be too late to address if groups waited until the end to inform her (Interview Transcripts, p. 47).

With language level being a potential contributing factor to free-riding, participants anticipated stronger students would take the lead and operate as a kind of quality assurance of the final work. Fernand noted that in the cooperative assignment he had experience of, where students had principle responsibility for one section of the essay and where the introduction and conclusion had to be written collaboratively “what normally happened was the stronger students would say “I'll write that” (Interview Transcripts, p. 131). This was also seen by Laura, when she was talking about higher level students and observed “when and where necessary, that student intervenes, and does the necessary edits” (Interview Transcripts, p. 224). She went on to discuss a student who had completed 80% of an assignment as he knew the other students were laid back and would not do it (Interview Transcripts,

p. 227) and an extreme case where a student came to submit the essay and elected to delete the work of the other two students in the group without informing them, explaining, when asked, “You know, they were, they were very poor in terms of English and lexis. And I decided to erase them” (Interview Transcripts, p. 228). This concern about assessment leads into consideration of high achievers and their reaction to collaborative assignments. Fernand would consider collaborative assignments harder as “they would much prefer to write an essay on their own” (Interview Transcripts, p. 43) and recounted

they're not going to take a chance of some student producing basically a piece of crap that's going to drag the whole essay down, you know, one third of it's going to be completely incoherent. So they're gonna write for them and that has happened. We know basically that that goes on. (Secondary Interview Transcripts, p. 97)

Barry's experience mirrors this, with him also describing situations where “if they're allowed to kind of work in a group but work unchecked, they've bound themselves together in a group of four where one very strong person can be the dominant voice and the others can coast” (Interview Transcripts, p. 13). Humberto also recognised the potential of a stronger student dominating felt “there is nothing intrinsically wrong with one person dominating if the other person is learning alongside” (Interview Transcripts, p. 164), although this sympathetic reading of the phenomenon was an outlier. More widely it was seen that, to get a good grade, high achievers may circumvent the collaboration.

3.5.6.3.5 Circumventing collaboration

More able students taking the lead and assuring the quality of the final work is not the only way in which participants thought students could circumvent collaboration. Students taking a cooperative approach and producing disjointed work was also discussed. Ava noted a lack of cohesion in some student work, and one extreme case where “in the body paragraph, they would, they would have one

attitude that they would negate in the conclusion” (Secondary Interview Data, p. 13). She also saw evidence of cooperation in presentations where

it's easier for them to just meet up at the very beginning, divide up the work and say, okay, you do the introduction, and you do the body, you do the conclusion, etc, etc. And then it turns out to the individual projects, eventually because each of them would have one part, and all, all they're doing is just to introduce the next person, so they don't always know what each other are doing. (Interview Transcripts, p. 43)

which was also the case for Bertha, who saw

that students seem to produce something, like, you could tell that belongs to that student that belongs to that one and that belongs to that one. So it was something like three individual sort of like, stuck together, rather than one thing that was seamlessly worked on collaboratively (Interview Transcripts, p. 56)

This is not an unusual observation, with Kyle similarly experiencing presentations where “it was almost as if each of the students were just giving an individual presentation about a different thing. There was really no element of them ever having sat down as a group and worked out” (Interview Transcripts, p. 214). Wendy also had experience of disjointed presentations, describing them as a “Frankenstein presentation” (Secondary Interview Data, p. 284), which is a particular danger when students work without supervision. She also saw the breakdown of collaboration in written work, where lack of communication was the issue and, in an online group writing task, she observed “them all writing these different sentences, and it was just a mess, but they weren't talking to each other” (Secondary Interview Data, p. 285). Van was the sympathetic ear in relation to students taking a cooperative approach, explaining “it's still a collaboration because they, you know, they

collaboratively, collaboratively set those cooperative roles” although he did recognise “that those cooperative things, they just kind of compartmentalise, more likely to compartmentalise and take less responsibility for what other people are doing and less collaboration” (Interview Transcripts, p. 273). Kyle also saw potential for a more cooperative approach, which either the students naturally adopt or which

sometimes as a teacher, you might guide students towards that. If they're having difficulties, if they're not working very effectively as group and it's a case of needing to get it done, it's easier to say you do this, you do this, you do that. (Interview Transcripts, p. 206)

Even within assignments that are clearly set up collaboratively, teacher intervention may not be enough for students to truly work collaboratively, with Cristobal trying to discuss students having responsibility for the whole essay and being told “Oh, that, em, that's Fatima's, Sir. It's not mine. You need to talk to Fatima about it” (Secondary Interview Data, p. 53). Laura counselled students against taking a cooperative approach, stressing the value of collaboration, but encountered a group of strong very able students who insisted on a cooperative approach, which ultimately led to their work having problems (Interview Transcripts, p. 223). Students adopting a cooperative approach in a collaborative assignment which, crucially, will be *assessed* as a collaborative assignment, is problematic and could potentially invalidate grades.

3.5.6.3.6 Grade focus and performative collaboration

Taken in sum, these expectations help shape models of assessment for collaboration and the need to assess the process in addition to the product. Process assessment offers a mechanism to mitigate free-riding and the circumvention of collaboration, by putting the process of collaboration under the spotlight. However, for practitioners, grades carried further expectations. A noticeable expectation was, as Barry noted, that “students are only interested in the bottom line” (Interview

Transcripts, p. 14), a sentiment shared by Fernand, who had observed that “they're just marks oriented all the time. It's like, what marks am I getting for this? So, you know, once they hear they're getting marked for it, then they'll make an effort” (Secondary Interview Data, p. 104). Olga was also very aware of students only being interested in “the grades they get, that's all” (Interview Transcripts, p. 253) and Bertha saw this as a means of focussing students as “the students take it more seriously, when they know they're being assessed on it” (Secondary Interview Data, p. 29) and talking about learner diaries she felt that “the students need to know that that's assessed as well or else they wouldn't bother keeping it” (Interview Transcripts, p. 68). Cristobal also recognised that “to get students to take anything seriously, you have to give them marks for it” (Secondary Interview Data, p. 41). Wendy too felt “students are often very grade driven. That's all they care about. You know, it'll be, what's my mark, first. Before looking at any of the comments usually” (Interview Transcripts, p. 293). Remember, these are practitioners working in different institutions in different parts of the world, suggesting this grade focus is not culturally bound. With an alternative viewpoint, Fay contested that grades could be counterproductive because “it all becomes for the grade instead of for natural communication and collaboration” (Interview Transcripts, p. 117). Dolly had a more negative take on this, suggesting that when grades came into play, students could become more competitive (Secondary Interview Data, p.74), rather than collaborative.

The student focus on grades leads, in turn, to an argument that the process of collaboration and collaboration itself should be assessed, to demonstrate the value of collaboration to students. As Fay explained, “the grade reflects the amount of importance the teacher gives to the process” (Interview Transcripts, p. 120). Assessing collaboration also potentially offers a mechanism to mitigate free-riding by putting the process of collaboration under the spotlight. However, if collaboration itself is assessed, that raised further expectations. Fay would go on to suggest

I think grades have a destructive power. So that something as valuable as reflection, if students know they're going to be given grade, then they will try to answer what they think I *want* and not necessarily have a *true* reflection, to have personal growth. Because the aim is different, you know, one is personal growth together with me, the teacher, whatever, another one is a grade. So of course, what will happen is very different. (Interview Transcripts, p. 126, emphasis in original)

and she was not alone in this concern that the grading could affect student behaviour. Lorenzo thought there was scope for assessing collaboration but wondered “are the students producing evidence of collaboration just because they think that's a part of their grade?” (Interview Transcripts, p. 249). Building on this, Van referenced Goodhart’s Law (Secondary Interview Data, p. 265), meaning the measure of collaboration e.g. the evaluation rubric would become the target of students’ behaviour and he felt there could be fake collaboration (Interview Transcripts, p. 278). Laura also saw a potential problem with a collaboration grading rubric, explaining

if my students have a criteria like this, some students will, will just think, you know, "Okay, if I write anything there, if I write anything mumbo jumbo, still, it will look like collaboration," you know, still will be a kind of sign off and work there. So it's about quality or quantity. (Interview Transcripts, p. 226)

Van also saw the potential for grading collaboration to increase competition between students in an attempt to be seen as the ‘better’ collaborator. This was partly informed by personal experience of a group job interview where “everyone was kind of out, trying to out smile each other and, I don't know, it just became artificial, you know. It wasn't real collaboration because of all these people around the room scoring us with clipboards” (Interview Transcripts, p. 277). Performative collaboration was also a concern for Dolly, who had found “they will they will make, they will, they

will, they will perform perform for you right in class while you are observing them collaborating. But then, you know, it's another it's another situation altogether when they're not in class" (Interview Transcripts, p. 107). Karen took a more optimistic view of "I like to think my students don't mimic, don't act" (Interview Transcripts, p. 196), although she conceded it may depend on the motivation the students have to genuinely learn something rather than get the grades for the assignment.

Another issue with a grade focus lay not with the students, but with the teachers. For Kyle, one of the sticking points in a collaborative assignment was the standardisation of markers, where considerable time and effort had to be spent standardising teachers for what was, ultimately, a small part of the overall grade of the course (Interview Transcripts, p. 204). Van took a similar line, feeling universities

they just put put in too much effort to producing too many scores with too fine a level of granularity. And it's just a waste, you know, debating, getting into arguments of whether something's a 63 or 65. When over the, overall, that's only 10% of their overall grade for the year anyway or whatever. (Interview Transcripts, p. 270)

Olga also felt a grade related tension, explaining "you not only you have to give a fair grade, you also have to evaluate the learning, but also you have to justify to the students" (Interview Transcripts, p.254). For course leaders running standardisation, teachers conducting evaluation and students, grades matter.

3.5.6.3.7 Students as assessment agents

Connected to the issue of grades is the question of who conducts assessment and the role of the students in the assessment of collaborative assignments. Whatever the value placed on student-led assessment in other types of assignment, it is a particular consideration in collaborative assignments since, as Dorian put it, "it's only the people who have collaborated that can best really

assess how they've collaborated" (Interview Transcripts, p. 27), with Bertha developing this theme when asking "Is that the teacher's role, to assess the collaboration if they weren't actively involved in the collaboration?" (Interview Transcripts, p. 64-65). Assessment being conducted by the people involved in the collaboration led to discussion of peer-assessment. While participants saw value in peers giving each other feedback, full involvement in summative assessment gave more pause, with Lorenzo's position of "I don't know how comfortable I am with the student, him or herself and a peer, you know, providing points" (Interview Transcripts, p. 248) being quite common.

For summative assessment, Ava had a system of students providing peer feedback (amounting to 5% of the final grade, which was also the maximum amount Van could see peer-assessment contributing to a final grade (Interview Transcripts, p. 279)) but was clear that for Ava it was "action/behaviour based" (Interview Transcripts, p. 51) where students were addressing objective observable behaviours, such as attending meetings or meeting deadlines. Her rationale for self- and peer-assessment is worth recounting in full

I actually observed the dynamics of groups in class when they're working in class, and also from talking to people, but I actually, I really tried to shy away from grading them based on my perceptions because I don't think that's fair in the sense that some groups are just more open, they talk to me more, I'm more involved in their work. Or some students are just more talkative. They come to me more asking more questions. But I don't think those are objective. So even though for the peer and the self, it's not objective, purely objective, either, but I think for at the, at least for the peer aspect for the survey, I get a fair amount of understanding on whether this person is actually doing work at all or not. (Interview Transcripts, p. 52).

This illustrates the struggles of practitioners knowing their limitations of objectively assessing a collaboration and acceptance that flawed assessment tools (in this case peer-assessment) can be preferable to tools which are still more flawed.

Olga could also see limited scope for summative peer-assessment, “if it’s not heavily weighted,” (Interview Transcripts, p. 259). An alternative role for peer-assessment was in it informing teacher assessment, so rather than direct peer-assessment it is more teacher-mediated peer-assessment (see Beard, et al., 1989, p. 34). This was the approach favoured by Tanya, where students’ formative self- and peer-assessment informed the final summative grading, conducted by the teacher (Interview Transcripts, p. 39) and Bertha suggested students could propose and discuss grades for collaboration with the teacher (Interview Transcripts, p. 67), with Cristobal also seeing peer-assessment as the grounds for discussion and teacher decision making (Interview Transcripts, p. 93). Teacher-mediated peer-assessment was something Dolly had employed but had gradually come to rely on it less and less as she found reading and comparing all the comments was becoming too much work. Lorenzo would have another issue with teacher-mediated peer-assessment, based on an experience he had where everyone in the group had negatively assessed one member but that particular student self-assessed positively, leaving Lorenzo to question

so how do you, how do you include that in the assessment, like, that student thinks that he did a great job, his peers think he didn't do a great job. So how am I assigning like, do I believe the student more than the peers or do I believe the peers more than then the student. (Interview Transcripts, p. 248)

He was not the only participant to raise the reliability of peer-assessment. The impact of culture and students being reticent of criticising each other was raised by Jerry (Interview Transcripts, p. 177), Cristobal (Interview Transcripts, p. 91) and Laura (Interview Transcripts, p. 228-229). Dolly did not make a link to culture, but did acknowledge

The students will talk to each other, and they will not mark another student, they will not give a low mark. So even with that case that I told you were like three people were like, this student did not the fourth student didn't do his or her work. And then their grade for that student was 85. Because they just feel, they feel bad, right? (Interview Transcripts, p. 107)

The impact peer-assessment could have on other students was a concern shared by Bertha, who seemed to value peer-assessment but admitted “my concern is always that someone will say, like, they're hard done by or that there's something unfair about it” (Interview Transcripts, p. 65), although she had also been surprised by how honest students were in their peer-assessments. Fay had also had positive experience with peer-assessment, finding students grades were in line with the marks she gave and they were not “giving each other a lot of very high grades, just to help their friends” (Interview Transcripts, p. 121). Van, in contrast, likened anonymous peer feedback and assessment to a prisoner's dilemma and game theory, and if it was part of final assessment it had to play a minimal role (Interview Transcripts, p. 279).

If engaging with peer-assessment, Fay maintained that there needed to be both student training and that “the institution should be positive toward it” (Interview Transcripts, p. 228), with Jerry adding “students would really have to be given very clear rubrics and criteria of how to do this” (Interview Transcripts, p. 178), and Bertha favouring checklists, although she admitted she was “always concerned about what the students are left feeling after these kind of tasks” (Interview Transcripts, p. 65) and she also felt that “sometimes it's difficult for students to look at their own work objectively” (Interview Transcripts, p. 66). For Cristobal, it was the need for learner training that was one of the barriers to incorporating peer- and self-assessment, explaining in his institution

trying to bring in other things like peer-assessment and self-assessment is just a bridge too far, if you like, you know, there's, we don't have time, we're, we're trying, how do we do this, let alone do anything else. (Interview Transcripts, p. 91)

Dolly also appreciated learner training could take time, and questioned its usefulness, saying “I think even if I train them on that they will in the end still feel bad. And, and would not, would not give, like an accurate assessment of the collaboration grade” (Interview Transcripts, p. 107). This may explain why she has relied less on peer-assessment, even when teacher-mediated.

Self-assessment carried many of the same concerns as peer-assessment: objectivity, honesty, reliability, need for training, impact of rubrics. However, as with peer-assessment, there were also bright spots, with Laura describing her experience with self-assessment realising “that they're really critical of themselves, and they're really honest. And they use the criteria, they mark themselves and they really mark themselves in the right way” (Interview Transcripts, p. 229). It must be noted however, that she was talking about self-assessment for formative rather than summative purposes and this positivity was counterbalanced by Olga who felt “I just can't be convinced that the grades they get will be justified” (Interview Transcripts, p. 258) since students could not understand the expectations of the assignment as fully as the teacher so were less well equipped to assess themselves. When discussing self-assessment Dolly added consideration of the Dunning Kruger effect to the other issues found in peer-assessment, explaining

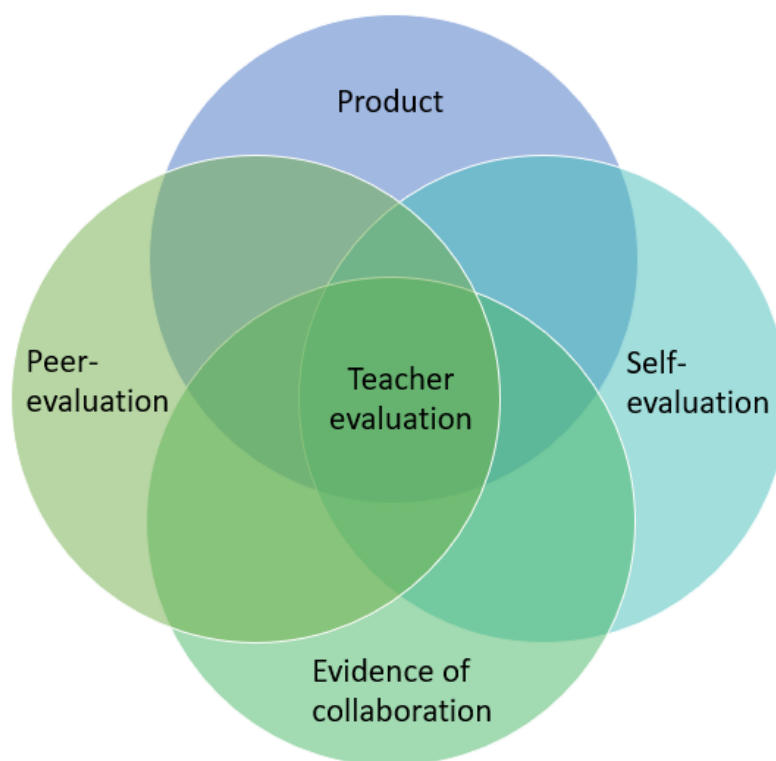
people who are novices at something, they may think, "Oh, I'm actually pretty good at this," because they don't know how good the best people are. But once you've had a lot of training, and you are an expert in the topic, you actually know how much more you don't know. So then they have a tendency to underrate their abilities. (Interview Transcripts, p. 110)

and then applied this to self-assessment, which could result in high-achievers underrating themselves and lower ability students overestimating their achievements. Gonzalo also felt some students may over estimate their achievements and others may under-rate themselves to fit with what they thought the teacher's expectations were (Interview Transcripts, p. 154). Fay felt there was a cultural dimension that could problematize self-assessment as “people in Protestant countries tend to be much more

rigorous in their self-assessment than then people in the Middle East” (Interview Transcripts, p. 122) and she also felt students may struggle to be objective and truthful when evaluating themselves.

Rather than peer- or self-assessment, for summative assessment there was a preference for teacher-led assessment, with the impact of environment clear. The participants are aware they are responsible for high stakes evaluation decisions that could have serious consequence on their students’ academic careers, recalling the importance of ‘consequence’ in classroom-based teacher assessment (Hamp-Lyons, 2007). This awareness of consequence tended to put participants in the role of ‘reluctant overlord’, where assessment was mainly within the purview of the teacher and the teacher has “the guiding hand” (Interview Transcripts, p. 18), as illustrated in Figure 3-8.

Figure 3-8 *The central role in assessment of the reluctant overlord teacher*



The use of ‘reluctant’ is deliberate as the participants did not seem fully comfortable with teacher-led assessment, acknowledging “I know that might be kind of seen as top down,” (Interview Transcripts, p. 178), “old school” (Interview Transcripts, p. 38), and “Is that a bit old-fashioned? It’s a bit old school”

(Interview Transcripts p. 294). Part of the motivation for this was as a result of student expectation that assessment is the teacher's responsibility, with Jerry explaining in relation to undergraduates, "the teacher still has that authority and usually respect from the students" (Interview Transcripts, p. 178). Laura agreed, explaining "the teacher, not always, but usually, is seen as the authority as the person who has more information than the others. Students, in my context, expect the teacher to be the authority and to mark those papers" (Interview Transcripts, p. 228). This fulfilling student expectations was also conveyed by Wendy, who felt "the students want the teacher's kind of final word on is this right or wrong as well" (Interview Transcripts, p. 294), and who also argued that the teacher had the expertise and knowledge to be able to assess, for example, features of successful academic writing and language accuracy (Interview Transcripts, p. 294). This point is reflected in Bertha's experience of students rejecting the value of peer feedback, telling her, "Why am I bothering to ask, you know, my classmates for feedback. Sure, they don't know any more than I do" (Secondary Interview Data, p. 36). Van concurred with Wendy, arguing "the teacher is the expert and it's got to be, part of the, the years and years of expertise the teacher builds up" (Interview Transcripts, p. 279).

The concern about students' ability to self-assess are also mirrored when it comes to using reflection as an assessment lens, as previously discussed in 3.3.3.2 and as illustrated by Karen's concern that "how do I really know that they've done what they've said they've done" (Interview Transcripts, p. 190). This problematizes reflection as an assessment lens due to the possibility of the students adopting the role of unreliable narrator and either sharing what they thought happened (rather than what actually happened) or simply telling the assessor what they assumed would be graded most favourably.

3.6 Defensive avoidance, acceptance and adaptation

The defensive orientation had an impact on how participants conducted, or would conduct, collaborative assignments. There were three main types of impact this orientation triggered: defensive avoidance, defensive acceptance, and defensive adaptation. Avoidance can be seen in

participants not taking on certain assessment lenses due to potential problems employing them. This is captured in the quote from Fay below, when discussing utilising measures to look inside a collaboration.

It might be possible. And yet, I don't know if this is where I would like to spend my energy and resources. Again, this does not happen in the real world. The collaboration itself is NOT evaluated. (...) If there was an easy way that was transparent enough and fair enough I would, you know, love to also have that aspect. But I can't see how because of what you're saying, yes, it's, it's SO complicated and so complex and multi-level, you know, that it seems like an impossible job that would need so much that it might not be worth it.

Here we can see the practical orientation overcoming the fairness orientation, leading to avoidance of attempting to directly assess collaboration itself. Avoidance can also be seen in Fernand's preference for collaborative assignments not to take place in students' first semesters but later in their academic careers, when they "have found their feet a little bit" (Interview Transcripts, p. 138).

Acceptance is evident in participants who are aware that while the way they assess collaborative assignments may be imperfect, there is little more they are able to do to improve things with Dolly recognising she has to "make a compromise with myself" (Interview Transcripts, p. 109) and incorporate elements into an assessment scheme she was not fully happy with. Acceptance is also seen in Cristobal rationalising a product-focused assessment approach with "Well, I can't do everything" (Interview Transcripts, p. 84-85).

Adaptation develops through experience as Wendy outlined:

I think the more you, the more you see different ways of doing it, that hasn't necessarily come from me, as though, that I've set up differently, that the different things that the students bring to it, and the more you can understand what might work well, or what might not work well, or how you can help them because you've seen these people doing this well. So actually, you can build that into your instructions next time or into the task brief or something. (Interview Transcripts, p. 286)

Adaptation can also be seen in Dolly's desire to "vaccinate my students in advance" (Interview Transcripts, p. 103), meaning make students aware of various pitfalls of collaboration and how to avoid them.

3.7 Pertinent findings of the companion dissertation

Assessment does not exist in isolation, and nor does this dissertation. It is part of a wider study with the companion dissertation (Bolster, 2025) which looks more deeply at EAP practitioner beliefs and attitudes to collaborative assignments generally. Some of the findings of that dissertation also have implications for assessment, with the most pertinent being: 'jagged expertise', 'collaborative dissonance', 'collaboration ripples', 'experience as training', 'witnessing it work', 'discipline deficit' and 'practitioner-led pedagogy of collaboration' (Bolster, 2025).

The duality of 'jagged expertise' of EAP practitioners, being skilled in some areas but relatively novice in others, is certainly applicable to collaborative assessment. Collaborative assessment is an area of uncertainty where the participants recognised they lacked expertise. So too is 'collaborative dissonance' evidenced in the findings of this dissertation, with participants holding contradictory views about collaboration. This can be seen in participants being sure collaboration should be assessed if it is a learning outcome and yet being unsure how this could be done. It is also visible in the belief oral collaborative assessment is fairer than written, despite acknowledging all the potential problems of a written assignments could still be apparent in an oral one. The concept of collaborative dissonance

also recalls the tensions around collaborative assignments for the I-in-teamer, who wants the students to engage in collaboration but also has the desire to be able to differentiate between students and reward them for their work.

Bolster's (2025) 'collaboration ripples' could also certainly impact collaborative assessment identity. If a teacher had a strong negative experience of collaboration when they were students, it might strengthen the defensive orientation through awareness of what could go wrong and how damaging that can be, with Wendy being a case in point. In contrast, a positive personal experience could make teachers more amenable to collaborative assignments, as was the case with Fay. It is also clear that, when it comes to assessing collaborative assignments, 'experience as training' is very much the case. Bolster (2025) identifies experience as the main way teachers learn about teaching collaboration and this can also be seen in assessment, where practitioners lack formal training in collaborative assessment and come to understand how to approach it through the experiences they have with it. This accounts for the shifting collaborative assessment identity as each experience may encourage a different orientation. We can see this in Lorenzo becoming more defensive after experiencing a student group breaking down during a collaborative assignment, making him more wary of them.

It is Bolster's (2025) 'witnessing it work' that has the most resonance with the practical output of this dissertation however. Conducting collaborative assignments successfully encourages teachers to engage with them further. Sharing experiences with other teachers helps the development of a 'practitioner-led pedagogy of collaboration', where through sharing experience, expertise develops. The work of this dissertation can be a contributor to the dialogue that should take place around collaborative assessment, encouraging teachers to engage with the challenges and come to rationalised and justified compromises to develop fit for purpose collaborative assessment schemes. In light of 'discipline deficit', where student collaboration is felt by EAP practitioners to be under-supported in faculty courses, it is vital such discussions take place, allowing for phronesis informed solutions.

3.8 Evaluations of the starter guide to assessing collaborative assignments

In order to contribute to those discussions to develop a practitioner-led pedagogy of collaboration, selected participants were asked to review the starter guide to assessing collaborative in terms of need, coverage and usefulness. The first draft of the starter guide was reviewed by Anna and Claudette, who attended a presentation of the findings of this dissertation at the NFEAP 2024 conference in Oslo, and Elsa, who was a contact made during discussion of an EAP conference in the UK in 2025. Their feedback and comments led to editing of the guide, with the second draft being reviewed by three of the original participants: Bertha, Jerry and Lorenzo.

3.8.1 Evaluations of the first draft of the starter guide

Anna, who works in a university in Norway, was the first respondent and echoed a sentiment which recalls how useful the participants found taking part in the interview, saying she enjoyed reading and thinking about the guide. In terms of need she felt “There is definitely a need for something like this to help practitioners really think through assessment practices and learning outcomes before they assign group projects,” (Reviews of the starter guide, p. 4) and she was very favourable toward it, explaining “as a guide that can help determine both learning outcomes and course design and helping think through some of the issues involved, not just in terms of practicalities, but pedagogical principles this works well” (Reviews of the starter guide, p. 4). Claudette, who works in a university in the UK, was the other NFEAP presentation attendee to review the first draft of the starter guide and also commented on enjoying reading it. Her evaluation of the need for the guide is worth recounting in full, as it captures the essence of this dissertation.

There is a clear and growing need for such a guide. As collaborative assignments become more common in higher educational settings, especially in courses focused on EAP, educators require structured methods to effectively assess these assignments. The guide addresses this need and provides a comprehensive list of strategies and techniques to evaluate students' collaborative efforts, which are often very challenging to measure fairly and accurately. (Reviews of the starter guide, p. 7)

Collaborative assignments happen, their assessment is challenging, and teachers need help, which the starter guide provides.

In terms of the usefulness of the guide, Anna felt, “this would be very useful as an awareness raising tool and a great resource for discussion among colleagues and teams of teachers” (Reviews of the start guide, p. 5). Claudette also felt that that it would “be highly useful for EAP practitioners involved in assessing collaborative assignments. It provides practical tools and frameworks that can help educators develop fair and effective assessment schemes” (Reviews of the starter guide, p. 7). She particularly appreciated the sample questions and prompts “as they offer ready-to-use resources for immediate implementation” (Reviews of the start guide, p. 7).

In terms of coverage, Claudette felt the starter guide “offers an extensive range of contents that cover key aspects of collaborative assessment” (Reviews of the start guide, p. 7) although she felt more examples and case studies could help illustrate different parts of the guide. She also thought there could be more coverage of dealing with “students' varied language skills (esp. Lower IELTS) and cultural predispositions/values about collaboration” (Reviews of the starter guide, p. 7). Anna, in relation to coverage, also raised the issue of student attitude, influenced by educational culture and language competency, thinking students might need to be persuaded to take a collaborative approach as they feel a cooperative ‘task-splitting’ approach reflects real-world teamwork in industry. She also raised an interesting take on performative collaboration, considering students hiding group dysfunction from the teacher as another example of this.

Elsa, who works as the English language centre manager at a university in the UK, and was contacted through discussion about a forthcoming conference, also reviewed the first draft. She felt

the guide provides a great roadmap for EAP practitioners considering introducing a collaborative task and provides some very practical pointers and suggestions. I think is clear and accessible, and does not shy away from the key challenges – providing sufficient scaffolding and support for the learners and the workload implications for the teacher. (Reviews of the starter guide, p. 8)

The opening vignette was found particularly useful for illustrating the challenges in assessing collaborative assignments, but did reflect upon reading the guide that “I felt that if I was an EAP practitioner considering a collaborative assessment, I might feel that some of the pitfalls were insurmountable!” (Reviews of the starter guide, p. 8). She also clearly articulated why the way we prepare and facilitate collaborative assignments matters when commenting on why the guide was useful, explaining

the guide is very useful in terms of highlighting the need for careful consideration prior to the introduction of an assessed collaborative task. In my experience the preparation stage is often given too little time and attention and students who have not had the time to consider the nature of collaboration and how to negotiate the challenges successfully will quickly resort to any strategy that will get them through the assessment and out the other side. Unfortunately, this often has a negative impact on the learning experience and negates so many of the benefits of working collaboratively. Perhaps providing some clearer guidance for the practitioner in terms of resources they can use to prepare their students might be helpful? (Reviews of the starter guide, p. 8-9)

This emphasises the need to put careful consideration into the start of a collaborative assignment so students can be well prepared for the road ahead, particularly given “EAP practitioners are often working with groups of students who may have very different ideas about and experiences of group work” (Reviews of the starter guide, p. 8). She valued the sample assessment lenses provided in the guide and another important aspect she highlighted was developing a shared conception of collaboration, which she felt needed more exploration “to drill down into what we might mean by ‘successful collaboration’” (Reviews of the user guide, p. 8).

Outside of the guide itself, her response also spoke to other issues related to this dissertation. She discussed the choice not to use collaborative assignments on pre-sessionals due to the need to attest to individual language level and how they were used more on in-session courses where the focus was “more on developing a wider range of skills that will enable students to perform better on their degree programmes” (Reviews of the user guide, p. 8). When using such assignments, she explained how the group work produced got a nominal grade, with most of the grade

being awarded to an individual reflective essay in which students discuss (with reference to relevant theory) the challenges and success they encountered during the group work (collaborative) element of the task, and what they have learned through the process. (Reviews of the user guide, p. 8)

which reflects the task type of assessing awareness of collaboration more than the collaboration itself. She also recommended potential resources that could be included in the guide to help the user better understand how to set up and facilitate groupwork.

As a further comment, Anna felt the user might need help was some guidance on how to use the guide, and the purpose of some of the information e.g. how the collaborative assessment archetypes could be used. She also raised an important point in that the reader needs to know what the guide is based on. As she explained

I realize it is not a research article – and that it is intended as a practical guide in practitioners' very busy schedules, so I am not suggesting including literature reviews, tons of references, etc. But I think it would be important for the readers to know that the guide is not simply based on your experiences or beliefs about what works, but on reading the literature and collecting data. (Reviews of the starter guide, p. 5-6)

The PhD behind the guide had deliberately been backgrounded to keep the guide accessible and guard against it becoming too long and unwieldy. However, it was an important concern to address so for the second draft a foreword was added, outlining the GT research behind the guide. The guide was also clearly sectioned into three parts, moving from the theoretical to the practical, with Part 1 being expanded slightly to cover the key principles and guiding orientations of EAP practitioner collaborative assessment identity. The potential impact of student culture was also raised, highlighting the need to orientate students into, and support students through, collaborative assignments and the evaluation scheme based on collaborative assessment identity was also added. This resulted in a second draft of the starter guide (Appendix 13), which was the version sent to original research participants for their review.

3.8.2 Evaluations of the second draft of the starter guide

Although the starter guide was edited based on feedback on the first draft, note that the second draft of the guide is slightly out of synch with the collaborative assessment algorithm proposed in this dissertation as the first two layers of the algorithm are amalgamated within the third layer of 'readiness' in the starter guide. This is because the algorithm was still under development and refinement at the time the guide was revised. The reviews of the original participants will be discussed in turn.

3.8.2.1 *Bertha's review (Reviews of the starter guide, p. 9-10)*

Bertha felt there was a clear need for the guide, particularly one “with a practical focus but with theoretical underpinnings” since such a guide “was definitely an omission within the field, especially the level of detail, and the thought process an EAP practitioner is likely to go through in approaching a collaborative assignment.” In terms of coverage, she felt the guide

got the level of detail just right. It's detailed enough, but not too much. There are real-life, tangible scenarios; guidance on how to think the process through; all the aspects and perspectives that need to be considered, which are comprehensive

although she concurred with Elsa in reflecting “if I had known all of this, would I have continued with a collaborative assignment, seeing all these complexities and considerations laid out?” Looking back at collaborative assignments she had conducted in the past she recognised “that I didn't realise all the thought that should have gone into collaborative assignments I have done in the past” so she suggested a useful addition to the guide would be some more advice on how to get started on collaborative assignments. As a means of doing this she suggested starting with a collaborative oral assignment like a group presentation since that could be “more accessible and manageable (and low stakes) as a starting point” than collaborative writing. This reflects the general greater level of comfort practitioners have with collaborative speaking rather than collaborative writing tasks as there is the clear and obvious superficial differentiation of oral fluency built in.

Further additions she felt could be beneficial were “a set of marking criteria/ a template to guide a practitioner on coming up with a rubric which integrates things like language considerations/ writing skills with collaboration” with the weighting of that being of particular interest. A diagnostic quiz could also be added to help the user identify their collaborative assessment archetype and more guidance could be given on how to use the evaluation criteria provided on p. 36 of the guide. This brings to mind the need to clarify the place of criteria in forming the basis of discussion and negotiation between

teachers rather than setting out the absolute truth (Davidson, 2004). So too did she feel the idea of positive washback could be explained in more detail. On the whole, however, in terms of usefulness she felt

The guide is definitely useful in terms of guiding the process, being accessible in terms of practicality and readability (the tone is just right) and the links to further literature. I found the scenarios very useful, as well as the sample reflection/ tutor interview prompts (p. 34), and the self-assessment/ peer-assessment prompts (p.36)

and closed the feedback saying “I think you have done a brilliant job on creating something with real-world value, that was definitely a gap, through your PhD!”

3.8.2.2 Jerry’s review (Reviews of the user guide p. 12-13)

Jerry also saw a need for such a guide, with the need extending beyond EAP to HE more generally. As he explained, “I think this should be extended beyond EAP, although, obviously that’s your focus for this particular research. But I don’t really see why this is disaggregated from other disciplines, really, I think it’s for all disciplines” and he felt it could be delivered through some other medium, “like a podcast or annotated PowerPoint presentation”. This was prompted by consideration of the audience since “, teachers, especially when it comes to guides, they do want quite snappy, you know, to the point, brief”. Counter to that, he also felt the guide could be extended to book length, to become part of a comprehensive toolkit for collaborative assignments. When speaking of the usefulness of the guide as is, he felt it was very useful, particularly for relatively new EAP practitioners. However, he added the caveat that “I can imagine [assessing collaborative assignments] is really quite a minefield, and probably not done very well, although I’ve been teaching for 25 years, and I’m not so sure I do it very well now, either”, suggesting the guide could be useful for any EAP practitioner.

In terms of coverage he felt it was “all necessary and good”, particularly the sample assessment lenses at the end of the guide. Where he felt coverage was lacking was dealing with grouping, particularly in light of “EDI, equality, diversity and inclusion”. Speaking specifically of the UK context he thought “a guide of how to set up groups and particularly to ensure diversity in groups, in terms of genders, in terms of linguistic ability, and taking into account anything like student support plans” would be particularly important so all students could be catered for.

3.8.2.3 *Lorenzo’s review (Reviews of the user guide p. 14-15)*

Lorenzo felt the guide was “clear and concise, and also provides a lot of useful information on the topic of collaborative assessments”. Speaking of need, he felt the guide “may be needed by programs that are not currently using collaborative assessments and would like to integrate them but may be unsure how to do so effectively.” He also considered the guide useful, explaining “I think this can help teachers, particularly new EAP teachers (or really any teachers) with collaborative tasks and assessments since “it gives practical ideas, different perspectives, guidance on assessment criteria, and examples.” While he felt the coverage was comprehensive, he felt adding case studies or examples to demonstrate the advice in use could be beneficial and also thought generative AI could be discusses given its impact on HE assessment more widely.

3.9 Chapter conclusion

This chapter has discussed the main findings from analysis of the study data. It opened with the participants’ conception of EAP and collaboration, which are in line with the discussions about the nature of EAP and collaboration in the literature, albeit the participants’ conceptions of collaboration are organic, rather than product of formal learning or training. This is significant as it speaks to a well-informed set of participants which, in turn, highlights the lack of knowledge and confidence EAP practitioners have in the particular area of assessing collaborative assignments. Although well-versed professionals, collaborative assessment remains an area of uncertainty.

The chapter moved on to specific discussion of the four models of assessment in the pre-interview task. Somewhat surprisingly for the researcher, this indicated that there were sound justifications for the three main approaches to assessing collaborative assignments: group product only, group product and individual complementary assignments, and group product and the collaboration itself. What is more striking, is that there were also sound justifications *against* different approaches. This underscores the complexity of summatively assessing collaborative assignments and also the need for any assessment scheme to be bespoke to a particular course or institution, dispelling the concept of a general use assessment framework that was initially envisaged when embarking on this dissertation

After review of the evaluation of the four models of assessment the chapter turned to the key principles and guiding orientations shaping practitioner attitudes and approaches to assessing collaborative assignments. Fairness emerged as the strongest principle, allied with transparency and simplicity. While these diverge from traditional assessment principles (e.g. validity and reliability) they held sway and helped in shaping the guiding orientations of the participants.

These orientations were practical, reward and defensive. The first two speak to the desire practitioners had to help students in their learning and ensure positive outcomes, fulfilling the principles of simplicity, transparency and fairness. The major orientation was the defensive orientation, which is not a negative conception but rather one of awareness and preparedness. Practitioners carried considerable expectations around student collaboration, incorporating personality, educational maturity, language level, free-riders and high achievers, the circumvention of collaboration, grade focus, performative collaboration and what may happen if students take on the role of assessment agents. Due to this orientation, practitioners avoid or accept problematic aspects of collaborative assignments or adapt to overcome them.

The findings of the companion dissertation also help illuminate the findings and work of this dissertation. In heed of its call for a practitioner-led pedagogy of collaboration, a starter guide to assessing collaborative assignments was developed and it was reviewed by three participants new to

the research and three of the original participants who had taken part in earlier rounds of data collection. The need and usefulness of this guide was apparent to the reviewers with suggestions for various additions and extensions to make a guide that fully supports EAP tutors working with collaborative assignments. The need for such a guide, however, is not limited to EAP. As Jerry contended, “I think beyond the EAP practitioner, I think it would be useful to any academic or tutor who is working in teaching and learning with collaborative assignments” (Reviews of the starter guide p. 13).

The next chapter will see how these findings come together to inform practitioners collaborative assessment identities and will propose a collaborative assessment algorithm, which will outline the thought processes and decision trees it will be necessary to navigate when developing a collaborative assignment for summative assessment purposes in EAP.



CHAPTER 4

DISCUSSION AND THEORY GENERATION

4 DISCUSSION AND THEORY GENERATION

4.1 Chapter overview

The previous chapter drew from the data to identify the major concerns and the thought processes EAP practitioners went through when considering how to evaluate collaborative assignments. It identified three key principles, allied with three guiding orientations, developing the basis of the theoretical output of this dissertation: collaborative assessment identity.

This chapter will explore collaborative assessment identity, demonstrating the relationships between the key principles and guiding orientations and how they work together. This will lead to discussion of the collaborative assessment archetypes, which embody this identity. These archetypes are important to recognise as they speak to the varying attitudes practitioners may bring in to collaborative assignments and their major concerns when conducting assessment of such assignments. This has implications for course leaders and teaching centres when implementing and rationalising a collaborative assignment assessment scheme to teachers, who will in turn rationalise the assessment schemes to students. These archetypes lie on a continuum from the Never-everers, who are opposed to collaborative assignments, to the theoretical construct of the All-seeing-eye, who is fully comfortable and confident in assessing collaborative assignments. The aspiration is that through professional development and dialogue between peers, practitioners can become Best-I-can-doers.

To aid with that, and as the practical output of this dissertation, the chapter will propose a collaborative assessment algorithm, which takes practitioners through the decision-making process of designing an assessment scheme for collaborative assignments. This is a five-layer process, involving consideration of the motivation for adopting a collaborative assignment, the purpose of the assessment, the readiness of students and teachers to engage in such an assignment, the selection of assessment lenses and the evaluation of the potential assessment scheme. The final evaluation step illustrates the connection between the theoretical and practical outputs of this dissertation, with the

dimensions of collaborative assessment identity providing the basis for an evaluation scheme to test collaborative assignment assessment schemes, looking at them through the criteria of simplicity, fairness, transparency, reward, practicality and defensibility.

The dissertation opened with a vignette, exemplifying the complexity of collaborative assessment in extended collaborative assignments in EAP and the collaborative assessment algorithm is applied to the vignette, demonstrating how different assessment lenses could have been utilised to aid with assessment. The collaborative assessment algorithm was also actualised as a starter guide to assessing collaborative assignments, developed as an accessible, teacher facing text, which could be used to assist teachers in the assessment of collaborative assignments. The evaluation of the guide by select participants will also be discussed in this chapter, demonstrating it is necessary and useful both within EAP and, potentially, across HE more widely.

Finally, the work of this dissertation will then be evaluated, first on a face validity level, then through the classic GT lenses of grab, work, fit and whether it is modifiable (Glaser, 1978). This leads to the main evaluation with the constructivist criteria of Charmaz (2008): credibility, originality, resonance and usefulness.

4.2 Collaborative assessment identity

In the original conception of this research, identity was not an anticipated area of discussion, but it was to identity the data led. Through the data analysis and development of the Findings, it became increasingly clear that the participants were bringing a lot of themselves into the assessment of collaborative assignments. This is not unexpected, but the interplay between beliefs and practices and the potential for a gap between the two, as identified by James and Pedder (2006), took on particular resonance, as did Brown's contention that conceptions about assessment influence practice (2011). What we believe about assessment impacts the way we approach and enact it. When conceptualising teacher assessment identity, Looney et al. (2018) closed their paper saying

teacher assessment identity encompasses not only a range of assessment strategies and skills, and even confidence and self-efficacy in undertaking assessment, but also the beliefs and feelings about assessment that will inform how teachers engage in assessment work with students, and focuses not simply on what teachers do, but on who they are. (p. 456-457)

This quote encompasses the dimensions of teacher assessment identity, making the connection between teachers' knowledge and beliefs and their subsequent engagement with assessment. Looney et al. illustrated teacher assessment identity as overlapping discs (see Figure 4-1), giving a broad stroke view of what shapes a teacher's assessment identity, incorporating knowledge, beliefs, feelings and experience. It became clear during the data analysis that the same way that in the way teachers hold an assessment identity, so too do they hold a collaborative assessment identity, influenced by the unique features of collaborative assessment. While EAP practitioners could hold a high degree of confidence in their assessment abilities in one sphere (i.e. several of the participants were experienced IELTS examiners), they could feel very differently about assessing collaboration meaning we can more properly talk about assessment identities rather than identity.

Figure 4-1 *Reconceptualising teacher assessment identity (Looney et al., 2018, p. 456)*

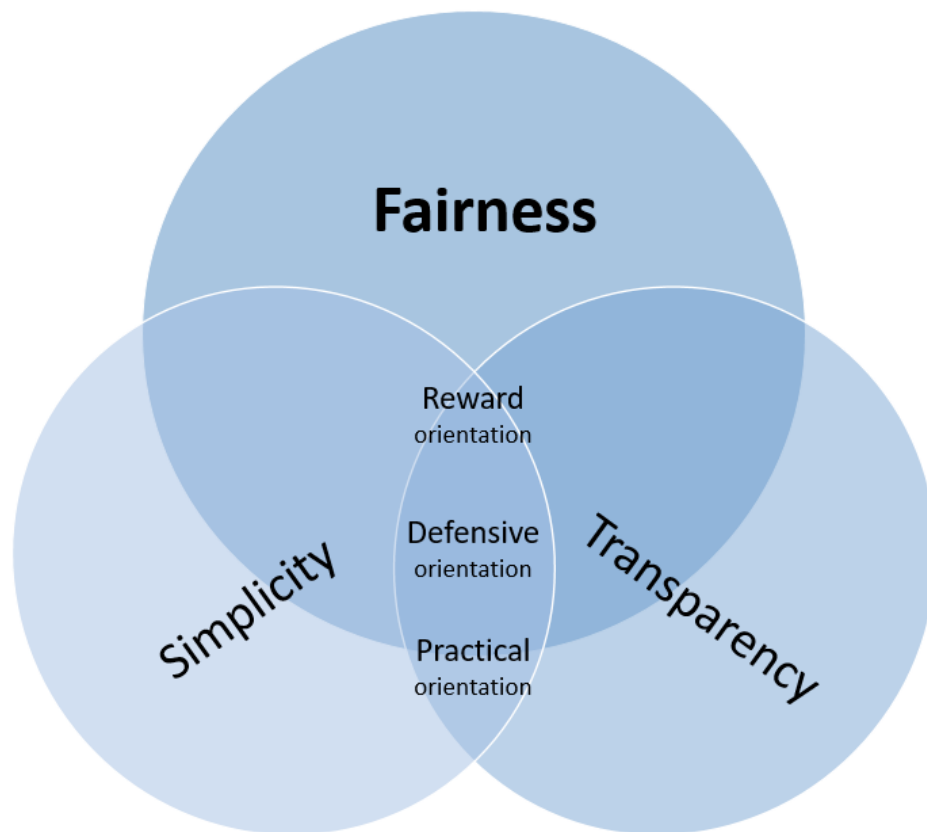


The components of teacher assessment identity are still apparent in collaborative assessment identity in that it is influenced by knowledge, beliefs, experience and feelings. However, when diagramming the collaborative assessment identity of EAP practitioners, more specific dimensions are brought to the fore. As discussed in the Findings, these are the three overlapping key principles of fairness, transparency and simplicity, which, in turn, impact the three guiding orientations and together encapsulate the collaborative assessment identity (see Figure 4-2).

Fairness, as the most widely discussed and strongest principle, is illustrated as larger and bolded, with significant overlap with transparency and simplicity. Assessment schemes being transparent to students is an aspect of fairness (as recognised in the GLAID framework (Hei, Strijbos et al., 2016)) and having assessments that are not over-complex and convoluted aids in transparency and also speaks to fairness. While there is a synergy and overlap between the principles, transparency and simplicity are not fully subsumed in fairness. Although they often work in concert, they can also operate in opposition. A product-focused assessment is the simplest approach to assessing a group assignment. It is also, arguably, the least fair means of assessing collaborative work, particularly if it is the collaboration itself which is being assessed. The rationale for an assessment scheme, regardless

of its fairness, could be made fully transparent to students, or could kept opaque. In this way, the three principles have to be seen as distinct.

Figure 4-2 *The overlapping key principles and embedded guiding orientations on EAP practitioners for collaborative assessment*



It will be clear that the classic assessment principles of validity and reliability do not feature in the visualisation of collaborative assessment identity. Validity is not explicit in

Figure 4-2 but, as Herman and Cook (2019) contend, fairness “is a fundamental validity issue: a fair test provides valid information for a given purpose for all students” (p. 244). In this quote, fairness is seen as subservient to validity. However, returning to Walter’s (2012) discussion of the relationship between fairness and validity, the importance of fairness to the participants places this dissertation on the side of the argument of Xi (2010) and Kunnan (2000) in seeing fairness “as an all-embracing concept that *subsumes* validity” (Walters, 2012, p. 417, emphasis in original). Consequently, the

validity of a collaborative assignment depends on the interpretations of the results of the assessment being fair. This turns the Herman and Cook quote on its head with validity, fundamentally, being a fairness issue.

In the way validity is not a distinct principle in collaborative assessment identity its own right, nor is reliability. That is not to suggest reliability is not important, but it is complex given the contradiction between DA and the traditional assessment conception of reliability (see 1.4.5), the uncertainty of the stability of collaboration, and reliability coming through teacher judgement and negotiation rather than repeatability. The aim of a collaborative assignment in EAP would be that the student performs better in future assignments, negating reliability as referring to an assessment giving a clear indication of future performance.

The three principles of simplicity, transparency and fairness set the foundation for collaborative assessment identity. While all three are positive principles, dogmatic zealotry toward any of them is problematic. Full transparency, for example, could lead to fine point criteria which actively deform natural collaboration through codifying and therefore encouraging very specific behaviours. Absolute simplicity could see the product being the only means of assessment. The quest for complete fairness could create an incredible amount of work which could again have a deformative effect on the collaboration in question. The principles have to be tempered. Within collaborative assessment identity this is done by the three guiding orientations, which sit within the frame of those principles. No orientation lies fully within only one principle, and nor does an orientation lie fully outside a principle. The placement of the three orientations in

Figure 4-2 indicates that while all three principles will be in play for each orientation, some orientations are more strongly influenced by certain principles. So too do the orientations influence each other.

In this conception, the principle of fairness is a major driver of the reward orientation, with participants wanting to ensure students get the credit their work deserves, demonstrating sweat assessment matters in the reward orientation. It is not fair for a free-rider to get the same result as

students who fully participated and collaborative assessment requires differentiation so higher-performing students and students who make good effort can be rewarded. Consequently, while the reward orientation is driven by the fairness principle, close adherence to fairness can run counter to collaborative assessment where such differentiation can be difficult, so problematizing such assignments for the reward orientated. The reward orientation is also limited by the practical orientation. To differentiate between students a teacher could invest a lot of time and energy in tracking the collaboration, which would enable fairer rewards, but the practical orientation limits how much resource can be dedicated to assessing collaboration in light of other course demands.

That is not to say fairness has no role in the practical orientation, just that the degree of fairness has to be in line with what is practically possible. For the practical orientation, assessment has to work and be of benefit and, for this to be the case, it needs to be simple and transparent. Fairness is still a concern, but it can be compromised to some extent to keep assessment deliverable and manageable. This does not mean that unfair practices are embraced, but rather there can be acceptance of some weaknesses in an assessment scheme if the scheme overall is felt to be deliverable and generally effective and if the nature of the assessment scheme is transparent to the students. Concerns of overload for teachers and students mean not every assessment lens is brought to bear, adhering to the principle of simplicity. However, the assessment still has to be robust and justified, feeling the pull of the defensive orientation.

The defensive orientation necessarily sits at the confluence of all three principles and is central to EAP practitioner collaborative assessment, sitting between the other two orientations, holding them in check. The defensive orientation requires assessment that adheres to the three key principles and also satisfies the other two guiding orientations, as far as is pragmatically possible. It is where practitioner experience and expectation come into play and where assessment is weighed and measured to ensure it is fit for purpose for the HE environment. So too is it where idealism meets realism and the necessary compromises for collaborative assessment are negotiated. Through defensive avoidance, acceptance and adaptation, collaborative assessment identity and the defensive

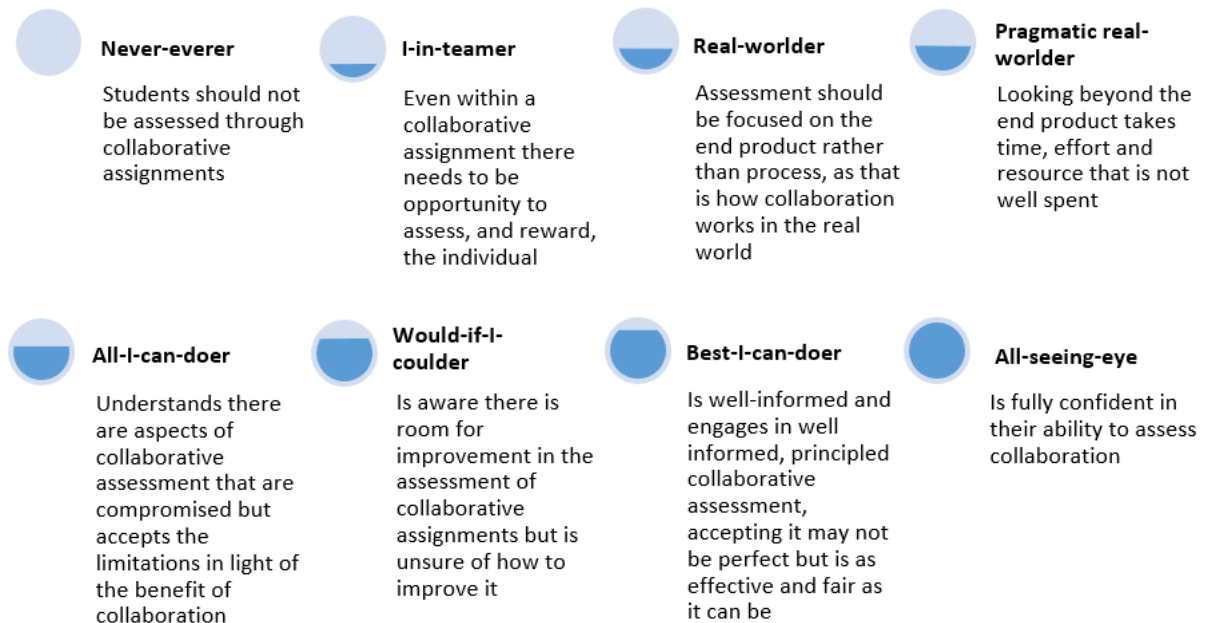
orientation directly impact practitioner approaches and practices in collaborative assessment, as Brown (2011) contended.

4.2.1 Eight collaborative assessment archetypes

During the process of focussed exploration, it seemed possible to identify participants characterised by a driving attitude toward collaborative assignments. The initial nomenclature was ‘collaborative assessment tribes’, which were defined during focused coding as “Categories of teachers who have superficial sameness in approaches for assessing collaborative assignments, albeit their motivations for doing so may differ,” (see Table 2.11). Rather than ‘tribe’, the term ‘archetype’ is now preferred as it brings together both the Platonic concept of ideal forms (indicating these are the purest conception and that practitioners are unlikely to fully inhabit a single archetype), allied with the Jungian conception of archetypes as “a representation of fundamental human motivations” (Domeng, 2021, para. 3). This means if we are able to identify someone within an archetype it gives us an understanding of what they may or may not consider acceptable when assessing collaborative assignments. This is in line with the contention that “who teachers are in the process of assessment is as important as what they know and are able to do” (Looney et al., 2018, p. 456). Through the process of data analysis and theorising, the archetypes in Table 2.11 have been refined, with some names changed and two new archetypes defined.

It was envisaged EAP practitioners could be placed on a spectrum from those who were diametrically opposed to collaborative assessment to those who embrace collaborative assessment and are fully confident in their ability to assess collaborative work (see Figure 4-3). Even in the early conception, the characterisation of the latter end of the spectrum was seen as a theoretical possibility rather reflective of a particular participant in this study, as no-one exhibited full confidence in their collaborative assessment practices. The other archetypes, however, could be identified among the participants. It is important to note the archetypes are not directly linear, with a teacher embodying the first, then the second, then the third. A practitioner may sit anywhere on the spectrum and can shift to any other archetype, as experience and professional development have an impact.

Figure 4-3 *The spectrum of collaborative assessment archetypes*



Wendy is the clearest illustration of a Never-everer, who advocates for collaboration and sees value in it, but rejects the notion that it can be a valid assessment. As she explained in the interview for the companion dissertation, which took place after her interview for this dissertation,

So I would be really positive about collaborative writing in the classroom, which isn't for assessment. And then if it is a form of assessment, quite negative, increasingly negative after the conversation with Peter ... I just thought about it in so much more depth and the more I thought about it, the more I thought, 'oh, just no'. (Secondary Interview Data, p. 284)

For a Never-everer, assessing collaborative assignments will always be fraught with tension, with the driving principle behind this being fairness, with the idea that assessing a collaboration is not seen as a fair means of assessing students. As Wendy explained

I think just the crux of it is seeing how frustrated the students get with with it, and they have no control who they're working with, and I know that that does prepare them for the real world quite well. And that is a helpful experience. I'm just not sure that they should be, that that should impact on their academic grades. (Interview Transcripts, p. 284)

She went on to discuss all the interpersonal challenges that could arise through a collaboration, students with different working styles, students who don't contribute, students who are domineering and felt

all of that's fine, in the kind of personal development sense. But again, should that be something that then impacts on a grade? If, if all of that happens, and then they're graded on a piece of writing? Is that fair? (Interview Transcripts, p. 284)

What is particularly interesting, and illustrative of how people may shift between archetypes, was that during the initial interview for this dissertation Wendy stood out as an I-in-teamer. The I-in-teamers value collaboration but what is paramount for them is that a student can be recognised and rewarded for the work they do, with the reward orientation coming to prominence. This means such practitioners will be looking for opportunities within collaborative assessments to differentiate and award individual grades. Further reflection, however, saw Wendy more fully inhabit the Never-everer archetype. In this move between archetypes we can see two different guiding orientations in play. When discussing the pre-interview task, the practical orientation was dominant, problem solving and trying to determine the best way to use different assessment lenses. With time, however, the fairness orientation came to the fore, leading to the shift in archetype and the rejection of collaborative assessment.

Van, Cristobal and Humberto seem to illustrate the Real-worlder, who are more accepting of collaborative assignments, but would be fully focused on the product to assess such assignments,

rather than considering the collaborative process. This stems from two main arguments. The first is that, in the real world, it does not matter how work is produced, just what the quality of the final work is. This brings together simplicity and practicality. While the collaborative process may be lightly monitored in case of issues, when it comes to assessment, the product rules. The second argument, also driven by simplicity and practicality, is that a successful collaboration will result in a successful product. Consequently, there is no need to try and assess the process of the collaboration as the quality of that is already reflected in the final work.

Pragmatic real-worlders, like Lorenzo, share a superficial sameness with Real-worlders, in that they too focus on the product. However, the rationale differs, with practicality taking on a greater influence. Pragmatic real-worlders are aware that there likely needs to be more to the assessment of collaborative assignments than solely looking at the final product, but are wary of the time and effort it would take to look further into the process, given the limited gains that may lead too. While recognising solely assessing the product may not be ideal, it is all Pragmatic real-worlders can do and they could resist efforts to increase workload by adopting different assessment strategies.

As Pragmatic real-worlders accept logistical limits to collaborative assessment and maintain a product focus, so too do All-I-can-doers accept there may be flaws in their approach to assessing collaborative assignments. This can be seen in Fay knowing there may be unequal contribution in a collaboration but rationalising it through Pareto's Law and in Ava accepting that all she can do is set the standards and expectations of a collaborative assignment and hold to them. Although not perfect, the All-I-can-doer is at relative peace with the compromises necessary in a collaborative assignment. However, Fay's excitement at the idea of individual complementary assignments or assessing awareness of collaboration show how a All-I-can-doer can shift in to the next archetype: the Would-if-I-coulder. As they learn more, All-I-can-doers are open to change and improved practices.

The majority of participants fell under the Would-if-I-coulder archetype. Within this archetype is awareness that there are opportunities to assess collaborative assignments more robustly but appreciation there are barriers to doing so, be that limitations in time, institutional constraints, or lack

of knowledge and ability. In this we can see the defensive orientation at play. This awareness speaks to assessment literacy as it recalls Stiggins' concept of knowing when there is something wrong with an assessment (Looney, et al., 2018), although it does not full meet his expectation that the assessment literate additionally know what to change to improve the assessment. While the Would-if-I-coulder know that they could develop and perhaps should improve their collaborative assessment practices, they lack the knowledge to do so, recalling Humberto's recognition that "if you can do it, then it's good. But I'm just not sure how you would do it" (Interview Transcripts p. 167) and Gonzalo's contention that "I think if you come up with a way to assess collaboration effectively, you could make millions" (Interview Transcripts, p. 155). This is not a critique of the participants. The literature also currently offers little in the way of solutions.

The All-seeing-eye is a true archetype and was not encountered in this research, but it is the logical end point of the continuum, named after a quote from Bertha: "I don't like the idea of the teacher being this like, all seeing eye who can just swoop in at the end and then award some arbitrary mark when they weren't involved" (Interview Transcripts, p. 68). The All-seeing-eye has full confidence in their ability to accurately and fairly assess a collaboration. Rather than that ideal, perhaps the best that can be hoped for, after practitioner collaboration and training, is the rise of the Best-I-can-doer, a well-informed practitioner actioning collaborative assignments in a principled, fully rationalised way. While this archetype was not directly encountered, with all participants needing fuller knowledge of collaborative assessment, it does not seem unreasonably idealistic or ambitious. The practical output of this dissertation, the collaborative assessment algorithm, would be a useful tool in terms of helping practitioners transition into being Best-I-can-doers.

The value in identifying such collaborative assessment archetypes is considering the introduction of collaborative assessment into a teaching centre and recognising not every teacher is going to have the same initial response or be equally invested in making a success of such an assignment. EAP practitioners may feel quite strongly about how such an assignment should be assessed, if it should be assessed at all. Wendy may not be alone in thinking "I see the value of

[collaborative assignments] but as a form of assessment, I think they're a bit of a cop out" (Interview Transcripts, p. 288). If this is not recognised, and the proposed assessment scheme not rationalised, there could be considerable resistance, as was the case in Fernand's teaching centre, and that resistance could have negative washback on the teacher and student experience, triggering a vicious cycle of negative experience hardening negative attitudes toward collaborative assignments.

Having gained greater understanding of the attitudes of EAP practitioners to collaborative assignments and developed the theoretical construct of collaborative assessment identity, embodied by collaborative assessment archetypes, the question of this dissertation then remains. Addressing this takes us to the practical outcome of this dissertation: the collaborative assessment algorithm.

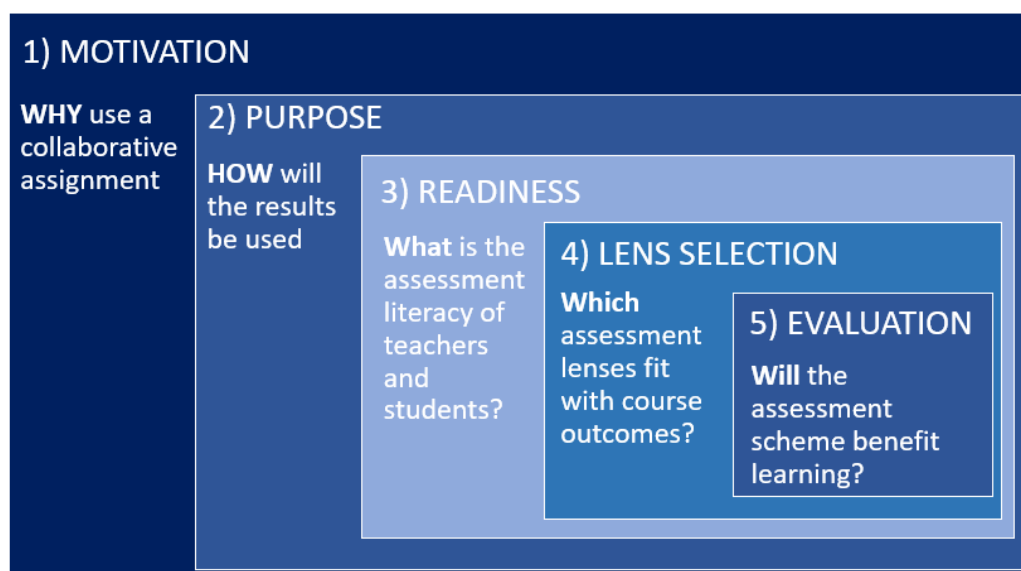
4.3 A collaborative assessment algorithm: Designing a principled assessment scheme for a collaborative assignment

So, how can we most effectively and fairly assess collaborative assignments in EAP contexts? There is no simple answer, no one-size-fits-all answer, no answer which will satisfy everyone. Collaborative assessment, in its nature, is messy, and there is no fully fair solution available. That is not mean, however, that collaborative assignments should be abandoned. They have utility and value and, given EAP's preparatory role, they are also necessary. The design of an assessment scheme for collaborative assignments matters, with Barry recognising collaborative assessment "suffers from the way we set up those assessments, I think it suffers from the fact that they are not truly worked on as a process" (Interview Transcripts, p. 18). As Mueller (2005) stated, when discussing AA, "the development of good assessments of any type begins with the development of meaningful goals and standards" (p.2). It is also important to remember Norton's contention that "assessment needs to be used to drive the learning process in a way that will encourage students to actually engage in meaningful learning rather than just perform the necessary assignments" (Norton, 2004, p. 693). The design of assessment for a collaborative assignment has to be geared toward maximising the

affordances of such tasks so that they can be a “vehicle for educational improvement” (Shapiro, 2003, p. 425).

This dissertation will not offer a ‘do this’ answer, but rather propose a collaborative assessment algorithm, meaning a series of considerations which should help each teacher/course leader/teaching centre/institution through the thought processes and decision trees of developing a suitable assessment scheme for collaborative assignments. A working title for an early conception of this dissertation was ‘Towards a framework for assessing collaborative assignments in EMI EAP Programmes’. The word ‘framework’ speaks to structure but, as the findings of this dissertation developed, ‘algorithm’ became the preferred term as it speaks more to procedure and problem solving. The dimensions of collaborative assessment identity could be said to provide a framework for designing a collaborative assessment, and, as will be discussed in 4.3.5, do provide a useful evaluation framework for a collaborative assignment assessment scheme. However, the development of that assessment scheme is more algorithmic and the collaborative assessment algorithm is designed to help EAP practitioners navigate the procedure of developing an assessment scheme for collaborative assignments. That procedure is divided into five layers of thought processes, as illustrated in Figure 4-4.

Figure 4-4 A collaborative assessment algorithm: 5 layers in designing an assessment scheme for a collaborative assignment



Note that rather than being discrete the layers are sequential, with each layer building on the preceding ones and the considerations of each layer resonating in later layers. For instance, lens selection will be informed by how the results of the assignment will be used and the assessment literacies of the teachers and students.

The collaborative assessment algorithm will go further than de Hei, Strijbos et al. (2016) suggested in the GLAID framework, guiding the practitioner further through the process of developing a collaborative assignment assessment scheme. It will also hopefully go some way toward helping EAP practitioners develop their assessment practices for collaborative assignments, meeting one of the CFTEAP competencies (see Table 1.3). There is a recognised need for guidance for teachers concerning collaborative assignments and Bertha is likely not alone in thinking

I would like to do, to work more on that, these kind of collaborative tasks, but you kind of think sometimes, "where do you start with this," like, and it's probably just easier to not go there, but it will be really valuable. I think it would be really valuable thing. (Interview Transcripts, p. 70-71)

While absolute fairness in collaborative assessment may be unobtainable, this dissertation stands with Gipps and Stobart (2009) and Fay and Gonzalo in hoping for an assessment scheme that can make collaborative assignments fairer. It is also important to recall Beard et al.'s (1989) assertion that how a collaborative assignment is assessed can impact student engagement, which means an ongoing concern throughout the process of designing a collaborative assignment assessment scheme should be looking for opportunities for positive washback on student behaviour and engagement since assessment can encourage collaborative effort (McDonald, 2003). Positioned as AaL, collaborative assessment should be designed to be a positive learning experience through the assessment process.

4.3.1 Layer 1 – Motivation: Rationalising collaborative assignments in EAP

EAP has a responsibility to prepare students for their academic studies and, since students will encounter collaborative assignments during those studies, EAP has a role in preparing them to collaborate effectively. Collaboration also falls in the remit of EAP since it necessitates communication, providing more opportunity for authentic interactions between students. There are sound reasons to do collaborative assignments, be that meeting learning outcomes or enriching the student experience. However, if the decision is purely logistical and driven by too many students and too little time, collaborative assignments may well not be the panacea hoped for.

It is essential to remember that in order for students to become better collaborators they need support and guidance in collaboration. As Cristobal contended, “if you're going to start assessing students for their collaboration, you *have* to have taught skills and what you expect from them in the first place” (Interview Transcripts, p. 82, emphasis in original). However it may be assessed, collaboration cannot be unsupported. As Wendy said,

there's almost an assumption that students will be, they'll be able to do it, you know, people can work things out, they know how to work together, or they can, they know how to, you know, share your WhatsApp number and make a group and like, there's an assumption that they'll have the social skills necessary to make this thing happen. And I just don't think they always do. I really don't think they do. (Interview Transcripts, p. 285)

Students will need support and the exhortation of Poehner and Lantolf (2005) is highly pertinent in collaborative assignments: “the teacher needs to be prepared to provide appropriate types of mediation and to know when to withdraw this mediation as the learners begin to mediate themselves” (p.242).

The need to support collaboration is especially important since negative collaborative experiences can affect student attitude to future collaborations. This was something highlighted by Karen in her interview when she spoke of the importance of preparing students for collaboration since

I think for the most part, they are very ... they're not ambivalent towards it, they do they either hold really, really positive feelings or quite sort of negative feelings. But again, that just goes, I think, to past experience. (Interview Transcripts, p. 187)

A bad experience can colour a student's attitude to collaboration and allowing students to work out how to collaborate by themselves as they go along perpetuates the lack of knowledge and discourse around collaboration. This does not serve a preparatory or positively developmental function, potentially leading to a collaborative experience that does more long-term harm than good.

Consequently, a collaborative assignment does not lower workload, for either students or teachers. Students will be engaged in more discussion, negotiation and debate, along with more logistical challenges of coordinating and working with other people. While the final marking load for teachers may be lower, more support to students will be required setting up the assignment and monitoring the collaboration throughout the process to mitigate against dysfunction. This would resonate with Olga who felt "I think for the teachers the stressful part is the time you have to put into it. And I think I do put into more time in this whole planning thing" and she went on to recognise "I do have to put more into that in terms of planning, organising, coordinating" (Interview Transcripts, p. 255-256). As Laura explained about facilitating a collaborative assignment

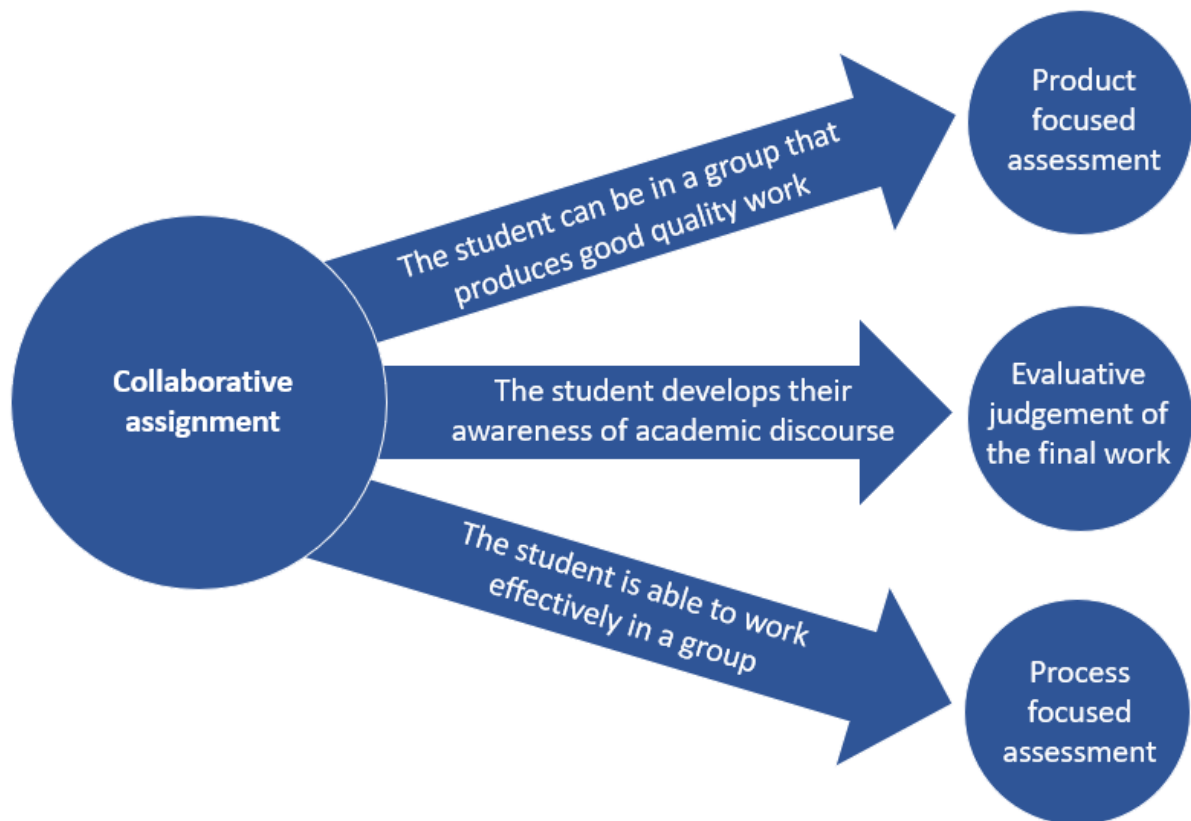
But you need to be proactive, and you need to follow them but not in a very, you know, dictating manner. But you need to give, give them the message that you're there to help, to guide. In case of emergency, you're always there. (Interview Transcripts, p. 223)

While they have benefits, collaborative assignments are going to be hard work which is why understanding why you want to commit to that work is so important.

4.3.2 Layer 2 – Purpose: Conceptualising a valid collaborative assessment

Once it has been determined why a collaborative assignment might be desirable, it is necessary to clarify the purpose of the assessment. Although traditional conceptions of validity may not be fully fit for collaborative assessment since it sits within AaL at the nexus of DA, CBA and AA (see 1.4.5), the validity of a collaborative assignment is still a consideration. Using Messick's (1989) conception of validity, relating to how results are interpreted, it is essential to understand the purpose of the collaborative assignment and how the results will be used. As Wendy questioned, "what are we actually marking them on? Is it the piece of writing Is it is it the project as a whole? And do students fully understand how all of that plays out?" (Interview Transcripts, p. 285). If students engage in writing a group essay, what information will the assessment of that assignment provide and how can that information be understood? The means of assessing a collaborative assignment will change, depending on what the results will be used to evidence. Figure 4-5 (below) illustrates how assessment purpose impacts assessment focus. It is not a definitive illustration, as a collaborative assignment could be used for different evaluation purposes, but is meant to highlight the importance of clarity of purpose from the outset.

Figure 4-5 *Impact of assessment purpose on assessment focus*



The first purpose, demonstrating students can be in a group that produces good quality, work was determined by considering what purpose wholly product-based assessment could serve. If there is no consideration of the process and the product is the only assessment lens, the purpose the assessment can serve is necessarily limited. If the assignment is intended to demonstrate the student has developed their awareness of academic discourse, that cannot be measured through the product as there is no clarity as to what the individual student contributed and how that impacted their academic discourse knowledge. However, if the student has to evaluate their work, illustrating ways in which it fits the target academic genre, that evaluation does speak towards their awareness of academic discourse. If they are able to identify strengths and weaknesses in the work they understand what that work should be. The final purpose, if the student is able to work effectively in the group, requires assessment of the collaborative process.

An ill-defined purpose for assessment “muddies the waters” (Interview Transcripts, p. 168) and challenges the validity of the assessment. It is essential to understand how the results of the assessment are going to be used to ensure the assessment scheme is designed in a way to get information relevant to that purpose. If there needs to be differentiation between students and clear, evidenced reporting of student’s individual competencies and capabilities, collaborative assignments may not be the most appropriate assessment tool, or will need assessment lenses that do facilitate evaluating individual competence.

4.3.3 Layer 3 – Readiness: Considerations before implementing a collaborative assignment

Figure 4-6 (below) asks three key questions to guide considerations when implementing collaborative assignments, or reviewing existing one.

Figure 4-6 *Readiness considerations before implementing a collaborative assignment*

1) What does good collaboration look like?

How is it hoped or envisaged students to work together? What would indicate a successful collaboration?

2) How does a collaborative assignment fit with the other course assessments?

What other assignments are there in the course? What time is available for a collaborative assignment? How would a collaborative assignment grade impact a student’s overall course grade?

3) What time and resources are there for teacher and learner training?

Are teachers well prepared to facilitate collaborative assignments? Are students equipped to reflect or self- or peer-assess? Do students and teachers have competence in the desired collaborative platforms and workspaces?

The first question addresses the nature of collaboration itself. It is essential to develop a clear, shared conception of collaboration so various stakeholders are on the same page. The divergence of

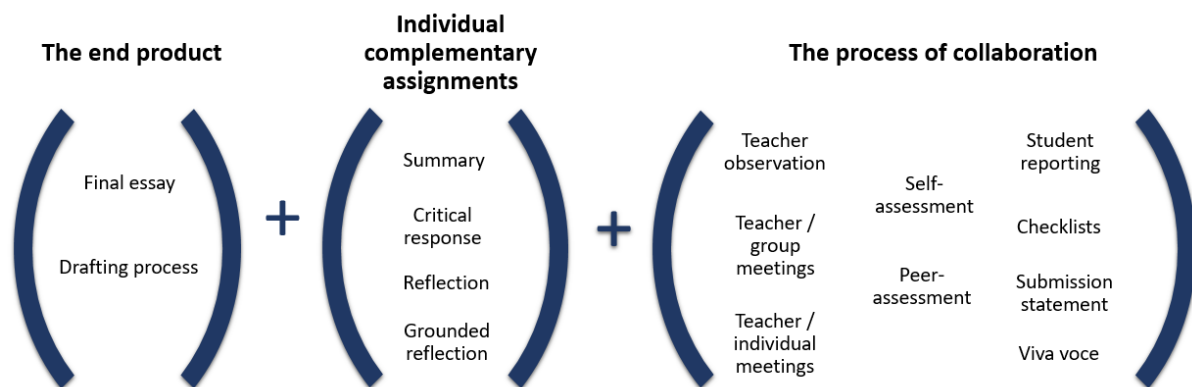
understanding of collaboration is an issue, as Kyle recognised when discussing assessing collaboration itself, asking “what do we mean by collaboration? What does that mean to, well, the teachers, the markers, first of all, who might have different ideas about it, the students who might have different levels of experience?” (Interview Transcripts, p. 209). Having an idea of what successful collaboration looks like is the first step in identifying indicators for assessment.

Following this, it is also necessary to locate the collaborative assignment within the other assignments in the course to ensure all course goals are being met and that there is not over-assessment in one area or under-assessment in another. The third question leads to practical consideration of workload for teachers and students, both in the preparation and undertaking of collaborative assignments, and the potential learner training that might need to be factored in. Once it has been determined that a collaborative assignment is a potentially suitable and beneficial assessment tool, consideration can turn to how that assignment could be assessed.

4.3.4 Layer 4 – Lens selection: Selecting assessment lenses

There is a constellation of assessment opportunities around collaboration, with a variety of assessment lenses which could be utilised, each of which have the opportunity for learning affordances. As discussed earlier, no single lens will bring a collaboration fully into focus and each lens in itself is an imperfect measure. Cumulatively, however, they do provide more insight and more potential for fair assessment. The lenses are broadly categorised into three main areas: the end product, complementary individual assignments and assessing the collaboration itself. The type of assessment lenses which could be utilised in each category are illustrated in Figure 4-7 below with Figure 2-14 in the Methodology section illustrating the potential relationship between these lenses. The illustration should not be considered exhaustive, as there are various potential assessment lenses of collaborative assignments available that would provide opportunity to engage students in productive, developmental work, addressing a range of learning gains and outcomes.

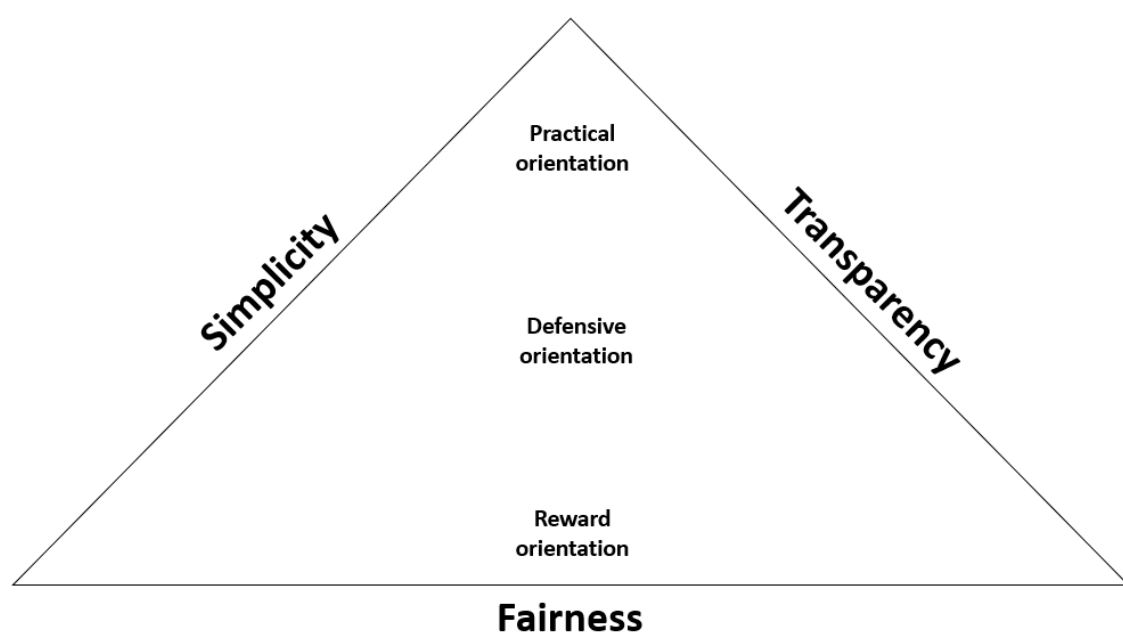
Figure 4-7 Potential assessment lenses for collaborative assignments



These assessment lenses are discussed in turn below, through the prism of the collaborative assessment identity. Where

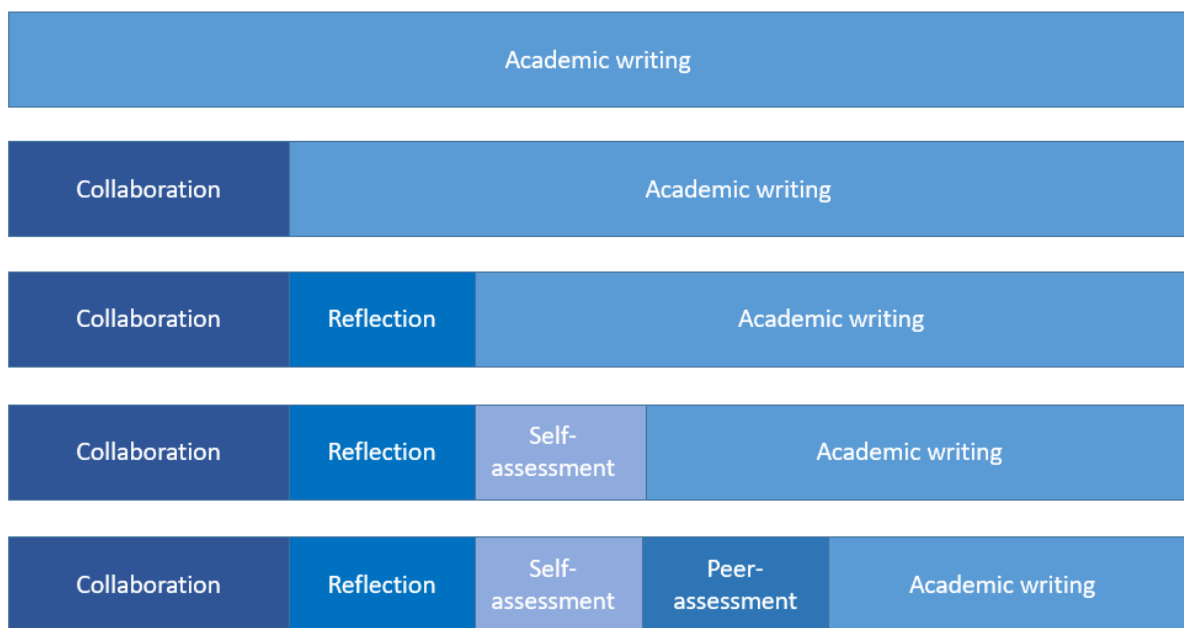
Figure 4-2 emphasises the overlapping nature of the three key principles, Figure 4-8 (below) represents a prism through which we can look at assessment lenses, where the three connected principles encompass the three guiding orientations, with fairness serving as the principle that grounds everything else.

Figure 4-8 The collaborative assessment identity prism: key principles encompassing guiding orientations



Before that discussion, however, it is important to note that while there are an array of assessment lenses available to evaluate a collaborative assignment, which lenses to employ will be a case of considering wider course outcomes, the profile and assessment literacy of the student cohort, and what is practically possible for students and teachers for that particular assignment as part of the fuller workload of the whole course and the full semester course load. A prime consideration is to try to avoid the $1+1+1+1+1=1$ problem identified in Table 2.11. Multiple lenses *could* be employed in service of assessing the same collaborative assignment, all of which contribute to a single overall grade for the assignment. To assuage Van's fears of duplication and unnecessary assessment and falling into Torrance's conception of AaL wherein assessment dominates learning (see 1.4.2), there must be a value to an assessment lens in and of itself, with each lens aligned to the learning outcomes of the course and leveraged for positive washback and increased learning gains. Each assessment lens will have a learner training opportunity cost, as illustrated below (Figure 4-9), which will take time away from the learning training for the assignment itself. Figure 4-9 is not meant to be taken literally in terms of expected percentage of time spent preparing students for each assessment lens, but is illustrative of task dilution and the time demand learner training for various assessment lenses may have.

Figure 4-9 Learner training opportunity cost of assessment lenses



For an example of maximizing an assessment lens, consider groupmates being required to give feedback to each other on the work they produce. There is potential for this to descend into busy work if the task is phrased as “Provide feedback on the contributions of the other members of the group”. However, if it is phrased as “Provide feedback which should have a positive impact on the quality of the final product” then students will necessarily be thinking about the consequence of their feedback and be orientated toward providing feedback to improve the essay. Further consideration of the different potential assessment lenses follows, drawing on the recommendations for implementing different models of assessment from the Findings and the literature.

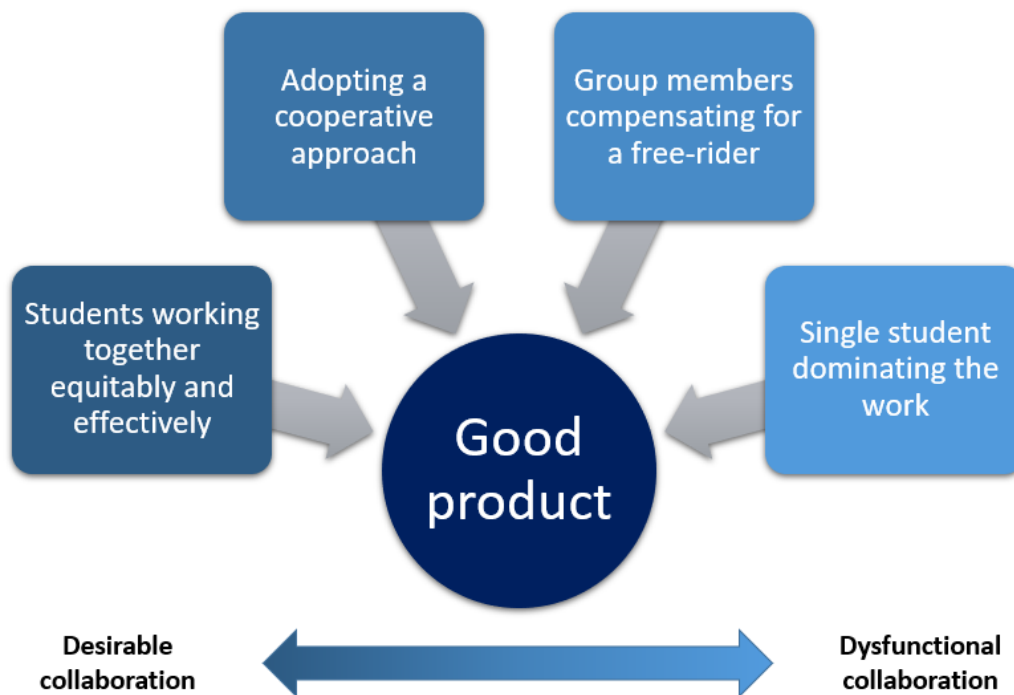
4.3.4.1 The final product as assessment lens

The final output of a collaboration has to be assessed, as the first assessment lens. If students are working together to produce a piece of work, the quality of that final work matters and if the work is not graded, students could have limited motivation to engage effectively with it. Even if the assignment’s prime concern is the collaboration itself, the product of that collaboration is an indicator of the work that was done. As Van argued, the coherence of a text is “a kind of inferred feature of the,

how good or bad the collaboration was” (Interview Transcripts, p. 279). The final product is not, however, a perfect indicator.

Assessing the final product meets the principles of simplicity (one assessment measure for one assignment) and transparency (what matters is the quality of the final product) but falls short in terms of the leading principle of fairness. As Storch and Knoch (2023) and Van Aalst (2013) hold, an individual’s abilities are not easily assessed through a piece of collaborative writing. Nor is the final work a clear measure of collaboration, especially given the issue of product glare (see Figure 4-10).

Figure 4-10 Product glare masking quality of collaboration



A good quality final product could mask problematic collaboration e.g. the final work being the product of an asymmetric collaboration dominated by the most competent student, supporting Williams’ (2017) contention that the process of collaboration also has to be considered as the product alone is not enough. The quality of the final product only evidences the quality of the final product, not how it was produced. As Barry explained, assessing the product is “limited in terms of actually

producing any, if you like, outcome for recognising the value of collaboration” (Interview Transcripts, p. 15). This is true of both written and oral assignments.

In the case of an oral collaboration, like a group presentation, a quality final product could still be the result of a dysfunctional or asymmetric collaboration. While there is scope for differentiation in terms of students’ speaking or presentation skills, the final presentation does not show how the final work was produced. This could be particularly unfair if the majority of the work in the development of the presentation (research, thesis generation, organisation, generation of visual aids), was done by a student with lower presentation delivery skills. In a marking system that rewarded surface level differentiation, focusing on individual speaking and presentations skills, the more able presenter, who only delivered what they were told to deliver, would gain a higher grade than the less able presenter, who did the majority of the academically sound work developing the presentation. A potential mitigation of this is the inclusion of a submission statement, as used by Tanya (Interview Transcripts, p. 33), where students attest to the work done and contributions made, although the veracity of that statement could be open to question.

Neither does the final product capture all the learning that could take place during a collaboration. Assessing the end product provides a summative measure, but neglects any formative learning which took place, an issue further complicated given Strijbos’s (2016) recognition that different students could learn different things from the same collaboration and Betha’s contention that the final product is “likely not where a lot of learning has taken place” (Interview Transcripts, p. 70). A problematic collaboration, resulting in a flawed final piece of work, could actually serve as a better learning experience for the students than an unchallenging collaboration. As Kyle put it, “If you’re genuinely sort of reflective about it you probably, the old thing, you learn more about failing than you do from succeeding” (Interview Transcripts, p. 216). From a bad experience, students could better understand how to collaborate more successfully in future, or what kind of behaviours can lead to dysfunction in a collaboration. Gonzalo reflected that he has probably learned more from his failures than successes and recognised

a potentially negative collaborative experience could really help students understand, "Okay, this is what, what not to do." And then it may not show them what the right thing to do is, but it'll show them what not to do. And then they can try, you know, whatever other approaches that they that they think might work in the future. (Interview Transcripts, p. 154)

If the product is the only lens of assessment, such learning gains, and such gains that would evidence development of collaborative skills as a learning outcome, would be invisible.

With the process writing approach being common in EAP, assessment of the end product could also encompass the drafting process since, as Kyle argued, "from an EAP perspective, I think the real value is in the process," (Interview Transcripts, p. 205). This could involve looking at how students incorporated teacher feedback on drafts, or how effectively they engaged in the peer feedback process. There is potentially more scope here for getting some indication of the quality of collaboration, particularly if early drafts evidenced a lack of cohesion which subsequently gets edited out during the drafting process. As clear in the definition of collaboration for EAP provided in the background study to this dissertation (see 1.3.1), there is potential for students working on an extended collaborative assignment to adopt a cooperative approach at times during the task. This could be more likely early in the process, where each individual is building their understanding of the topic and task before generating a shared understanding within the group. As such, incoherence in a first draft could be expected but, as the collaborative editing process progresses, the work should take on a more coherent shape and singular voice, indicating positive collaboration.

However, such editing and development of a more uniform voice could be the result of a single student taking responsibility for redrafting. Through looking at the drafting process a teacher can see what kinds of changes were made, but not necessarily how those changes came to be. In this sense, while there might potentially be fairness gains in assessing the drafting process, it adds complexity as there are different ways to edit based on feedback. A feedback suggestion could be ignored or it could be carefully considered and discussed by the group, but rejected. The end result will be no change in

the text but in the first case this would demonstrate a failure to engage with feedback and, in the latter, full engagement with feedback. So too could a change be made in response to feedback that actually lowers the quality of the final work. The question then becomes if the attempt to respond to feedback should be rewarded, despite being ineffective.

A product-only approach would suit a practical orientation but perhaps be difficult to accept for a reward orientation as there is no scope to genuinely reward individual students for the work they have done. It would also have to be considered from the defensive orientation in terms of how such an assessment approach would be rationalised and explained to students. Consideration would have to be given as to how it could be enacted in a way that fosters collaboration rather than opening opportunities for students to circumvent collaboration to get the best possible end product. In any case, it is worth recalling Kyle's comment on the quality of the work reflecting the quality of the collaboration quoted earlier. "We would assume a better end product means a more successful collaborative process *if we set it up right* [emphasis added]" (Interview Transcripts, p. 209).

If the product is going to be the only assessment lens used, diligent work need be done at the outset so students understand the task, the benefits of collaborating and the expectations. As recognised in Table 3.3, a product only assessment approach still requires focus on the process and ongoing teacher involvement and feedback. Fay would concur, considering the development of a collaborative assignment being "a reverse engineering type of thing. Okay, where do I want them to end up? And like what skills? What knowledge? What can-dos do they have at the end? And how do they, how do I get them there?" (Interview Transcripts, p. 118). Focusing on the product for assessment purposes does not negate the need to facilitate and support the collaboration and, indeed, perhaps makes doing so even more necessary. As Ava asked when talking about needing to talk about collaboration with students from the outset, "If we as instructors didn't teach them anything, and didn't uphold the standard, then what exactly are we doing?" (Interview Transcripts, p. 47). This allies with Karen's desire to provide "transformational feedback" (Interview Transcripts, p. 185), locating

collaborative assignments very much as DA (see 1.4.4). If one of the purposes of assessment is to help students improve, the process of collaboration needs to be supported, whether it is assessed or not.

4.3.4.2 *Individual complementary assignments as assessment lenses*

The second set of assessment lenses encompass individual complementary assignments, which were generally favoured by the research participants. From the outset, it is important to note that the limitations inherent in solely addressing the final product are still apparent in an assessment scheme incorporating individual complementary assignments. The assessment of the main product itself will carry the concerns outlined in the previous section. Although the individual components *may* be an indicator of engagement with the main collaborative task, they add little additional insight into the collaborative process unless specifically designed to do so. As Wendy put it, through complementary individual assignments “you’re not assessing their collaboration at all actually” (Interview Transcripts, p. 303). The discussion that follows is also referring to mandatory individual complementary assignments that are fully part of the assessment scheme. This is in contrast with Karen’s use of voluntary individual components, which were available to students who wanted to increase their grade from the collaborative assignment through taking a quiz or doing some extra work (Interview Transcripts, p. 186), which leads to a scalability and workload concern if more students than anticipated take up the option.

However, mandatory individual complementary assignments provided scope for differentiation between students and an opportunity to broaden the scope of the collaborative assignment, meaning further learning aims could be met. There is an increase in fairness, as a student’s individual work will be assessed and that will be a truer reflection of *their* work than the group product. At the same time, simplicity suffers, since each component added to an assessment adds more complexity. The impact on transparency is dependent on if the individual assignment is assessed on its own merits, or if it is used as an indicator of the quality of the collaboration, which could be more problematic. Considering the kind of individual complementary assignments which could be incorporated, there are those that

are directly related to the main task and those that are more tangentially related to main collaborative task, but which have no direct impact on it, with the recommendations of Table 3.5 holding true in terms of each task needing to build toward a learning outcome and the individual tasks being supportive of the main collaborative task.

A direct complementary task could be students individually finding a relevant source, summarising it or writing a critical response to it, which engages them in the research process and knowledge-building needed in order to develop the group essay. Such assignments provide examples of students' academic writing and can demonstrate students can write to the expected level, following the norms of an academic genre. It is important to try and mitigate the potential of these complementary assignments to be counterproductive to the main collaboration. For example, assignments that serve a background knowledge building purpose should be completed before the main collaborative task begins and could even, potentially, take place before group formation has occurred. So too will measures need to be taken in learner training to ensure when writing the group essay students are synthesising sources rather than being wedded solely to *their* source. Nonetheless, complementary assignments can meet specific learning goals the collaborative assignment may not and, additionally, provide opportunity for individual assessment within the context of a wider group task.

The other individual complementary assignment type would be one that has a connection to the collaborative task, but does not directly feed into it. Reflection, would be a prime example of this. An individual reflection after the collaboration is complete was generally valued by the participants, although it must be clear what such an assignment is being assessed for. A reflection could be assessed as a measure of a students' ability to reflect, or as an indicator of the quality of the collaboration. The latter purpose is problematic due to the concerns the participants had of the student as unreliable narrator, although a way to mitigate this could be to conduct a grounded reflection. In this type of reflection task students are given some learning theories, in this case theories related to collaboration. After they complete the collaboration the reflection is geared toward how theory explains their

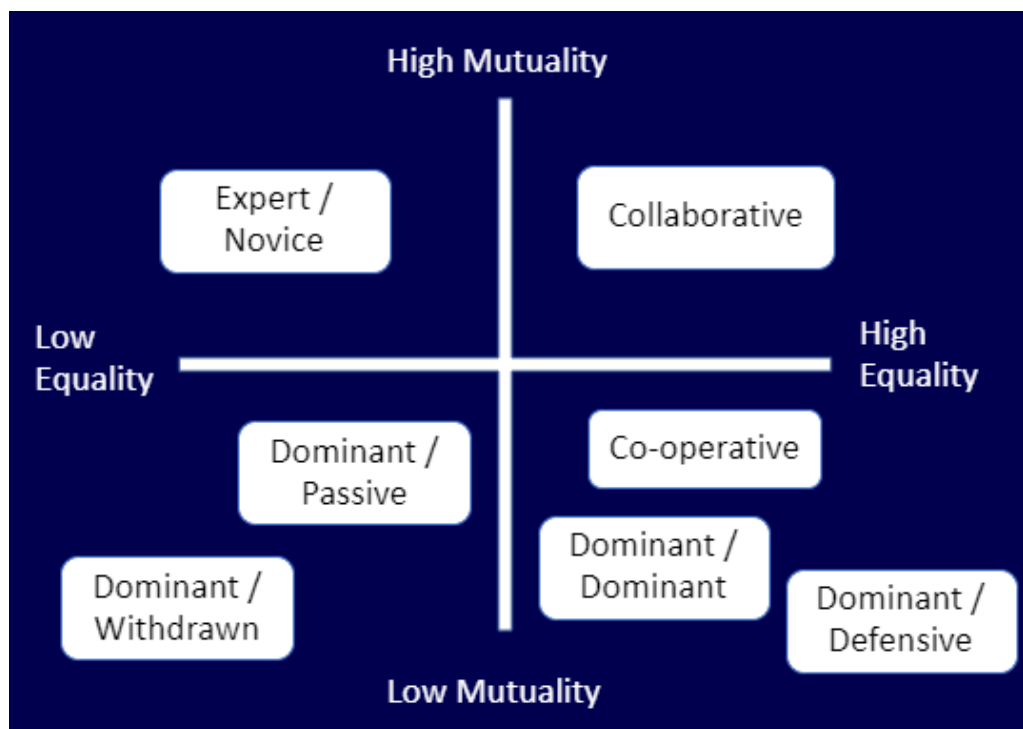
experiences, or where their experiences differed from theory, and why. In this kind of grounded reflection, the concerns of unreliable narrator may be mitigated because even if there is a performative nature to a students' reflection, they will still be incorporating theory into their writing and juxtaposing their experience and theory, which is a useful academic skill in itself. This kind of task could also speak more to student's broader awareness of collaboration, rather than the quality of their collaboration during a specific assignment, which could have particular utility if collaboration is not a stable construct.

Complementary individual assignments would satisfy the reward orientation but would need to be considered from the practical orientation and defensive orientations. Each task would have to have a purpose that either a) helps the main collaborative task, b) meets a distinct learning aim or c) ideally, both of the former. The workload, for teachers and students, would need to be considered to ensure no overload. This would need to be considered for both the complementary assignment itself and for any teacher or learner training that would need to precede it. A final consideration would be Bertha's concern about "mixed messages" (Interview Transcripts, p. 59), when having individual components in a collaborative task. The relationship between assignments and the positive washback from one to the other would need to be clear to students, and teachers, from the outset.

4.3.4.3 *Collaboration itself as an assessment lens*

The third set of assessment lenses are those focused on the collaboration itself. The assessment of collaboration is particularly challenging, although, as Humberto recognised "If you *can* do it without screwing up the entire operation and making it too surgical, then yes, I think it could and should [be assessed]" (Interview Transcripts, p. 170. emphasis in original). The multitude of dimensions to collaboration were outlined earlier and before any attempt to assess collaboration there needs to be clarity on exactly what is meant by collaboration within in the context of a particular assignment. One potential measure of collaboration suggested in the literature comes from research into interaction patterns, evaluated on the axes of mutuality and equality (see Figure 4-11).

Figure 4-11 Interaction patterns based on Storch, 2002; 2013, Li & Zhu, 2013, 2017; Edstrom, 2015



In this analysis, more positive collaborations occur when there is high mutuality between participants. Mutuality “refers to the level of engagement with each other’s contribution. High mutuality describes interactions that are rich in reciprocal feedback and a sharing of ideas,” (Storch, 2002, p. 127).

When looking to utilise mutuality as an assessment measure, concerns abound. The first is the visibility of such interactions to the teacher in cases where students will not be conducting all of the assignment within the context of the classroom, bringing to mind Dolly’s assertion she cannot assess what she cannot see. Students documenting such communications is an option, but one which recalls Van’s distain for busy work and box ticking. If students have to document reciprocal feedback and sharing of ideas to gain credit, reciprocal feedback and sharing ideas will certainly be documented, although the value of such behaviours and reporting may be questionable. Would students be giving feedback because they genuinely believe there are ways to improve the work, or would each feedback comment be a step towards meeting the required number of feedback interactions? As Tanya noted

when discussing using technology to look into a collaboration, “students are smart enough that they could kind of game that” (Interview Transcripts, p. 39).

The other axis of the analysis is equality which “refers to the degree of control or authority over the task. Equality describes more than merely an equal distribution of turns or equal contributions but an equal degree of control over the direction of a task” (Storch, 2002, p. 127). Contribution will be discussed in the next section but the latter aspect, students having equal control over the direction of the task, speaks to the shared ownership and responsibility for the whole work. Teacher/group meetings could be a way of looking into this aspect of collaboration, ensuring everyone is involved and aware of what their work is saying. This is one of the features that gave participants more confidence in assessing group presentations since the Q&A provided an opportunity to really see what the students knew. As Fernand explained, during the Q&A of a presentation you could see “well, actually, maybe they did do, they did know what they were talking about or, more often the case, no, they hadn't a clue what they were talking about, you know, they were trying to wing it” (Interview Transcripts, p. 132). Particularly in an age of LLMs bringing the authenticity of written submissions into question, there could well be a rise in oral defence of written work at all degree levels. Discussion of the development of a collaborative piece of work, or the defence of a group assignment, could shed light on the relative knowledge and investment of each individual.

4.3.4.3.1 What is “contribution”?

The question of “contribution” is a vital one when considering the assessment of collaboration. Participants wanted to see equal or equitable contribution between students in a good collaboration, echoing the sentiment of the above Storch quote. It is important to note from the outset, however, that ‘contribution’ and ‘collaboration’ are not synonymous. A student who finds three random sources for an essay has contributed, a student who negotiates with their group mates about how to use three potential sources they have found in the essay has collaborated. A student who pastes three hundred words from ChatGPT into the essay has contributed, a student who has written three hundred words

in service of furthering the group's thesis has collaborated. Not every contribution is part of a collaborative effort. For a contribution to have value, it should be meaningful and further the work of the group.

Where things are further problematized is in identifying the various ways in which students could make meaningful contributions to a group task. Kyle provided an example of how two students could work together, making different contributions according to their strengths, recalling Boud and Bearman (2022) in linking collaboration and collusion.

Student A perhaps has, has, is not as proficient in English as Student B. But Student A has done a lot of the reading and got the you know, the notes and examples. So then you ask Student B, so well, why don't you proofread it? And then actually, well, that's like ooh, is that collusion or plagiarism? Or is that effective collaboration? (Interview Transcripts, p. 214)

The issue of the range of ways students can contribute rose as an increasing concern for the researcher, prompting the interview question

How do you weigh the relative contribution of the kid who actually understands the topic and comes, comes up with the thesis and says, "That's the argument that we're making," the kid that can go off and do the research and find the support for that argument. And then the student who doesn't contribute that much to the essay, but if they weren't there, the other two would kill each other? (Interview Transcripts, p. 71)

This recalls Belbins' (1981) breakdown of team roles into cerebral roles (e.g. generating the thesis), action-orientated roles (e.g. doing research or drafting) and person-oriented roles (e.g. maintaining group dynamics). All of these contributions are necessary to the effective working of the group and the quality of the work of that group. As such, contribution cannot be captured in raw numbers, where

each student makes so many contributions. Within an academic writing assignment, which encapsulates research, critical thinking, organisation, planning, thesis development, drafting, dealing with feedback, redrafting and proof-reading, there are a multitude of ways a student could contribute. As Olga put it, when pondering assessing collaboration, “how do you, how do you define, you know, in terms of how much percentage of the writing they do, would that count as a collaboration or the background research, should that count as collaboration?” (Interview Transcripts, p. 257). Some types of contribution, like words on the page, are more visible and measurable. Others, such as making the logistical arrangements to get the group together, or motivating the group and keeping it energised and focused, are harder to observe and quantify. Laura was able to well illustrate this, when investigating a student who did not seem to be contribution to their group’s collaborative workspace. When she investigated, the other students explained

she may not be here, she may not be on Google Docs, but she attends our regular Zoom meetings. And she's providing us a significant amount of food of thought. She supports us so much. You cannot see her on document, but she's here, she's supporting and she's acting like a mentor. And she's also providing us with some research findings. (Interview Transcripts, p. 231)

It could be the case that students might not have the opportunity to contribute if a single student is dominating the collaboration. Gonzalo pondered this lack of opportunity to collaborate, explaining

you might have some bossy people or difficult people in the group and then the other ones are suffering, who might want to collaborate or might be good at collaborating, but they can't because they're getting bullied or shouted down or ignored. (Interview Transcripts, p. 92)

This could go so far as a student unilaterally deleting the contributions of others, as was the case with one of Laura's collaborations where a student decided the work of the other students was not good enough (Interview Transcripts, p. 228).

Bertha questioned the quality of contribution when discussing self-assessment and thinking "maybe one student could say, "well, I spent the most time on this. Therefore, I should get the highest grade." Is that, you know, is spending the most time on it the most important factor?" (Interview Transcripts, p. 65). This demonstrates awareness that the measurable (in this case time spent) may not be the best measure of collaboration. However, participants also engaged in sweat assessment when it came to collaborative assignments, considering what students put into the work and wanting to see students make an effort to work together to produce the required output.

An interesting thought experiment is considering a free-rider, who contributes nothing but has a more positive impact on the quality of the product than a demotivated student who produces poor quality work. This is illustrated by Laura, describing students submitting late work to the group, saying

we got these students, extremely laid-back students who are doing things the last minute, you know, the other students, they tried everything in the best way and the others, they just include a couple of things the last minute, and they ruin the whole coherence of the paper. (Interview Transcripts, p. 224)

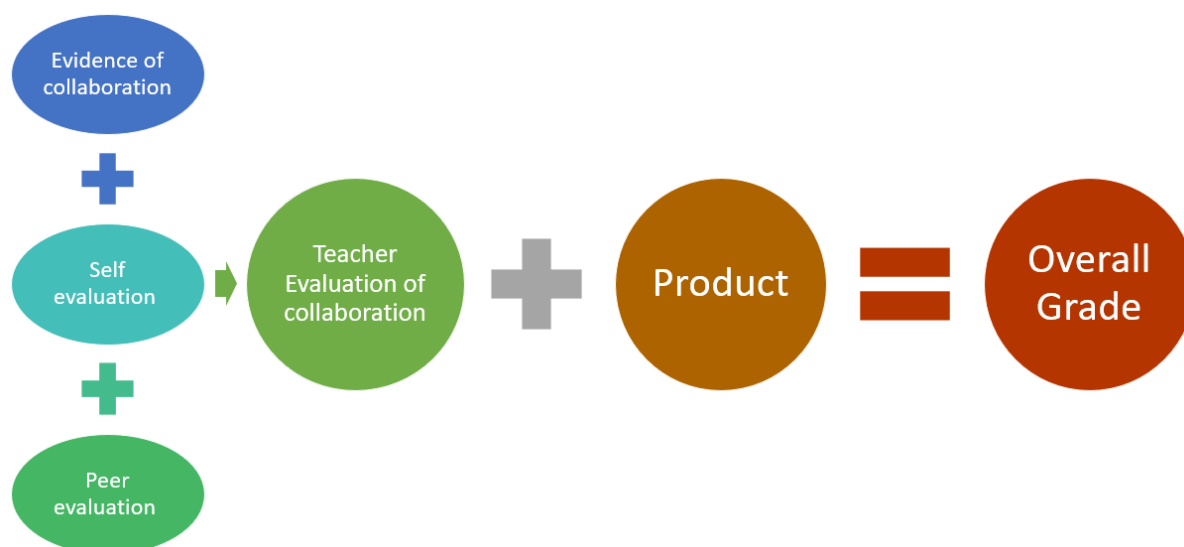
The question then becomes whether no contribution is better than a bad contribution.

Mechanisms to try and capture contribution (teachers reviewing workspaces, students self-reporting, diarying, teacher observation) would offer incomplete understanding and only reflect certain types of contribution. An over-emphasis on measurable contribution could, as with other documenting measures, lead to artifice and busy work. When considering a reporting mechanism, effort must be made to try and generate positive washback, with the reporting having some learning benefit in itself.

4.3.4.3.2 Utilising self- and peer-assessment to evaluation collaboration

Rather than contribution, two other potential lenses for assessing collaboration are self- and peer-assessment. As was clear in the discussion about students as assessment agents (see 3.5.6.3.7), within the high-stakes realm of EAP assessment, self- and peer-assessment need to be treated with caution, despite Kearney et al. (2016) advocacy for peer-assessment as an learning opportunity in itself. In a context where students have the assessment literacy to self- or peer-assess, such assessment lenses could have great value. In contexts where students may lack the assessment literacy to engage in such assessment a cost/benefit analysis of the learner training required would need to be made. In assessment literate cohorts, self- and peer-assessment would satisfy the principle of fairness and the reward orientation. So too would the practical and defensive orientations be satisfied as there is a straight forward assessment mechanism the students are already au fait with, that engages them in the process. However, the initial caveat of “In assessment literate cohorts” cannot be understated. EAP practitioners may not find themselves with such assessment literate students. As Fernand explained when talking about his context, self- and peer- assessment would be “something that's completely alien to the students” (Internet Transcripts, p. 138). As opposed to direct self- and peer- assessment, it may be preferable to employ teacher-mediated self- and peer-assessment as that may be more defensible. In this case, the evaluation of a collaborative assignment can be visualised as in Figure 4-12, which illustrates teacher-mediated assessment of a collaborative assignment.

Figure 4-12 Teacher-mediated assessment of a collaborative assignment



Teacher-mediation in itself is not a panacea. For example, consideration must be given how a teacher interprets conflicting reports of a collaboration, as was the case for Lorenzo.

4.3.4.3.3 The challenges of assessing collaboration and the case for cooperation

When looked at through the collaborative assessment prism, the challenges of assessing collaboration are manifest. As collaboration itself is not transparent, the assessment of it lacks transparency. Attempts to look inside collaboration greatly increase complexity of assessment, although there may be some pay-off in terms of the potential for fairer marking. The reward-orientation provides the best motivation to attempt assessing collaboration since even an imperfect measure is fairer than no measure at all. Set against that, assessing collaboration could quickly become onerous from a practical orientation and knowingly implementing a flawed assessment tool raises tension for the defensive orientation.

Given the thorny nature of assessing collaboration, the reader may be tempted to abandon the thought abandoning collaborative assignments altogether, or at least abandoning the inclusion of collaboration in an assessment scheme. Considering the former option, rather than a collaborative assignment, the task type could be changed to a cooperative one. This is a valid consideration and, in

some cases (e.g. if there is need for close reporting of students' individual competencies), may be the best way forward. This choice could be driven by practicality as well assessment concerns. As Gonzalo noted when discussing collaboration

if everyone tries to work on everything at the same time, it often takes a lot longer. And that, while that may be a good thing, it often isn't the most expedient or practical way to go about doing things (...). But assigning chunks to it I think, well may not be perfect, it might be the more expedient practical way to kind of move a collaborative project along. (Interview Transcripts, p. 144)

Rather than a unified, indivisible whole, students can work together with specific responsibility for one part of the assignment. Work could be divided on a content level, or at a task level although, as Kyle noted, "I think that the great conundrum is also finding a task in which all contributions can be equal" (Interview Transcripts, p. 213).

Dividing work on a content level could be as straightforward as each student taking responsibility for a particular section of the essay. Dividing work on a task level could involve one student taking responsibility for research and summarising, another for planning and organisation, a third the actual writing of the piece. The challenge, in a cooperative assignment, is trying to ensure equity between the responsibilities students are given. Gonzalo discussed trying to balance this and reach parity in a cooperation, with some students doing slightly more, but engaging with less challenging tasks and others doing more complex tasks but fewer of them (Interview Transcripts, p. 144). Even the relatively straightforward content division, with students being given responsibility for a particular section or paragraph of the essay, is not as easy as it seems as different parts of an essay carry different a different load. Introductions and conclusions do not require the same research as a contextualising or background paragraph, which in turn are more descriptive and less requiring of critical thinking than later discussion or evaluation paragraphs. It would also be important, in a

cooperative assignment, to emphasise the importance of coherence and cohesion to try and overcome patchwork essays developed by students who have no idea what is happening in other paragraphs. Gonzalo discussed the value of shared documents to enable this, so students could see and comment on each other's work. It would also have to be considered if each individual 'task' would lead to the desired learning goals so all the students reach the course objectives.

Since the learning gains of collaboration outweigh cooperation (see 1.2.4.1) and students are likely to encounter collaboration in future, abandoning collaborative assignments altogether may not best serve the students' needs. This brings us to the second temptation, that of not including collaboration in an assessment scheme. Since collaboration cannot be fully or accurately assessed, perhaps the solution lies with Nouri et al. (2017) and in seeing collaboration as a way of working rather than a competence to be measured. Humberto may agree as he felt the most value for students would come from the feedback they got rather than any mark for assessment purposes. The benefits and learning gains of students collaborating together will still emerge if collaboration itself is not assessed and, arguably, the collaboration itself may be more natural and authentic without the pressure assessment brings. In another teaching domain, ignoring the evaluation of the collaboration itself may be acceptable.

However, in the realm of EAP, with a responsibility to prepare students for their future studies and working lives, and in a course where collaboration is a stated learning outcome, collaboration should be assessed, however problematic that may be. This is the general position taken by the participants in this study, that if collaboration is a learning outcome, it has to be assessed (see 3.4). Part of the rationale for this is drawing the attention of the student to the value of collaboration since it carries weight in assessment. As Olga put it "if I tell them very specifically, like what my expectations are, they might work harder in that direction" (Interview Transcripts, p. 261). If collaboration is assessed, it matters. If it is not assessed, it does not. Assessing collaboration focuses students, and also focuses teachers. Part of the rationale for assessing collaboration is drawing teacher attention to collaboration so they may try and facilitate positive collaboration between students.

4.3.4.3.4 Criteria driven assessment of collaboration

During the interviews, participants spoke of the need for clear assessment criteria which are shared with students as this helps inform the students about the expectations of the assignment. Norton discusses the prevalence of this concept in HE arguing it “is pedagogically sound on a common sense level and on principles of equity, fairness and accountability” (2004, p. 689). However, they go on to explore the negative impact of criteria in the sense that such criteria may become the full focus of student attention and ultimately have “a deleterious effect on students’ learning” (Norton, 2004, p. 487). This reflects the participants’ and Strijbos’ (2016) concern about performative collaboration and the concern that the criteria shape the student experience and are ultimately reductive (Davidson, 2004; Lantolf & Frawley, 1985). When behaviours are codified in criteria they become the target behaviour for both teachers and students at the expense of other potentially natural and beneficial behaviours.

Given the conflicting views on the need for and the dangers of explicit criteria it is useful to recall Davidson’s placement of criteria not as an end in themselves but as the basis for discussion. Criteria are unlikely to be able to capture the full complexity of collaboration and the concern prompted by Model D of needing to distinguish between below average, average and above average collaboration comes to the fore. If communication with the group were a criterion, what would an above average performance look like and how could that be expressed so it did not become a target for students, prompting them to send unnecessary messages to be seen to be high performing communicators?

However, the process of developing criteria, of teachers negotiating criteria, will be highly beneficial in engaging them with careful thought of what ‘good’ collaboration looks like in their context. Developing criteria makes us consider how we want students to work with each other, what we value. The criteria below (Table 4.1), taken from the second draft of the user guide produced as a practical output of this dissertation (see Appendix 13, p. 35) reflect what is valued by this researcher.

Table 4.1 *Sample criteria for assessing a collaborative assignment*

	Satisfactory	Unsatisfactory
Involved throughout the process		
Contributed to the work		
Engaged with the contributions of others		
Accepted and worked with feedback		
Supported teammates		
Maintained clear communication		
Maintained positive relationships with teammates		
Took accountability for the whole work		

Note, it tries to capture the two dimensions of contribution, that of contributing to the work, and contributing to the process of the groupwork. That is not to suggest these criteria are definitive. They are not, and as Olga suggested, “perhaps, there must be more studies into [assessment criteria for collaboration]” (Interview Transcripts, p. 260). However, until such research is conducted, co-constructing assessment criteria for collaboration with other teachers would be a valuable exercise in developing collaborative assessment literacy. As Kyle also notes, through careful design of criteria you can mitigate against unwanted student behaviours (Interview Transcripts, p. 215).

We know, from the outset, that the assessment of collaboration is going to be flawed and necessitate compromises, no matter how many assessment lenses are employed or who conducts the assessment, be that the teacher-assessment, self-assessment, peer-assessment or teacher-mediated self- and/or peer-assessment. We also know the assessment of collaboration could include a lot of work and effort for very marginal gains. This necessarily means collaboration cannot be weighted too heavily in an assessment scheme, but should be weighted at a significant enough level to matter. To mitigate defensive concerns, it is worth recalling the quote which opens this dissertation: “In a

programme of assessment, individual assessments are purposefully chosen in such a way that the whole is more than the sum of its parts. Not every individual assessment, therefore, needs to be perfect” (Van de Vleuten et al., 2015, p. 641). The evaluation of the collaboration itself is not going to be perfect, but, taken in concert with assessment of the final product and, potentially, individual complementary assignments, a collaborative assignment can productively sit amongst other course assessments.

4.3.4.3.5 Assessing awareness, not performance

Another way forward may be linked to the discussion of grounded reflection as an individual complementary assignment. Rather than assessing a specific collaboration, which has already been found to be problematic (Boud & Bearman, 2022; Pitt, 2000), assessment could be geared towards awareness of collaboration. Humberto was against assessing collaboration itself, but felt it was necessary to talk about collaboration with students because “if I can build an awareness of these things in the student, then that prepares them for later in life” (Interview Transcripts, p. 172). Assessing awareness sidesteps the issue of collaboration being a stable construct and each collaboration being a different performance. Although reticent about assessing collaboration itself, Fay felt assessing awareness of collaboration was a potential way forward, saying “I love that idea. Yeah, I think that's wonderful” (Interview Transcripts, p. 125). Gonzalo concurred, feeling assessing awareness of collaboration “would seem more, more feasible, and probably easier to measure” (Interview Transcripts, p. 152). Cristobal also agreed, feeling assessing awareness of collaboration would be preferable to trying to assess collaboration itself and could also recognise the learning that could take place during a problematic collaboration with a poor final product (Interview Transcripts, p. 90). Laura was an outlier, feeling assessing awareness of collaboration was even more abstract than assessing the collaboration itself (Interview Transcripts, p. 237) but Fay saw such assessment preparing students for future collaborations, contending “the more you think about it, the more aware

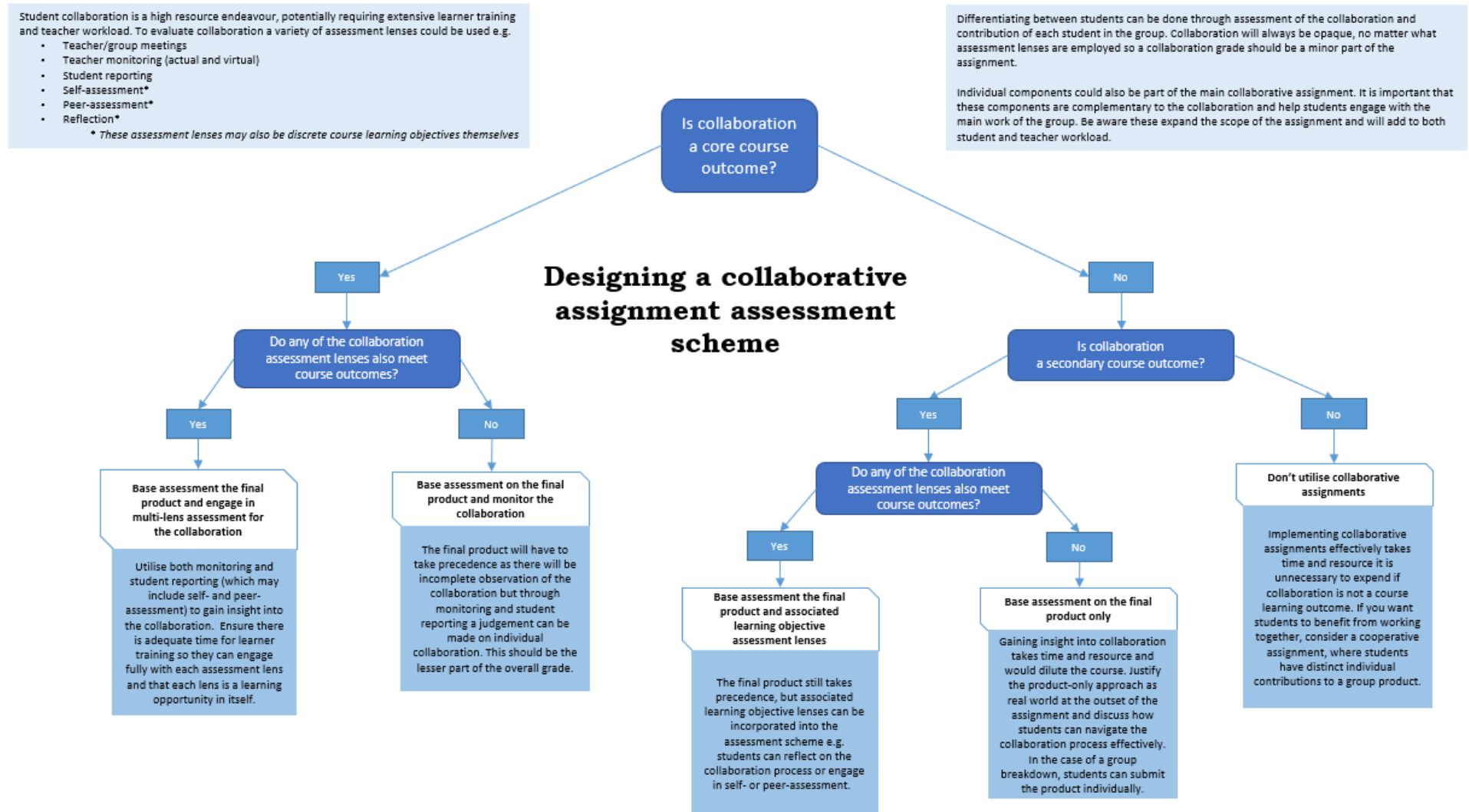
you are of what happened and why it happened and your role within that, not just it's other people's fault, the better prepared you might be for the next encounter" (Interview Transcripts, p. 125).

4.3.4.4 *Learning outcome directed lens selection*

Decision making for the design of a collaborative assignment marking scheme should be driven by the importance of collaboration as a learning outcome, as that is what determines how thoroughly collaboration has to be assessed. The need to assess collaboration grows the more central it is as a stated outcome. As Andrea put it, the design of an assessment scheme "all depends on the objective of the teaching plan" (Interview Transcripts, p. 9). Figure 4-13 (see also Appendix 12, which is a larger version of the image) illustrates the decision tree and choices that have to be made, depending on the importance of collaboration as a learning outcome and wider learning objectives that could be met through the utilisation of different assessment lenses. There is also a recommendation for an assessment scheme for each branch of the decision tree.

Exploring EAP practitioner collaborative assessment identity

Figure 4-13 Designing a collaborative assignment assessment scheme



If collaboration is not a learning outcome, the recommendation is that collaborative assignments are avoided. This is because collaboration takes time and resources and in already busy EAP courses the work that would go into facilitating collaboration could well be better spent elsewhere. In the case where collaboration is a secondary learning outcome and potential assessment lenses do not help meet other learning outcomes, the recommendation is to base assessment solely on the product. This is an unexpected recommendation that was not envisaged at the outset of this dissertation. However, if there are no relevant learning gains that can come from individual complementary assignments or assessing the collaboration itself, then a product-only approach is justified. However, that does not mean collaboration itself can be neglected. It would still have to be facilitated and supported, with mechanisms in place to address group dysfunction. This reflects Fay's approach when talking about the need for teacher-group meetings, classroom observation and potentially group mediation in case of a problem during the facilitation of collaborative assignments (Interview Transcripts, p. 124), seeing them as vital elements despite their not being graded. Humberto also takes this line, discussing the collaboration with students although "I don't grade them, according to that, but I do, I am interested in it and I want to raise their awareness" (Interview Transcripts, p. 162). There should likely also be scope for individual submission of final work in the case of full group breakdown.

Where collaboration is a secondary learning outcome and some assessment lenses would help meet other learning outcomes, then lenses which have learning benefit can be employed alongside the main group product. This would have to be clearly rationalised to students so they understand the purpose of and need for each part of the assessment. When collaboration becomes a core course outcome, it necessarily carries more weight and will require more robust assessment. If other assessment lenses do not contribute to expected learning outcomes, assessment will be product-focused again. However, collaboration will also have to be assessed as a core learning outcome. There are two ways this could be done. The first is trying to assess the collaboration that happened, which can be done through observation and student reporting. Given that these will give incomplete insight,

the collaboration grade would necessarily be the smaller part of the assignment. Rather than assessing the collaboration that was, attention could turn to awareness of collaboration, incorporating a individual reflection assignment geared towards reflecting on collaboration. If a student is able to demonstrate understanding of their own roles, contributions and behaviours and how their team worked, this would indicate their awareness of collaboration. The final branch of the decision tree is when collaboration is a core learning outcome and potential assessment lenses also have utility in addressing other learning outcomes. In this case, a multi-lens assessment approach can be used, with each lens rationalised to students so they know the purpose of each and how they all cumulatively fit together as a full scheme of assessment.

As is clear, no single assessment scheme will fit every context or course. In each case the appropriate approach will have to be rationalised 'in-house'. The benefit of this is that the assessment scheme will be fully understood by the teachers enacting it and the negotiation of it could prompt fuller engagement. Teachers from across different collaborative assessment archetypes will better understand and potentially shape the rationale for the assessment scheme.

4.3.5 Layer 5 – Evaluation: Using the dimensions of collaborative assessment identity to evaluate a collaborative assignment assessment scheme

It was while initially drafting the evaluation of this dissertation that the need for this fifth layer in the collaborative assessment algorithm became apparent. This algorithm provides a pathway of decision-making for the EAP practitioner designing an assessment scheme for a collaborative assignment. There was initially, however, no mechanism to evaluate the assessment scheme that was developed other than potentially putting it to use and seeing what happens. Yet in the way the work of this dissertation had to be evaluated, so would the output of the collaborative assessment algorithm.

In terms of evaluation, there is value in considering a potential collaborative assignment through the lens of test usefulness (see 1.4.5), particularly in relation to impact, since the decisions

made in designing an assessment scheme for a collaborative assignment are going to have washback on teaching and learning. However, there is also the possibility of unifying the theoretical and practical output of this dissertation, with collaborative assessment identity providing evaluation criteria for the output of collaborative assessment algorithm. Once an assessment scheme is developed it can be evaluated, pre-use, via criteria based on the key principles and guiding orientations of collaborative assessment identity. Table 4.2 (below) has guiding evaluation questions for the evaluation criteria of simplicity, transparency, fairness, reward, practicality and defensibility.

Table 4.2 *Using collaborative assessment identity as criteria to evaluate an assessment scheme for a collaborative assignment*

Simplicity	Will the assessment scheme be easily understood by the institution, exam boards, the teacher, and students?
Transparency	Does the assessment scheme make it clear to students how and why they are going to be assessed? Is the purpose of each assessment lens clear? Is the relationship between the assessment lenses clear?
Fairness	Are students going to get fair credit for the work they do? Is the assessment scheme designed so students should do equitable amounts of work?
Reward	Will the work students do be recognised? Will the assessment bring learning gains to the students? Will the assessment lenses have positive washback?
Practicality	Is the assessment manageable for students and teachers? Are the teachers and students familiar with any collaborative workspaces that should be used? Is the effort required to conduct the assessment worth the information gained?
Defensibility	Will the assessment help students meet and demonstrate desired learning outcomes? Is the assessment scheme robust, justifiable and fit for purpose?

Utilising the dimensions of collaborative assessment identity as evaluation criteria for a collaborative assessment means the assessment scheme is considered from the perspective of the concerns and beliefs of EAP practitioners.

4.4 Application of the algorithm to the opening vignette

This dissertation opened with a vignette, illustrating the challenges of assessing a collaborative assignment in EAP, which could take place over an extended timeframe with students working face-to-face, virtually, synchronously and asynchronously. There was never an intention to suggest an assessment scheme for such a case and the exact type and nature of appropriate assessment lenses cannot be known without full knowledge of the course, the learning objectives and the other assignments in the course. However, considering the vignette in light of the findings of this dissertation it is clear that various assessment lenses could help address some of the challenges. A submission statement may have encouraged Tomas to engage with the feedback more thoroughly, as might a viva voce, as he would have had to take responsibility for the final version of the work. Alternatively, an evaluative judgement task would have encouraged him to rationalise not engaging with the feedback and defend the quality of the first draft of the essay. A viva voce or teacher/group meeting mid-assignment might have helped bring Maria's valuable contributions to light, as would a reporting system, where Sara's dedication to referencing could have been reflected. Peer- and self-assessment may have also added to teacher knowledge of how the collaboration had worked, although the disagreement between Tomas and Maria may have skewed any summative peer-assessment grade.

Considering the situation through the collaborative assessment algorithm, the motivation for implementing a collaborative assignment is not apparent in the vignette, and nor is the preparedness of teachers and students to engage with such an assignment and self- or peer-assessment lenses. In terms of the second layer of the algorithm, purpose, there is not explicit purpose stated but the foundational purposes of a collaborative written assignment in EAP are likely two-fold. First is to develop student awareness of academic discourse and the genre of the academic essay, in preparation for future individual written assignments in the same genre. Second is to give the students the experience of working in a group to produce an assessed piece of work. Consequently, when it comes to assessment lenses, the following tentative mark scheme seems appropriate. It is qualified as tentative due to the lack of contextualisation in the vignette.

Table 4.3 *Tentative assessment scheme for the vignette assignment*

Assessment lens	Weighting	Rationale
The final product	50%	The quality of the essay is demonstrative of the success the group had in developing an academically sound piece of work.
Individual evaluative judgement of the final product	25%	This will demonstrate the student's awareness of academic discourse and the academic genre in question.
Grounded reflection on the collaboration	25%	This will demonstrate the student's understanding of themselves in the context of collaboration.

The greater weighting of the collaborative product of the assignment is to ensure the students give it the time and attention deserved. The weighting of the individual assignments is to recognise they will be substantive in themselves and will be the clearest indicator that the purpose of the assignment has been met. Learner training time should not be too onerous as preparation for the grounded reflection can come near the start of the course, when discussing collaboration with the students. A mid-course teacher/group meeting could also allow for discussion of the collaboration as it is happening. Evaluative judgement could be developed through the training regarding academic writing and adding a peer-feedback session for the second draft of the essay, after the initial teacher feedback on the first draft. Work on evaluative judgement would also be beneficial for the final written product as the focus of the training will be on academic voice, academic genre and appropriacy.

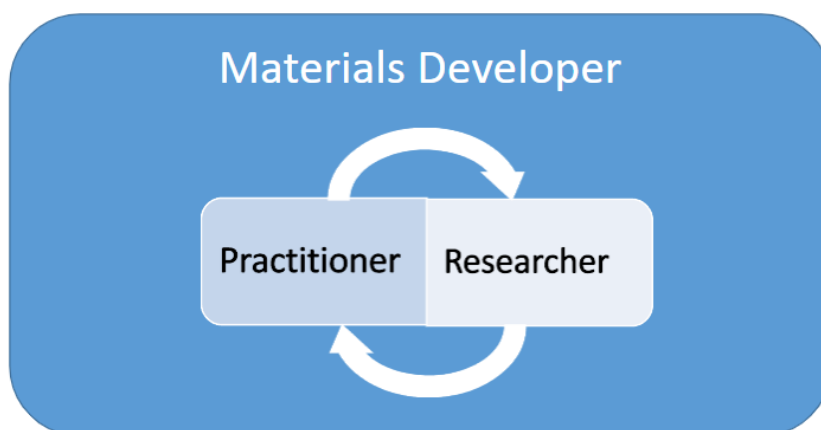
4.5 From theory to practice: From algorithm to guide

The process of this dissertation has been illuminating, not least in casting light on my own self-perception. Earlier in this work I located myself as an insider researcher (see 2.4.1.3) and in this I was maintaining my identity as practitioner, practitioner engaging in research, or practitioner-researcher. The role of practitioner-researcher is explored by Cullen et al. (2012), bringing together Mcleod's definition of practitioner research as "research carried out by practitioners for the purpose of advancing their own practice" (Mcleod, 1999, p. 8, in Cullen et al., 2012, p. 10) and Shaw's (2005)

contestation that practitioner research can be, as Cullen et al. (2012) put it, “rigorous and critically engaged with debates within policy and theory” (p. 10). Being a practitioner-researcher, I intended to engage with theory and explore practice for a practical benefit, not just for myself but for the wider EAP community. Beyond this role of practitioner-researcher a third dimension, a dominant dimension, has come to light. Rather than practitioner-researcher, I operate as a materials developer-practitioner-researcher.

The materials developer is a role championed by Tomlinson, who has done much to professionalise the field. He recognises that “materials development is one of the most important undertakings in applied research” (Tomlinson, 2016). After undertaking an MA in English language teaching and materials development, I argued for the role of the course book as trainer (Levrai, 2013), contending that teaching materials play a central role in teacher development, shaping their knowledge, practices and pedagogy. That belief in the ability of materials to help teachers become better or more effective teachers means my materials development identity encapsulates my roles as practitioner and researcher (see Figure 4-14). Practice should be informed by research, research can learn from practice, and the materials developer is the interface through which the knowledge and expertise of both is shared. This realisation re-emphasised the ambition for this dissertation to have a practical output, one which will benefit and serve as professional development for the EAP community.

Figure 4-14 The materials developer-practitioner-researcher



The collaborative assessment algorithm illustrates the pathway of decision making when developing an assessment scheme for a collaborative assignment and Figure 4-13 works as a snapshot of the considerations that have to be navigated. However, that figure alone would not be sufficient aid to practitioners in understanding the complexities of assessing collaborative assignments. Consequently, the starter guide to assessing collaborative assignments was developed. This is very much a work in progress and not intended as the final form of the guide.

4.5.1 Reactions to the starter guide

The draft starter guide was intended as proof of concept that a guide to assessing collaborative assignments would be appreciated by the EAP community and to indicate what would be valued in a guide to aid to practitioners when designing assessment schemes for collaborative assignments. The guide is the actualisation of the collaborative assessment algorithm, a practitioner-friendly distillation of the research of this dissertation. Participant evaluation speaks to the need for such a guide, given the uncertainties held around collaborative assessment. So too are both drafts of the guide considered a valuable resource that addresses that need, despite still being in the development stage. The suggestions for increasing the range of the guide are well taken, since assessment is only one aspect of practitioner work connected to collaborative assignments. While the guide, as it is, has value, practitioners need more support throughout the whole process of conducting collaborative assignments, from the initial planning, through preparation and facilitation, to assessment practices.

This need for expansion is particularly true in light of the concerns of some of the reviews that the guide could put practitioners off from engaging with collaborative assignments. This is taken as an encouraging sign, as it means the guide leads to careful consideration of the various challenges embedded in collaborative assessment. However, the guide would need to be fuller to assuage some of the concerns raised and better support teachers in meeting the challenges. Nonetheless, the response to the guide was very positive, indicating an already practically useful piece of work, fulfilling the intent behind its production.

It is envisaged that the final version of the collaborative assessment guide could follow the sequence of layers of the collaborative assessment algorithm, taking practitioners through the considerations of motivation, purpose, readiness and selecting lenses. Then an additional layer of material could be added to discuss the practicalities of facilitating a collaborative assignment (covering issues like grouping, set-up tasks and activities, and trouble-shooting) before the evaluation material of the final layer. In this way a comprehensive guide could be in reach.

4.6 Theory evaluation

As recommended by Birks and Mills (2023) evaluation is going to begin on a face validity level. As an EAP practitioner, I know I feel differently about collaborative assessment than individual assessment, indicating I hold a specific collaborative assessment identity, although I would not have been able to articulate that prior to this dissertation. The principles held by the participants are ones I recognise in myself when considering collaborative assessment and I have felt the tensions between the different orientations, wanting to be able to give fair reward for the work students have done, being practically hampered by not necessarily knowing what they have done, and understanding my assessment approach may be open to critique and must therefore be robust. Similarly, the collaborative assessment algorithm feels like it would have been a very useful tool to have access to when designing a collaborative assignment assessment scheme, providing a pathway of considerations and necessary decisions to fully rationalise an assessment approach. Feedback from early dissemination of the findings at conference has also encouraged the belief that these findings have value and utility.

Hadley, when defining the core methodological characteristics of GT (see Figure 1-9) had this to say about a Grounded Theory

The theory will be useful. It studies the concerns of a particular group of people and the ways in which they deal with these issues. The theory results in a heightened awareness, understanding, ability to predict and/or a greater sense of control, first for the group that was studied but then later potentially for people in other situations, but who are experiencing similar problems or concerns. (Hadley, 2017, p. 59)

This dissertation and accompanying starter guide to assessing collaborative assignments fulfils each aspect of that definition. The dissertation is centered on the concerns of EAP practitioners dealing with the assessment of collaborative assignments. The findings provide more awareness and understanding of the beliefs and practices of EAP practitioners, with the starter guide offering a greater sense of control in terms designing rationalised and justified assessment schemes for collaborative assignments. Initial response to the guide demonstrates that it is considered useful by EAP practitioners and objectively brings together findings from literature and research to generate a practitioner friendly guide through the complexities of designing a collaborative assignment. The assessment of collaborative assignments is by no means a uniquely EAP phenomena, meaning the approach to assessing collaborative assignments could be of value to other HE professionals engaged in collaborative assessment. In the way this dissertation fulfils Hadley's definition, so too does it meet Turner's in being "accurate, understandable and enlightening to individuals who have some familiarity with the social phenomena under investigation, either as participants or as 'lay' observers" (1983, p. 347).

Evaluation of the work using the classic criteria of Glaser (1978) can be seen in Figure 4-15. Looking through these lenses the theory is firmly grounded in the data and has implications for practice and further research.

Figure 4-15 Evaluation of the work using Glaser's 1978 criteria

Grab – the theory is interesting, memorable and people want to use it.

The theory of collaborative assessment identity makes sense. If teachers hold an assessment identity then it seems likely this identity will shift with different types of assessment. The three key principles, and the primacy of fairness, would also seem appropriate for complex assessment where student outcomes are impacted by other students. The pull between the reward and practical orientations also resonates, as does the centrality of the defensive orientation given the high stakes assessments EAP practitioners are involved in. The collaborative assessment algorithm also seems like an appropriate tool, drawing attention to the layers of thought processes that have to be navigated when designing a collaborative assignment assessment scheme.

Fit – the theory is indicated by the data and is applicable to the data and covers all the data.

The collaborative assessment identity was an unexpected outcome of this dissertation, and determining the dimensions of it was not the original intent of this study. However, the key principles and guiding orientations were evident from the data. All participants can be located in the prism of the collaborative assessment identity, although it must be noted that such an identity is dynamic and changing. So too does the assessment algorithm cover all the necessary steps and potential assessment lenses the participants would need to develop a collaborative assignment assessment scheme.

Work – the theory can explain, interpret and predict what is happening in the area under study and must be relevant to it.

Assessment of collaboration in EAP is an area of scant research so this dissertation necessarily adds to area of knowledge. Collaborative assessment identity helps us understand what drives EAP practitioners when it comes to collaborative assessment and the collaborative assessment archetypes pose food for thought for course leaders or language centres looking to implement collaborative assessment. Knowing a teaching body will have varied attitudes and beliefs towards collaborative assessment helps in the development of a justification for a specific collaborative assignment assessment scheme.

Modifiable – the theory can adapt to new data in an ever-changing world

This dissertation takes the first step in defining collaborative assessment identity, although its boundaries may change as the world changes. Collaborative assessment may take on more resonance in an era of LLMs, where students are given more human-human tasks to complete. As the place of collaborative assessment changes, so might the collaborative assessment identity, which is not fixed. Practitioners may move between archetypes as different orientations or principles come to the fore. So too can the collaborative assessment algorithm adapt to different teaching and educational contexts, with the intent of the algorithm being a modifiable rather than a fixed assessment scheme.

Consideration will now be given to the evaluation criteria laid out by Charmaz (2008): credibility, originality, resonance and usefulness. Credibility is vital and the foundation upon which this dissertation is built and contributing originality to the field of EAP and collaborative assessment is also highly valued. However, the greatest satisfaction for this materials developer-practitioner-researcher will come from the resonance and usefulness the work has for fellow EAP practitioners. The true evaluation of this dissertation will come when the findings are more widely disseminated and, hopefully, embraced by the EAP community.

4.6.1 Credibility

This dissertation draws on a wide body of literature to contextualise the focus of the research, taking in elements like EAP, collaboration, collaborative assessment and teacher assessment identity. The methodology has been clearly described, with transparency about challenges and changes. The data is richly presented in the paper, with participant voices fronted. The theory of collaborative assessment identity is believable and its dimensions rooted in the data. So too does the collaborative assessment algorithm bring together both literature and data from this research to provide a practical guide to the development of collaborative assignment assessment schemes.

4.6.2 Originality

Collaborative assessment is a recognised area in need of further research (Meijer et al., 2020; Storch & Knoch, 2023). There have been calls in the literature for research into how practitioners operationalise collaborative assessment, their rationales and practices (Meijer et al., 2020). The EAP practitioner is under-researched (Ding & Monbec, 2024), with their knowledge, beliefs and experiences rarely explored. The need for EAP teacher knowledge in assessment was the title of Schmitt and Hamp-Lyons' 2015 paper and in it they lament that "most of the assessment work of EAP practitioners is hidden away and thus contributes little to theory-building or wider understandings of EAP assessment in practice" (2015, p. 3). The approach and rationale of EAP practitioners toward

assessment generally is under-researched, and this is also the case for collaborative assessment. This dissertation contributes to each of those areas of knowledge, providing a new perspective on collaborative assessment by fronting the practitioner voice. It builds on the existing concept of teacher assessment identity to develop an understanding of the EAP practitioner's collaborative assessment identity. It draws on discussions of approaches to assessing student collaboration in the literature and from the research data from the participants to propose the collaborative assessment algorithm, a practical, useful tool.

4.6.3 Resonance

The findings of this dissertation resonate deeply with the author. While that may be expected, early dissemination has also indicated that the need for this work and the utility of its findings are keenly felt by fellow EAP practitioners. Participants' positive response to process of the interview and having opportunity to discuss collaborative assessment indicate this is an important area where professional development would be appreciated. The response to the starter guide shows it could be tool in aiding such development.

4.6.4 Usefulness

As with the previous criteria, the usefulness of the collaborative assessment identity and collaborative assessment algorithm are apparent to this researcher. Understanding of collaborative assessment identity and the collaborative assessment archetypes can help teaching teams understand the multitude of attitudes that surround collaborative assessment. It can aid in discussion and rationalisation of assessment schemes, and provides a means to evaluate such schemes through the criteria of simplicity, transparency, fairness, reward, practicality and defensibility. The collaborative assessment algorithm, as realised through the starter guide to collaborative assessment, provides a practical pathway through the decision-making process of developing an actionable and fair assessment scheme for collaborative assignments. The need for professional development around

collaboration was underscored by Dolly, who questioned the impact lack of knowledge on behalf of teachers could have. As she put it

some teachers who believe, who are proponents of collaboration and see the value, not just be, not only within English language courses, but also beyond, right? But then, you know, it's, this is not a firm belief for all language teachers, right? There are language teachers who are like, "Hey, I'm, I'm only here for the language. There's nothing else, like, I'm not gonna do anything else. I'm only here for the language." Right? So ... and then if you want to teach collaboration, do you not have to believe in collaboration? There are people who don't believe in collaboration, right? This kind of like, if I am a terrible discussion facilitator how am I supposed to teach that? How do I spot the student who is not facilitating when I am a poor facilitator and can't, can't facilitate? So, so some of that, you know, like we, even if we see the value are, in collaboration or in discussion, are teachers trained to teach those? (Interview Transcripts, p. 111-112)

Given that EAP practitioners will engage with collaborative assignments, increasing our understanding of how to approach such assignments is of tangible use, with implications for teaching practices and student experience.

4.7 Chapter conclusion

Important work has been done, examining the thorny issue of summative assessment in collaborative assignments. The EAP practitioner, an important player in practice but near silent voice in the literature, has been fronted in this dissertation. Through evaluation of different methods of assessing collaborative assignments key principles and guiding orientations came to light. Together they form the dimensions of collaborative assessment identity, which is distinct from teacher

assessment identity (Looney, et al., 2018), shaped by the specific concerns engendered by summative collaborative assignments in EAP HE contexts.

The three key principles underpin the guiding orientations, encompassing collaborative assessment identity. Such an identity is not fixed. It can change as different principles or orientations gain dominance, with a rise in fairness seeing a shift toward the reward orientation while a rise in simplicity will see a move toward the practical orientation. There is an ongoing pull between these principles and orientations, creating a dynamic collaborative assessment identity.

Understanding of the range of opinions and attitudes held by EAP practitioners toward collaborative assignments can be characterised through eight collaborative assessment archetypes, existing on a continuum from those diametrically opposed to collaborative assignment, the Never-everers, to those thoroughly comfortable with it, the All-seeing-eyes. Within those archetypes we can find the I-in-teamer, whose drive is being able to recognise the work done by an individual, strongly inhabiting the reward orientation. The Real-worlder, driven by simplicity and transparency, endorses product-only assessment, an approach also favoured by Pragmatic real-worlders, although they are guided not so much by belief in the validity of product assessment so much as the practical orientation of what is actually feasible and achievable. The All-I-can-doers understand aspects of their approach to assessing collaboration are compromised, but accept that compromise weighed against the benefits of collaboration. Most EAP practitioners inhabit the Would-if-I-coulder, demonstrating an appetite to improve their collaborative assessment literacy, to satisfy the defensive orientation and engage in principled, rationalised and justified collaborative assessment. Once that literacy is gained, practitioners will be able to move into the Best-I-can-doer archetype, delivering fully justified collaborative assessment they can believe in.

To aid with this, and to contribute to the discussion needed about assessing collaborative assignments, the collaborative assessment algorithm identifies the layers of thought processes that need to be navigated when designing an assessment scheme for a collaborative assignment. It provides a pathway of considerations and decision-trees that must be navigated when designing a

collaborative assignment assessment scheme. Beginning with fundamental questions like the motivation for establishing a collaborative assignment and questioning what “good” collaboration looks like, the algorithm raises the issues generated in the literature and through this research, encouraging asking vital questions. The algorithm was realised as a starter guide, a practitioner facing text designed to help EAP practitioners on a practical level. The need and usefulness of this guide was apparent to the reviewers with suggestions for various additions and extensions to make a guide that fully supports EAP tutors working with collaborative assignments. The need for such a guide, however, is not limited to EAP. As Jerry contended, “I think beyond the EAP practitioner, I think it would be useful to any academic or tutor who is working in teaching and learning with collaborative assignments” (Reviews of the starter guide p. 13).



CHAPTER 5

CONCLUSION

5 CONCLUSION

This chapter closes the dissertation, opening with a summary of the project and main findings. It will review the principles and orientations which came to light through data collection, and how these encompass EAP practitioner collaborative assessment identity and collaborative assessment archetypes, the theoretical outputs of this dissertation. It will then recap the practical output, the collaborative assessment algorithm, as embodied in the starter guide.

Consideration will then be given to the research question of this dissertation “How can we, if we can, best assess collaborative assignments in EAP?” and areas of interest that emerged during the research. Consideration will also be given to the research questions and hypotheses from the pre-pandemic research plan. Although data was not gathered with an aim to address these, the rich data set does allow for some comment.

No research work is perfect, and this dissertation is no exception. There will be discussion of the limitations of the research, which also illuminate future directions of research to help build on the work of this dissertation. There will also be consideration of the implications and significance of this work, both within EAP and HE more widely. The chapter, and dissertation, ends with some closing thoughts and cautious optimism for the assessment of collaborative assignments.

5.1 Summary of findings

This dissertation has explored EAP practitioner attitudes and beliefs around assessing collaborative assignments. The assessment of collaboration is no easy matter to explore, with literature either lacking or pointing to complexity and difficulty, particularly when it comes to summative assessment. We do not know how to best assess collaborative assignments, or what informs practitioners' collaborative assessment decisions. This is true of any field of study, EAP included. By following a constructivist GT approach, however, some progress has been made, bringing to light EAP practitioners' collaborative assessment identity and leading to the development of a collaborative assessment algorithm as practically actualised in a starter guide to assessing collaborative assignments.

This dissertation was shaped by the COVID pandemic, which necessitated a change of research plan. That change has been to the benefit of this research, evolving from a postpositivist validation exercise into a constructivist GT exploration. GT is well suited to under-researched phenomenon and the pandemic moved this dissertation from research looking at a single teaching context to a broader work, examining the lived experience of EAP practitioners in a range of teaching contexts. The breadth of experience of participants strengthens this research, as the experience of collaborative assignments does not seem to be tightly context bound, with participants in varying teaching cultures and contexts having the same kind of doubts and concerns, and sharing the desire to do better when it comes to collaborative assessment. It is also refreshing to have research into assessment fronting the teacher voice, as the teacher is an essential yet underrepresented assessment agent in the literature.

Through discussion of the pre-interview task, which asked participants to evaluate different models of assessment for a collaborative assignment, rationales for and against various assessment approaches were uncovered. Adopting a product only assessment approach found little favour, due to its limitations in terms of differentiation between students, which led to perceived unfairness, as illustrated by free-riders potentially getting benefit for the work of others. In favour of the product only approach is the realworldness of it. Outside of the academy, if you collaborate on a piece of work,

the only important metric is the quality of that work, not how it was produced. However, given the role of education, and EAP particularly, to support students in developing their skills, this argument falls short.

Moving beyond the product and assessing collaboration itself was an area of interest, although doubts abounded about how this could actually be done. Challenges involve insight into collaboration and the myriad of ways students can contribute to a collaboration, some of them invisible to the teacher. Potential assessment lenses to bring the collaboration into focus, like self- or peer-assessment, or reflection, were all greeted with a level of concern when it comes to high stakes summative assessment. The paradox emerged of participants feeling collaboration should be assessed if it is a learning outcome, yet not being confident in the mechanisms to assess it. The assessment scheme which found most favour was utilising individual complementary assignments alongside the main collaborative task, since this provided the opportunity for differentiation between students. There is need for caution though, as this expands the task, adds to the workload of teachers and students, and does not, ultimately, shed additional light on to the collaboration itself.

Through these discussions, three key principles which guided EAP practitioners in their assessment of collaborative assignments were illuminated. The foremost principle was fairness, which held the strongest sway. The nature of collaboration, with students dependent on the work and the working relationship with others, seem to particularly trigger concerns about fairness in relation to unequal contributions, asymmetric collaborations and students not getting the overall results they deserve. Allied to this was the second principle: transparency. It is essential that assessment practices are clear to students, that they know what is being assessed and how it is being assessed. Transparency holds a second concern in relation to collaborative assignments, in that collaboration itself is not a transparent process, problematizing its assessment. With this being the case, the appetite for transparency was not just that students would know what is being assessed, but that teachers would also have a clearer conception of what they are assessing when assessing collaboration and sound knowledge of how to go about it. The final principle is simplicity. This does not speak to a

desire for simplistic assessment, but rather that assessment is not over complex or convoluted. Assessment has its place alongside learning, and can be an aid to learning, but it must not gain ascendancy at the cost of learning. Assessment has to be clear and easy to follow in order to be transparent and, therefore, fair.

These three principles work in concert and encompass the collaborative assessment identity of EAP practitioners, informing three guiding orientations, which affect their decision making and assessment approaches. With the fairness principle to the fore, the reward orientation is clear, with a desire to ensure students both benefit from and get reward for the work they have done. This requires looking past the unified piece of collaborative work to the process of development in order to differentiate between students. Nonetheless, this could be a time-consuming and onerous process, which runs afoul of the second orientation: practicality. The practical orientation guides participants to assessment which is feasible, manageable and easy to operationalise. Workload for teachers, and students, is a concern, with assessment needing to have a positive impact on learning outcomes. The final orientation is the most complex, informed by experience, expectation and, perhaps most crucially, the HE environment in which EAP operates. The defensive orientation speaks to the need for assessment to be clearly rationalised so that it can be justified to different stakeholders. It requires anticipating problems or challenges and trying to mitigate or avoid them from the outset. The importance of environment can be illustrated by considering the role of self- and peer- assessment. The EAP participants in this study valued the role students could play in their own assessment and the assessment of others but were very aware of working in the HE context and having to make high-stakes assessment decisions. This ecological reality pushed them from their inclination toward valuing self- and peer-assessment for formative purposes and toward the role of reluctant overlord, advocating a teacher-centric assessment approach for summative grading. Teacher assessment is easier to explain and justify and is likely going to generate less scope for student complaint, satisfying the defensive orientation.

Further to collaborative assessment identity, collaborative assessment archetypes could be discerned. Teachers are not fixed within one archetype, and may move between them, but such archetypes are useful to broadly categorise varying attitudes to collaborative assessment. These archetypes lie on a continuum, from those fully against collaborative assessment as an assessment approach to those fully comfortable and confident in assessing collaboration. The former, the Never-everers, are guided by fairness and cannot reconcile collaborative assignments with fair assessment practices. The latter, the All-seeing-eyes, are a theoretical construct as the opposite end of the spectrum to the Never-everers. The researcher of this dissertation, after giving careful thought to the assessment of collaborative assignments, should be well placed to fully inhabit the All-seeing-eye but does not do so because the assessment of a collaborative assignment will always be compromised in some way. Those compromises might be well rationalised and justified, with measures taken to mitigate problem areas, but there has to be recognition that the assessment of a collaborative assignment will never be perfect. We can do better, we can be fairer, we can be Best-we-can-doers, but absolute certainty in assessing collaboration seems, to this researcher at least, out of reach.

Considering the other archetypes on the spectrum, the I-in-teamer values the benefits students get from assessment and the opportunity to reward students for their individual work. The individual has to be seen and recognised. This can be fully opposed to the aspiration of collaboration, in which the individual should not be distinguishable in the work of the group. For I-in-teamers, individual complementary assignments would be a welcome addition to a collaborative assignment assessment scheme. In contrast, Realworlders would have no need for individual complementary assignments, believing in the sanctity of the product. Good collaboration produces good work, therefore good work indicates good collaboration. The logic of this can be challenged through consideration of product glare, where the quality of the product blinds the assessor to the quality of the collaboration. Good work could also stem from asymmetric collaboration, with a single student dominating the work, or group members compensating for a free-rider, or it could result from students compromising the collaboration by taking a cooperative approach. Nonetheless, Realworlders consider it beneficial for

students to be assessed the way collaboration will be measured after graduation. Realworlders are the not only archetype who follow product-focused assessment. Pragmatic-realworlders share a superficial sameness in this regard, although they are less convinced the product is the best measure of a collaboration. However, the practical orientation comes to the fore, encouraging Pragmatic-realworlders to look for the most efficient assessment approach and dismissing other assessment avenues that would require additional time and resources.

In favour of collaborative assessment are the All-I-can-doers, who value collaboration and, while acknowledging there are flaws in their current approach to assessing collaboration, feel the benefit of collaboration outweighs issues in the assessment scheme. For example, they may be aware they are not certain the contribution each student made, but accept unequal contributions happen in collaboration. That is not to say All-I-can-doers are fixed in their practices and are not desirous of development. Given the opportunity they would take on the archetype inhabited by the majority of participants: the Would-if-I-coulder archetype, who demonstrate an interest in more effective ways to assess a collaborative assignment. Since collaborative assessment is an expertise gap not generally covered in teacher training or professional development, it was an area where participants feel an opportunity to improve their practices, demonstrating their assessment literacy. As an aid in developing this literacy with regards to collaborative assignments, the collaborative assessment algorithm was developed and, should it be effective, it will aid practitioners in reaching the final archetype, the Best-I-can-doer, who is able to facilitate effective, principled and fairer collaborative assessment.

The intent behind the collaborative assessment algorithm is to help EAP practitioners navigate the development of an assessment scheme for collaborative assignments, or evaluate the assessment scheme of an already existing assignment. It starts with the fundamental considerations of why a collaborative assignment should be employed. This sets the stage for understanding the purpose of the assessment and how the results will be interpreted and used, speaking to the validity of the assignment. The next layer of the algorithm leads to consideration of readiness. This involves

developing a shared conception of collaboration and weighing the assessment literacies of teachers and students in the potential assessment lenses that could be employed e.g. self- and peer-assessment. So too would teacher and student competence in target collaborative workspaces be a consideration at this stage. This layer helps identify areas of teacher or learning training that may be needed as part of the set-up of the assignment.

In the fourth layer practitioners are asked to consider the pros and cons of different assessment lenses, to determine which lenses might be appropriate for use in their contexts. The algorithm then introduces a decision tree to help decide which assessment lenses to employ, guided by the importance of collaboration as a learning outcome, and the other learning outcomes that could be addressed through complementary assessments linked to the main collaborative assignment. The final layer of the algorithm brings the theoretical output (collaborative assessment identity) and practical output (the collaborative assessment algorithm) of this dissertation together. A tool for evaluating an assessment scheme for collaborative assignments is proposed for layer five of the algorithm, utilising the dimensions of collaborative assessment identity as evaluation criteria. In this way we can look at an assessment scheme for a collaborative assignment and consider it in terms of fairness, transparency, simplicity, practicality, reward, and defensibility.

The collaborative assessment algorithm was actualised as a starter guide, written as a practitioner facing text. Although not a finished product, the draft was developed to gauge the potential of such a guide. The draft guide was disseminated to selected participants and their feedback spoke to the need for such a guide and the usefulness of the guide they reviewed. The coverage of the guide could and should be expanded, to encompass collaborative assignments more fully, from conception, through preparation, to delivery, and assessment. So too should the target audience of the guide be expanded beyond the EAP practitioner to the wider HE community of lecturers and faculty engaging with collaborative assignments.

5.2 Research question and emergent areas of interest

The main research question of this dissertation was “How can we, if we can, best assess collaborative assignments in EAP?” Clearly, there is no simple, readily apparent way assess collaborative assignments that is applicable in all contexts. In order to assess a collaborative assignment, it is likely more than a single assessment lens will have to be employed, with the ‘best’ lens(es) being dependent on specific factors like the assessment literacy of students, time and resources available, and importance of collaboration as a learning outcome. When Gonzalo was discussing assessing collaboration he noted that there might be an ideal way to do something but that there also had to be recognition of “what we can practically do in a course and how much we can actually observe in a course” (Interview Transcripts, p. 151). Humberto concurred, explaining the problem with assessing collaboration itself is that “even though that ideally would be a good way of doing it, a good way of assessing, I’m just not sure about the practicality” (Interview Transcripts, p. 168). That is to say, the best assessment scheme for a collaborative assignment will have to be decided on a case-by-case basis, following the collaborative assessment algorithm as a guide through the layers of thought processes and decision making.

Three further areas of interest which emerged during this research were if collaboration is a stable skill, if collaboration, as a complex and opaque construct, is actually assessable, and the question of what is needed to make teachers more comfortable and confident assessing collaborative assignments. It is not within the data or scope of this dissertation to answer the questions definitively but the stability of collaboration as a skill seems questionable since each collaboration has numerous complex factors (e.g. group members, task type, task topic) that will impact an individual’s performance as recognised by Fay and Boud and Bearman (2022). However, collaborative tendencies may be more stable, with some people generally preferring to work alone while others may generally like collaborating. For Kyle, adaptability is key, contending

I think, probably if someone is a good collaborator, in general terms, they're probably more likely to say that a big part of it is being adaptable. So they're more likely to adapt to different contexts and circumstances. But being able to adapt to different contexts and circumstances is one of the key skills of collaboration. (Interview Transcripts, p. 216)

Although adaptable, people may also tend to operate in certain ways in groups. This can be seen in Gonzalo's reflection on the question of stability when explaining his approach to collaboration tended not to change, but it would be more or less successful, depending on the people he was working with.

As to whether collaboration is assessable, that is dependent on what we see collaboration as and what we are assessing it for. The second layer of the collaborative assessment algorithm speaks to purpose and how the results of the assessment are interpreted. If one of the measures is the student's ability to collaborate then an essential consideration proposed in the collaborative assessment algorithm is asking what 'good' collaboration looks like. This is an encouragement to really think about the types of interactions, contributions and ways of working that are expected and valued. Once there is clarity on this, there are a range of potential assessment lenses available. That is not to suggest collaboration will become easily assessable. It will not. Indeed, the act of attempting to bring collaboration into focus through assessment lenses does carry the danger of distorting it further. No mechanism has been discovered during conducting this dissertation that makes the All-seeing-eye a likely archetype and there will always be ambiguity and a level of uncertainty about assessing collaboration. However, a flawed assessment may be preferable to no assessment at all, particularly if the assessment is designed with positive washback in mind, since assessment can change student behaviour (Wall, 2012) and the intent should be to change it for the better. As per Davidsons' (2004) call for more dialogue between teachers about assessment practices, the assessment of collaboration can be improved by teachers working together to explore their beliefs and practices and develop practical wisdom solutions.

For the final question, it is hoped the guide to assessing collaborative assignments will be a help in making teachers more comfortable and confident in assessing collaborative assignments, but the answer to the previous question is applicable to this issue too. Dialogue, debate, negotiation and community around collaborative assessment will be the most effective means of making practitioners more aware of an under-researched, under-taught and under-discussed phenomenon, answering Bolster's (2025) call for a practitioner-led pedagogy of collaboration.

5.2.1 Research questions and hypotheses from the pre-pandemic research plan

Although it was not in the scope of this dissertation to answer the research questions of the original conception of this dissertation (see 2.3.1) it is worth reflecting on them given the rich data set gathered. Below are reflections on each research question of the original plan, drawing on the findings discussed previously.

1. *Which models of grading collaborative assignments used in EAP are perceived by teachers and students to be the fairest?*

Solely assessing the product was critiqued for fairness and there were doubts about appropriate ways to assess collaboration itself so the marking scheme which found most favour was one where the collaborative assignment was complemented by individual assignments, which gives students the opportunity to shine and teachers the opportunity to differentiate. An important addendum is that if there were an effective way to assess collaboration, then that should be folded into the assessment scheme.

2. *How can an individual (best) be assessed in a summative collaborative assignment?*

This remains a problematic area. While there are mechanisms to look at the individual's performance in a collaboration (self-assessment, peer-assessment, student reporting, observation) all need to be treated with a degree of caution. As with the first question, individual complementary assignments likely offer the clearest route to individual assessment.

3. *Can, and should, contribution be assessed in a collaborative assignment and, if so, how and by whom?*

For the research participants in this dissertation, if collaboration is a learning outcome, it has to be assessed. There is no fully accepted means of doing that, but even a flawed measure is preferable to no measure at all. For summative assessment, participants took on the role of reluctant overlord, with teacher-led assessment preferred to self- or peer-assessment.

4. *To what extent does computer mediated communication provide the teacher with the affordance to assess individual contribution to a summative collaborative assignment?*

The computer mediation line of enquiry was not directly pursued during this dissertation. However, what discussion there was indicate computer mediation was not seen as a complete solution. While the ability to track student contributions online added a potential assessment lens there was awareness that what is entered into a document does not fully capture a collaboration and the various ways in which students can contribute to group work that are not captured in a key stroke. So too could the tracking of student online actions quickly become onerous, while still giving little insight, and change the role of the teacher.

At the outset of the dissertation, certain hypotheses were proposed, and they are discussed below.

- *Attitude to collaborative assignments will depend on the type of assignment. Teachers will perceive oral collaborative assignments to be fairer than written collaborative assignments due to the greater visibility of the individual in an oral assignment.*

The oral/written distinction was not a direct line of inquiry in this dissertation but participants did seem more comfortable and confident in assessing oral collaboration as there was a clearer means of differentiating between students, albeit only superficial differentiation based on factors like presentations skills, fluency and pronunciation. As Fernand noted “definitely as far as a presentation goes, I think it's easier to rate and rate more accurately” (Interview Transcripts, p. 132). There was awareness, however, that assessment of individual presentation skills did not provide any deeper insight into the collaboration itself.

- *Different models of grading collaborative assignments will be perceived differently by teachers.*

This was fully the case, and to a greater extent than anticipated. There were firmly held rationales for each of the models of assessment and justifications could be made for a range of assessment models. This underscores the complexity of introducing a collaborative assignment to a group of teachers as they may be coming to it with very different knowledge, preferences, and expectations.

- *When comparing assessment models, including individual tasks and clear grading of contribution will be favoured by teachers and students*

The student voice is not a part of this dissertation (see 5.3) and participants were not provided a model incorporating both individual complementary assignments and collaboration assessment. However, in discussion of how they would put an assessment scheme together participants generally favoured a combination of final work, individual complementary assignments and collaboration, albeit with various weightings.

- *Teachers and students will support the assessment of contribution and that this should be done by both tutor and peers*

Support was found for the assessment of collaboration itself, with the substantial caveat that it was not clear *how* to do this. While self- and peer-assessment could be employed, this would be more in the formative assessment realm than the summative assessment realm. With summative assessment, self- and peer-assessment could be employed, but filtered through teacher mediation.

- *Contribution will mainly be measured by the teacher in terms of visibility of production rather than acknowledging various ways an individual can contribute to a team effort*

There was considerable awareness that contribution in terms of word on the page was by no means the whole picture and did not capture the variety of ways in which students could contribute to a collaboration as recognised by Belbin (1981) and Williams (2017). So too was there awareness that the visibility of production would be compromised by the teacher not being privy to all the work to the students outside of that in the classroom or shared online workspaces.

- *Peer-assessment (both inter- and intra-group) will be perceived as more problematic the higher the stakes of the assignment, particularly for credit bearing or gate-keeping summative collaborative assignments.*

This was very much the case. In summative assessment in EAP, practitioners took the role of reluctant overlord, giving primacy to teacher-led assessment despite their belief in the value of peer-assessment for formative purposes.

- *Collaboration itself is difficult for the teacher to assess and although CMC increases the possibility of teachers assessing contribution to collaborative tasks it will not be felt to be adequate (some interactions and collaborative processes will remain invisible to the teacher) or the workload will be onerous.*

This was also clearly borne out by the data collected for this dissertation. Although numerous assessment lenses are available to bring collaboration into focus it can never be fully captured and attempting to do so may lead to performative collaboration rather than genuine collaboration. So too do the gains that might be made using each lens be outweighed by the workload and effort involved, for both students and teachers.

5.3 Limitations and future directions

In this section, limitations and future directions of research will be considered together. This dissertation took a GT approach, exploring an under-researched area regarding collaborative assessment and EAP practitioners. As purely qualitative research, it has not reached for statistical significance, although there are areas that may benefit from more quantitative study. For instance, from the original research plan of this dissertation, statistical analysis of student grades across a large cohort to determine the impact of middling out of grades in collaborative assignments would help clarify how much of a concern the impact of collaborative assignments on grades should be, particularly for students at borderline thresholds e.g. who could fail the course or get a distinction.

The initial limitation is that this was conducted as a global study, drawing on the experience of practitioners from a variety of backgrounds and teaching institutions. This width was valuable in developing a generalizable conception of EAP practitioner collaborative assessment identity, but also serves as a limitation as there is no opportunity to evaluate the role of institutional culture. Although some participants had experience in the same institutions, and took different understandings away from it, this research project is not robust enough in that area to speak to the impact of institutional

culture, which could have a considerable impact on collaborative assessment identity and is a necessary area of future research.

This research is also limited in the sense that it took a snapshot of practitioner thoughts at a particular time. Wendy's shift further away from collaborative assessment post interview for this dissertation suggests attitudes are not fixed. Indeed, one of the claims of this dissertation is that practitioner attitudes shifts through experience and, as such, a longitudinal study of practitioners through multiple iterations of collaborative assessment would be of value. It would also be worth investigating the impact of teacher training interventions and how collaborative assessment identity shifts as collaborative assessment literacy increases.

The evaluation of different approaches to collaborative assessment was a central part of this dissertation, but was purely theoretical. Participants were asked to consider a theoretical assignment and discuss their responses to it. There would certainly be value in getting practitioners to reflect on the experience of grading the same collaborative assignment through different assessment schemes, although the logistics of that (asking participants to run an assessment with multiple assessment lenses and evaluate the experience of using a variety of lenses) and the ethics (generating work for students that would sometimes, necessarily, not be purposeful) could be problematic.

This dissertation is focused on the voice of the EAP practitioner, under heard in the realm of collaborative assessment. But so too are the voices of university faculty staff and ELT professionals outside of academia found lacking. It would be enlightening to see how faculty staff in different institutions evaluated the four models of collaborative assessment, with their corresponding rationales. This could help indicate if the dimensions of collaborative assessment identity are generalizable or if they are specific to the EAP community. So too would it be of value to see the evaluation of ELT professionals not working in the HE sector, where assessment may not carry such high stakes and where self- and peer- assessment may find more favour.

An obvious voice missing from this dissertation is that of the student. This dissertation necessarily explored the experience of EAP practitioners as the conduit through which students

experience collaborative assignments since the approach of EAP practitioners would have an impact on the student experience. However, so too is the student voice vital. The teaching community could learn a lot from students' experience of collaborative assignments and their responses to alternative assessment practices. Indeed, the attitudes of some of the participants to collaborative assignments were product of their experiences of collaborative assignments themselves as students as much as, if not more than, any practical teacher education or training. Through incorporating the student voice, practitioner concerns may turn out to be unfounded or unconsidered factors may rise to prominence. Having noted that, the student attitude to collaborative assignments is already more prominent in the literature than that of the teacher or EAP practitioner (Zheng et al., 2021), not least in the evocatively titled "I'd rather vomit up a live hedgehog" (Strauss, 2001).

The theoretical output of this dissertation is collaborative assessment identity, where the guiding orientations of reward, practicality and defensive are bound by the principles of fairness, transparency and simplicity. Bertha's recommendation of developing a diagnostic tool to identify collaborative assessment identity would be a fruitful area of future work. It would also be of interest to try and uncover the principles and orientations of EAP practitioners for the assessment of individual written assignments. Without the complexity of collaboration, for example, it is possible the defensive orientation is not so pronounced. Further to this, it would also be of interest to see if wider teacher pedagogy beyond assessment are also impacted by the dimensions of collaborative assessment identity. Are lesson planning and course design also guided by practicality, reward and defensibility?

This dissertation also touched on some concepts that were not fully explored or within its scope. An essential area to explore is the stability of collaboration as a construct. Is a good collaborator always a good collaborator or is each collaboration unique? This has serious implications for assessment. If collaboration is stable, then an assessment can be an indicator of future performance. If collaboration is not stable, assessment becomes a snapshot of a particular instance but cannot speak to future performance. Collaborative tendencies, what they are, how they develop, would then become an important avenue of future research. We need a better understanding of what makes someone a good

collaborator, or collaboration ready, and how to nurture those tendencies. So too did the dissertation advocate for considering the positive washback of assessment lenses and assessment criteria. An example of criteria that may engender positive washback was suggested (see 4.3.4) but considerable work could be done in the area of assessment and assessment criteria design to most positively impact student behaviour in collaborative assignments.

Further to potential research, there is practical work to be done on the basis of this dissertation. The starter guide needs to be expanded and refined. As Jerry suggested, it could also be envisaged in different formats, as a lecture, a series of workshops, an online toolkit. The practical dissemination of the findings of this dissertation to the EAP community is essential to give it value.

5.4 Implications and significance

The evaluation of this dissertation (see 4.6) indicates this is a meaningful piece of work. It adds to our knowledge of EAP practitioners and their attitudes to different assessment schemes and assessment lenses that could be employed in assessing a collaborative assignment. This in turn illuminates collaborative assessment identity and collaborative assessment archetypes, building on extant theory of teacher assessment identity. So too did it build on the recommendations of the GLAID framework with the development of the collaborative assessment algorithm. The user guide which came from this is also demonstrably valued by EAP practitioners and, after further refinement, should go some way to addressing a gap in their assessment literacy.

The implication is that through the work of this dissertation, assessment practices of collaborative assignments in EAP can be improved. There can be better understanding of the concerns of EAP practitioners around collaborative assignments, which provides opportunity for professional development. That is not to say assessment literacy should be developed in a top down way, with teachers being told how to assess. Rather, that assessment literacy for collaborative assignments will improve if, and when, practitioners engage in more discussion about them, as suggested in Bolster's practitioner-led pedagogy of collaboration (2025). The utilisation of reflection as an assessment lens,

for example, is not a decision that can be imposed on a teaching centre and gain immediate acceptance. But if practitioners are well informed and are able to discuss how various assessment lenses could be best utilised, if utilised, a more broadly acceptable assessment scheme can be developed. As Manning recommended at the end of his thesis, “in order to improve practitioner confidence and enjoyment with principled EAP testing and assessment, it is worthwhile considering more widespread promotion of access to training or CPD mechanisms related to Assessment in EAP” (2013, p. 387). While this may be true of assessment generally, it is certainly the case of collaborative assessment. Collaborative assessment is not a problem that can be solved alone. Developing collaborative assessment literacy will have an added benefit as doing so necessarily develops broader assessment literacy. It will also impact course design decisions as assessment is not an add-on but fundamentally shapes an assignment.

Another implication is not directly connected to collaborative assessment, but stems from the rich data set gathered through interviewing EAP practitioners. They are a well-informed, thoughtful cohort. They raised concerns voiced in the literature without necessarily being familiar with the literature. They could articulate the challenges in assessing collaborative assignments and work towards solutions to those challenges. This suggests that if research is being conducted into an aspect of pedagogy in HE contexts, the EAP practitioner would be a good starting point.

Following this is also the question of collaborative assignments in HE more widely. Given EAP professionals dealing with class sizes typically organised to facilitate communication, how does a faculty member facing a lecture hall cohort of over a hundred students deal with collaborative assignments? The practicality and workload concerns of EAP practitioners dealing with a class of 24 students will be compounded by a significant factor by faculty lecturers with substantially larger groups. EAP has a role in preparing students for their other studies so understanding the kind of collaborations they will be facing is vital in light of discipline deficit (Bolster, 2025).

5.5 Closing thoughts

Collaborative assessment sits somewhat uneasily in HE assessment culture, which prioritises the assessment of the individual. Nonetheless, collaborative assignments are a feature of EAP and HE programmes, meaning their assessment matters. Assessment can have a significant impact on the student experience and can be integral to learning. If following Backward Design, that is absolutely the case as assessment cannot be divorced from course design and is integral to the design of tasks and materials. That does not mean assessment dominates learning, but it can shape it, and shape it for the better since assessment is, or should be, in service of helping the students attain the desired learning outcomes. Even in AoL contexts, assessment should be designed with Aal in mind, looking for opportunities to enrich the learning affordances of a course.

There are no simple plug-and-play solutions to assessing collaborative assignments but the outputs of this dissertation provide a pathway towards bespoke collaborative assessment solutions that will meet the criteria to satisfy EAP practitioner collaborative assessment identities. The assessment of collaborative assignments is a supraEAP issue and the findings of this paper can provide a pathway not only for EAP practitioners but also for HE lecturers and practitioners. As Anna noted in a follow-up email to her start guide review, “it strikes me as something that many teachers in higher ed (and not just EAP folks) would find very useful” (Anna, personal communication, July 2nd, 2024).

While there may not be a simple answer to assessing collaborative assignments, there are good, necessary, questions to ask and through discussion of those questions we can develop better, fairer, collaborative assignment assessment schemes. Through more discussion and engagement with the collaborative assessment algorithm and the starter guide, teachers will become more familiar with and, hopefully, more comfortable with assessing collaborative assignments, raising each other’s collaborative assessment literacy in the process. This, in turn, should lower any scepticism and engage them more positively in the process of teaching collaborative assignments, leading to better student experiences and learning outcomes. It is felicitous that collaboration between teachers is the route to determining how to best assess student collaborations.

If assessment is considered as a part of the learning experience, as it is under AaL, collaborative assignments offer a rich array of assessment lenses, which could help in the development of varied learning gains. Positively preparing students for, and supporting them through, collaborative assignments can leverage the benefits of collaborative learning, particularly if the assessment scheme is designed with this in mind. Fully fair assessment of collaborative assignments may be unobtainable but we can develop principled, informed assessment approaches for the benefit of student experience and learning. We can be Best-we-can-doers. And that is enough.



REFERENCES

REFERENCES

- Adachi, C., Hong-Meng Tai, J. & Dawson, P. (2018). Academics' perceptions of the benefits and challenges of self and peer assessment in higher education. *Assessment & Evaluation in Higher Education*, 43(2), 294-306. <https://doi.org/10.1080/02602938.2017.1339775>
- Alexander, O., Argent, S., & Spencer, J. (2008). *EAP essentials: A teacher's guide to principles and practice*. Garnet.
- Alexander, O., Bell, D., Cardew, S., King, J., Pallant, A., Scott, M., Thomas, D. & Ward Goodbody, M. (2008). *Competency framework for teachers of English for Academic Purposes*. BALEAP. <https://www.baleap.org/wp-content/uploads/2016/04/teap-competency-framework.pdf>
- Alghasab, M. & Handley, Z. (2017). Capturing (non-)collaboration in wiki-mediated collaborative writing activities: the need to examine discussion posts and editing acts in tandem. *Computer Assisted Language Learning*, 30(7), 664-691. <https://doi.org/10.1080/09588221.2017.1341928>
- Almond, R. (2009). Group assessment: comparing group and individual undergraduate module marks. *Assessment & Evaluation in Higher Education*, 34(2), 141-148. <https://doi.org/10.1080/02602930801956083>
- Angelova, M., & Zhao, Y. (2016). Using an online collaborative project between American and Chinese students to develop ESL teaching skills, cross-cultural awareness and language skills. *Computer Assisted Language Learning*, 29(1), 167-185. <https://doi.org/10.1080/09588221.2014.907320>.
- Anton, M. (2012). Dynamic assessment. In G. Fulcher & F. Davidson (Eds.), *The Routledge Handbook of Language Testing* (pp. 106-119). Routledge.
- Apramian, T., Cristancho, S., Watling, C., & Lingard, L. (2017). (Re) Grounding grounded theory: a close reading of theory in four schools. *Qualitative Research*, 17(4), 359-376. <https://doi.org/10.1177/1468794116672914>

- Ashenafi, M. M. (2017). Peer-assessment in higher education—twenty-first century practices, challenges and the way forward. *Assessment & Evaluation in Higher Education*, 42(2), 226-251. <https://doi.org/10.1080/02602938.2015.1100711>
- Bakhurst, D. (2009). Reflection on activity theory. *Educational Review*, 61(2), 197-210. <https://doi.org/10.1080/00131910902846916>
- Bachman, L., & Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests (Vol. 1)*. Oxford University Press.
- Battelle for kids. (2019). *Framework for 21st Century Learning*. Battelle for Kids. https://static.battelleforkids.org/documents/p21/p21_framework_brief.pdf
- Beard, J. D., Rymer, J., & Williams, D. L. (1989). An assessment system for collaborative-writing groups: Theory and empirical evaluation. *Journal of Business and Technical Communication*, 3(2), 29-51. <https://doi.org/10.1177/105065198900300203>
- Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education*, 20(2), 107-128. <https://doi.org/10.1016/j.tate.2003.07.001>
- Belbin, M. (1981). *Management teams: Why they succeed or fail*. Butterworth-Heinemann.
- Berry, E. (2007). Groupwork and assessment - benefit or burden? *The Law Teacher*, 41(1), 19-36. <https://doi.org/10.1080/03069400.2007.9959723>
- Birks, M., & Mills, J. (2023). *Grounded Theory: A practical guide (3rd edition)*. Sage.
- Bloxham, S. & Boyd, P. (2007). *Developing assessment in Higher Education: A practical guide*. McGraw-Hill Education
- Bolster, A. (2025). *A qualitative Grounded Theory study of EAP practitioners and collaborative learning: Identity and beliefs*. [Unpublished doctoral dissertation]. The University of the Basque Country
- Bolster, A. & Levrai, P. (2022). Are we talking about the same thing? Researcher and practitioner perspectives of student collaboration. In M. Evans, B. Bond & A. Ding (Eds.), *Innovation*,

Exploration and Transformation – Proceedings of the 2019 BALEAP Conference (pp. 96-104).

Garnet Education.

Borkan, J.M. (1999). Immersion/Crystallization. In B.F. Crabtree & W.L. Miller (Eds.), *Doing Qualitative Research* (pp. 179-194). SAGE Publications.

Boroditsky, L. (2017, April). How language shapes the way we think [Video]. Ted Conferences.
https://www.ted.com/talks/lera_boroditsky_how_language_shapes_the_way_we_think

Boud, D., & Bearman, M. (2022). The assessment challenge of social and collaborative learning in higher education. *Educational Philosophy and Theory*, 56(5), 459–468.
<https://doi.org/10.1080/00131857.2022.2114346>

Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>

Bowden, J., Hart, G., King, B., Trigwell, K., & Watts, O. (2000). *Generic capabilities of ATN University graduates*. Teaching and Learning Committee, Australian Technology Network.

Boychuk Duchscher, J. E., & Morgan, D. (2004). Grounded theory: reflections on the emergence vs. forcing debate. *Journal of Advanced Nursing*, 48(6), 605-612. <https://doi.org/10.1111/j.1365-2648.2004.03249.x>

Brannick, T., & Coghlan, D. (2007). In defense of being “native”: The case for insider academic research. *Organizational Research Methods*, 10(1), 59-74. <https://doi.org/10.1177/1094428106289253>

Breckenridge, J., & Jones, D. (2009). Demystifying theoretical sampling in grounded theory research. *Grounded Theory Review*, 8(2). <https://groundedtheoryreview.com/2009/06/30/847/>

Breen, L. (2007). The researcher 'in the middle': Negotiating the insider/outsider dichotomy. *The Australian Community Psychologist*, 19(1), 163-174. [https://psychology.org.au/aps/media/acp/breen_19\(1\).pdf](https://psychology.org.au/aps/media/acp/breen_19(1).pdf)

- Brown, G. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters*, 3, 45-70.
<https://search.informit.org/doi/abs/10.3316/informit.515755939850862>
- Brown, G., & Edmunds, S. (2011). *Doing pedagogical research in engineering*. Engineering Centre for Excellence in Teaching and Learning. CETL.
- Brown, S. (2005). Assessment for Learning. *Learning and Teaching in Higher Education*, 1, 81-89.
<https://eprints.glos.ac.uk/id/eprint/3607>
- Bryant, A. (2002). Re-grounding grounded theory. *Journal of Information Technology Theory and Application (JITTA)*, 4(1), 25-42. <https://core.ac.uk/download/pdf/301356818.pdf>
- Bryant, A., & Charmaz, K. (Eds.). (2007). *The Sage Handbook of Grounded Theory*. Sage.
- Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research*. 1(1), 8-22. <https://doi.org/10.1177/2345678906290531>
- Burrows, C. (2012). Washback in classroom-based assessment: A study of the washback effect in the Australian adult migrant English program. In L. Cheng & Y. Watanabe (Eds.), *Washback in Language Testing: Research contexts and methods* (pp. 113-128). Lawrence Erlbaum Associates.
- Chaiklin, S. (2003). The zone of proximal development in Vygotsky's analysis of learning and instruction. In A. Kozulin, B. Gindis, V. Ageyev & S. Miller (Eds.), *Vygotsky's educational theory in cultural context* (pp. 39-64). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511840975.004>
- Chan, C. K. Y. (2023). *Assessment for experiential learning*. Taylor & Francis.
<https://doi.org/10.4324/9781003018391>
- Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. SAGE.
- Charmaz, K. (2008). Grounded Theory as an Emergent Method. In S. N. Hesse-Biber & P. Leavy (Eds.), *Handbook of Emergent Methods* (pp. 155-172). The Guildford Press.

Charmaz, K. (2014). *Constructing Grounded Theory* (2nd Edition). Sage.

Charmaz, K., & Belgrave, L. (2012). Qualitative interviewing and grounded theory analysis. In J. Gubrium, J. Holstein, A. Marvasti & K. McKinney (Eds.), *The SAGE handbook of interview research: The complexity of the craft* (pp. 347-365). SAGE Publications Ltd. <https://doi.org/10.4135/9781452218403>

Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal*, 12(2), 68-82. <https://doi.org/10.26522/BROCKED.V12I2.38>

Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Wantanabe & A. Curtis (Eds), *Washback in language testing* 3-18. Lawrence Erlbaum Associates.

Chun Tie, Y., Birks, M., & Francis, K. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Medicine*, 7, 1-8. <https://doi.org/10.1177/2050312118822927>

Clarke, A. (2005) *Situational Analysis: Grounded Theory after the Postmodern Turn*. SAGE

Clarke, A. (2007). Grounded theory: critiques, debates, and situational analysis. In W. Outhwaite & S. Turner (Eds.), *The SAGE handbook of social science methodology* (pp. 423-442). SAGE Publications Ltd. <https://www.doi.org/10.4135/9781848607958>

Coffin, P. (2020). Implementing collaborative writing in EFL classrooms: Teachers and students' perspectives. *LEARN Journal: Language Education and Acquisition Research Network*, 13(1), 178-194.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education* (7th Edition). Routledge. <https://doi.org/10.4324/9780203720967>

Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research: Techniques and procedures for developing grounded theory: 3rd Edition*. Sage publications Inc. <https://doi.org/10.4135/9781452230153>

- Creswell, J. W., & Poth, C. N. (2017). *Qualitative inquiry and research design: Choosing among five approaches (Fourth Edition)*. Sage publications.
- Cullen, F., Bradford, S., & Green, L. (2012). Working as a practitioner-researcher. In S. Bradford & F. Cullen (Eds.) *Research and research methods for youth practitioners* (pp. 5-24). Taylor & Francis Group.
- Cushing-Leubner, J., & Bigelow, M. (2014). Principled eclecticism and the holistic approach to language teaching and learning. In S. Çelik (Ed.), *Approaches and principles in English as a foreign language (EFL) education* (pp. 254-263). Egiten.
- Davidson, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 307-334. <https://doi.org/10.1191/0265532204lt286oa>
- Davies, W. M. (2009). Groupwork as a form of assessment: Common problems and recommended solutions. *Higher Education*, 58, 563-584. <http://dx.doi.org/10.1007/s10734-009-9216-y>
- Day, C., Kington, A., Stobart, G., & Sammons, P. (2006). The personal and professional selves of teachers: Stable and unstable identities. *British Educational Research Journal*, 32(4), 601-616. <https://doi.org/10.1080/01411920600775316>
- de Chazal, E. (2014). *English for Academic Purposes - Oxford Handbook for Language Teachers*. Oxford University Press.
- de Hei, M. S., Sjoer, E., Admiraal, W., & Strijbos, J. W. (2016). Teacher educators' design and implementation of group learning activities. *Educational Studies*, 42(4), 394-409.
- de Hei, M., Strijbos, J. W., Sjoer, E., & Admiraal, W. (2016). Thematic review of approaches to design group learning activities in higher education: The development of a comprehensive framework. *Educational Research Review*, 18, 33-45.
- DeLuca, C. (2012). Preparing teachers for the age of accountability: Toward a framework for assessment education. *Action in Teacher Education*, 34(5-6), 576-591.
- Dey, I. (1999). *Grounding grounded theory: guidelines for qualitative enquiry*. Emerald.

- Dictionary.com. (2024). Principle. In *Dictionary.com dictionary*. Retrieved September 14, 2024.
<https://www.dictionary.com/browse/principle>
- Ding, A. (2013, April 25). Assessment in EAP: a continuously improved means to carelessly examined ends. *Teaching EAP*. <https://teachingeap.wordpress.com/2013/04/25/assessment-in-eap-a-continuously-improved-means-to-carelessly-examined-ends/>
- Ding, A. & Bruce, I. (2017). *The English for Academic Purposes Practitioner: operating on the edge of academic*. Palgrave Macmillan.
- Ding, A. & Monbec, L. (2024). A socio-analysis of English for Academic Purposes. In A. Ding & L. Monbec (Eds.), *Practitioner agency and identity in English for Academic Purposes* (pp. 11-45). Bloomsbury.
- Domeng, T. (2021 May 6) *What are archetypes?*. Medium. <https://sixtosilab.medium.com/what-are-archetypes-97a9838fbe6b>
- Donato, R. (1994) Collective scaffolding in second language learning. In J. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp.33-56). Ablex.
- Dörnyei, Z. & Malderez, A. (1997) Group dynamics and foreign language teaching. *System*, 25(1), 65-81. [https://doi.org/10.1016/S0346-251X\(96\)00061-9](https://doi.org/10.1016/S0346-251X(96)00061-9)
- Duff, P. (2010). Language Socialisation into Academic Discourse Communities. *Annual Review of Applied Linguistics*, 30, 169-192. <https://doi.org/10.1017/S0267190510000048>
- Dwyer, S. C., & Buckle, J. L. (2009). The space between: On being an insider-outsider in qualitative research. *International Journal of Qualitative Methods*, 8(1), 54-63. <https://doi.org/10.1177/160940690900800105>
- Earl, L. M. (2013). Assessment for learning; assessment as learning: Changing practices means changing beliefs. In Hong Kong Education Bureau (Ed.), *Assessment and learning Issue 2* (pp. 1–5). Assessment and Support Team.

- Eckstein, G., Chariton, J., & McCollum, R. M. (2011). Multi-draft composing: An iterative model for academic argument writing. *Journal of English for Academic Purposes*, 10(3), 162-172. <https://doi.org/10.1016/j.jeap.2011.05.004>
- Ede, L. S., & Lunsford, A. A. (1992). *Singular texts/plural authors: Perspectives on collaborative writing*. SIU Press.
- Edwards, A., & Daniels, H. (2004). Using sociocultural and activity theory in educational research. *Educational Review*, 56(2), 107-111. <https://doi.org/10.1080/0031910410001693191>
- Edstrom, A. (2015). Triads in the L2 classroom: Interaction patterns and engagement during a collaborative task. *System*, 52, 26-37.
- Elabdali, R. (2021). Are two heads really better than one? A meta-analysis of the L2 learning benefits of collaborative writing. *Journal of Second Language Writing*, 52, 1-16. <https://doi.org/10.1016/j.jslw.2020.100788>
- Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research*, 4(3), 193-220. <https://doi.org/10.1177/136216880000400302>
- Elola, I., Oskoz, A. (2017) Writing with 21st century social tools in the L2 classroom: New literacies, genres, and writing practices. *Journal of Second Language Writing*, 36, 52-60. <https://doi.org/10.1016/j.jslw.2017.04.002>
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215-251. <https://doi.org/10.1037/0033-295X.87.3.215>
- Fani, T., & Ghaemi, F. (2011). Implications of Vygotsky's zone of proximal development (ZPD) in teacher education: ZPTD and self-scaffolding. *Procedia-Social and Behavioral Sciences*, 29, 1549-1554. <https://core.ac.uk/download/pdf/82645123.pdf>
- Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, 21(1), 40-58. <https://doi.org/10.1016/j.jslw.2011.12.002>

- Fernández Dobao, A. & Blum, A. (2013). Collaborative writing in pairs and small groups: Learners' attitudes and perceptions. *System*, 41(2), 365-378.
<https://doi.org/10.1016/j.system.2013.02.002>
- Finnish National Board on Research TENK. (2019). *The ethical principles of research with human participants and ethical review in the human sciences in Finland*.
[https://tenk.fi/sites/default/files/2021-01/Ethical review in human sciences 2020.pdf](https://tenk.fi/sites/default/files/2021-01/Ethical%20review%20in%20human%20sciences%202020.pdf)
- Flowerdew, J. & Peacock, M. (2001) *Research Perspectives on English for academic purposes*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524766>
- Frykedal, K., Chiriac, E. (2011). Assessment of students' learning when working in groups. *Educational Research*, 53(3), 331-345. <https://doi.org/10.1080/00131881.2011.598661>
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. Routledge.
- Gammie, E., and Matson, M. (2007). Group assessment at final degree level: An evaluation. *Accounting Education: An International Journal*, 16(2), 185–206.
<https://doi.org/10.1080/09639280701234609>
- Gaur, S. B., & Gupta, T. (2013). Is group assessment a bane or boon in higher education? A students–teacher perspective. *International Journal of Fashion Design, Technology and Education*, 6(3), 141-148. <https://doi.org/10.1080/17543266.2013.795610>
- Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education: Choosing and using diverse approaches* (pp. 41-53). McGraw-Hill Education.
- Gillett, A. (2017). *What is EAP?* Retrieved from <http://www.uefap.com/bgnd/whatfram.htm>.
- Gipps, C., & Stobart, G. (2009). Fairness in assessment. In C. Wyatt-Smith & J. J. Cumming (Eds.), *Educational assessment in the 21st century: Connecting theory and practice* (pp. 105-118). Springer. https://doi.org/10.1007/978-1-4020-9964-9_6
- Glaser, B. (1978). *Theoretical Sensitivity*. The Sociology Press.
- Glaser, B. (1998). *Doing Grounded Theory: Issues and Discussions*. The Sociology Press.

- Glaser, B. G., & Holton, J. (2007). Remodeling Grounded Theory. *Historical Social Research / Historische Sozialforschung. Supplement*, 19, 47–68. <http://www.jstor.org/stable/40981068>
- Glaser, B. & Strauss, A. (1967). *The Discovery of Grounded Theory; Strategies for Qualitative Research*. Routledge. <https://doi.org/10.4324/9780203793206>
- Green, J., Willis, K., Hughes, E., Small, R., Welch, N., Gibbs, L., & Daly, J. (2007). Generating best evidence from qualitative research: the role of data analysis. *Australian and New Zealand Journal of Public Health*, 31(6), 545-550. <https://doi.org/10.1111/j.1753-6405.2007.00141.x>
- Hadley, G. (2017). *Grounded Theory in Applied Linguistics Research: A practical guide*. Routledge.
- Hamp-Lyons, L. (2009). Principles for Large-Scale Classroom-Based Teacher Assessment of English Learners' Language: An Initial Framework From School-Based Assessment in Hong Kong. *TESOL Quarterly*. 43(3), 524-430. <https://www.jstor.org/stable/27785034>
- Hamp-Lyons, L. (2011). English for Academic Purposes. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning - Volume II* (pp. 89-105). Routledge.
- Hastings, M. (2019). Neoliberalism and Education. *Oxford Research Encyclopedias*. <https://oxfordre.com/education/display/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-404>
- Hathorn, L. G., & Ingram, A. L. (2002). Online collaboration: Making it work. *Educational Technology*, 42(1), 33-40. <https://www.jstor.org/stable/44428720>
- Henri, F. & Rigault, C. (1996). Collaborative distance education and computer conferencing. In T. T. Liao (Ed.), *Advanced educational technology: Research issues and future potential* (pp. 45-76). Springer-Verlag.
- Herman, J., & Cook, L. (2019). Fairness in classroom assessment. In S. Brookhart & J. McMillan (Eds.), *Classroom assessment and educational measurement* (pp. 243-264). Routledge.
- Hey, J. (2019, July 7). *The Cobra Effect*. Sketchplanations.com. <https://sketchplanations.com/the-cobra-effect>

- Hill, K., & McNamara, T. (2011). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395-420. <https://doi.org/10.1177/0265532211428317>
- Holton, J. A. (2007). The coding process and its challenges. In A. Bryant & K. Charmaz (Eds.), *The Sage handbook of grounded theory* (pp. 265-289). Sage.
- Hood, J. C. (2007). Orthodoxy vs. power: The defining traits of grounded theory. In A. Bryant & K. Charmaz (Eds.), *The Sage handbook of grounded theory* (pp. 151-164). Sage.
- Humphrey, C. (2013). Dilemmas in doing insider research in professional education. *Qualitative Social Work*, 12(5), 572-586. <https://doi.org/10.1177/147332501244600>
- Hyland, K. (2006). *English for Academic Purposes: An advanced resource book*. Taylor & Francis.
- Hyland, K. & Hamp-Lyons, L. (2002) EAP. Issues and directions. *Journal of English for Academic Purposes*, 2, 1-12. [https://doi.org/10.1016/S1475-1585\(02\)00002-4](https://doi.org/10.1016/S1475-1585(02)00002-4)
- Ion, G., Díaz-Vicario, A., & Mercader, C. (2023). Making steps towards improved fairness in group work assessment: The role of students' self-and peer-assessment. *Active Learning in Higher Education*, 0(0), 1-13. <https://doi.org/10.1177/146978742311548>
- James, M., & Pedder, D. (2006). Beyond method: assessment and learning practices and values. *The Curriculum Journal*. 17(2), 109-138. <https://doi.org/10.1080/09585170600792712>
- Jaques, D. (2000). *Learning in groups: a handbook for improving group work (3rd edition)*. Routledge.
- Johnston, L. & Miles, L. (2010). Assessing contributions to group assignments. *Assessment & Evaluation in Higher Education*, 29(6), 751-768. <https://doi.org/10.1080/0260293042000227272>
- Jordan, R. R. (1997) *English for Academic Purposes: A guide and resource book for teachers*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511733062>
- Kane, M. (2010). Validity and fairness. *Language Testing*, 27(2), 177-182. <https://doi.org/10.1177/0265532209349467>

- Kearney, S., Perkins, T., Kennedy-Clark, S. (2016) Using self- and peer-assessments for summative purposes: analysing the relative validity of the AASL (Authentic Assessment for Sustainable Learning) model. *Assessment & Evaluation in Higher Education*, 41(6), 840-853. <https://doi.org/10.1080/02602938.2015.1039484>
- Khan, S. (2016). Process Writing Instruction and Assessment. *Langlit*, 2(3), 788-794.
- Kumaravadivelu, B. (1994). The postmethod condition: (E)merging strategies for second/foreign language teaching. *TESOL Quarterly*, 28(1), 27-48. <https://doi.org/10.2307/3587197>
- Kunnan, A. J. (2000). Fairness and justice for all. In A. J. Kunnan (Ed.), *Fairness and validation in language assessment: selected papers from the 19th Language Testing Research Colloquium, Orlando, Florida* (pp. 1-14). Cambridge University Press.
- Lantolf, J. P. (2000). Introducing sociocultural theory. *Sociocultural Theory and Second Language Learning*, 1, 1-26.
- Lantolf, J., P., & Frawley, W. (1985). Oral-proficiency testing. *The Modern Language Journal*. 69(4), 337-345). <https://doi.org/10.2307/328404>
- Lantolf, J., P., & Pavlenko, A. (1995). Sociocultural theory and second language acquisition. *Annual Review of Applied Linguistics*, 14, 108-124. <https://doi.org/10.1017/S0267190500002646>
- Lantolf, J. P., Poehner, M. E., & Swain, M. (Eds.). (2018). *The Routledge handbook of sociocultural theory and second language development*. Routledge.
- Lantolf, J. P. & Thorne, S. (2006). *Sociocultural Theory and the Genesis of Second Language Development*. Oxford University Press.
- Lea, M. R., & Street, B. V. (2006). The "academic literacies" model: Theory and applications. *Theory Into Practice*, 45(4), 368-377. <https://www.jstor.org/stable/40071622>
- Leach, L. (2012). Optional self-assessment: some tensions and dilemmas. *Assessment & Evaluation in Higher Education*, 37(2), 137-147. <https://doi.org/10.1080/02602938.2010.515013>

- Lee, I. (2007). Assessment for Learning: Integrating assessment, teaching and learning in the ESL/EFL writing classroom. *Canadian Modern Language Review*, 64(1), 199-213.
<https://doi.org/10.3138/cmlr.64.1.199>
- Lee, I. (2016). Putting Students at the Centre of Classroom L2 Writing Assessment. *Canadian Modern Language Review*, 72(2), 258-280. <https://doi.org/10.3138/cmlr.2802>
- Levers, M. J. D. (2013). Philosophical paradigms, grounded theory, and perspectives on emergence. *Sage Open*, 3(4), 1-6.
<https://journals.sagepub.com/doi/pdf/10.1177/2158244013517243>
- Levin, P. (2003). Running group projects: dealing with the free-rider problem. *Planet*, 9(1), 7-8.
<https://doi.org/10.11120/plan.2003.00090007>
- Levrai, P. (2013) The coursebook as trainer. *English Teaching Professional*. 4-7.
- Levrai, P., & Bolster, A. (2017, April 7th). Essays with Benefits: Undergraduate Collaborative Writing. BALEAP Conference: Addressing the state of the union: Working together = learning together, Bristol. <https://www.baleap.org/event/addressing-state-union-working-together-learning-together>
- Levrai, P., & Bolster, A. (2019). A framework to support group essay writing in English for Academic Purposes: A case study from an English-medium instruction context. *Assessment & Evaluation in Higher Education*, 44(2), 186-202.
<https://doi.org/10.1080/02602938.2018.1487024>
- Li, M., & Zhu, W. (2013). Patterns of computer-mediated interaction in small writing groups using wikis. *Computer Assisted Language Learning*, 26(1), 61-82.
- Li, M., & Zhu, W. (2017). Explaining dynamic interactions in wiki-based collaborative writing. *Language Learning & Technology*, 21(2), 96-120.
- Li, M., & Zhang, M. (2023). Collaborative writing in L2 classrooms: A research agenda. *Language Teaching*, 56(1), 94-112. <http://dx.doi.org/10.1017/S0261444821000318>

- Littlewood, W. (2004). The task-based approach: some questions and suggestions. *ELT Journal*, 58(4), 319-326. <https://doi.org/10.1093/elt/58.4.319>
- Looney, A., Cumming, J., van Der Kleij, F. & Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity, *Assessment in Education: Principles, Policy & Practice*, 25(5), 442-467, <https://doi.org/10.1080/0969594X.2016.1268090>
- Lowe, S., Carr, A. N., Thomas, M., & Watkins-Mathys, L. (2005). The fourth hermeneutic in marketing theory. *Marketing Theory*, 5(2), 185-203. <https://doi.org/10.1177/1470593105052471>
- Lumivero. (2024a). *NVivo12 (Version 12)* [Computer software]. www.lumivero.com
- Lumivero. (2024b). *NVivo20 (Version 20)* [Computer software]. www.lumivero.com
- Lund, A. (2008). Assessment Made Visible: Individual and Collective Practices. *Mind, Culture, and Activity*, 15(1), 32-51. DOI: 10.1080/10749030701798623
- Maiden, B., & Perry, B. (2011). Dealing with free-riders in assessed group work: results from a study at a UK university. *Assessment & Evaluation in Higher Education*, 36(4), 451-464. <https://doi.org/10.1080/02602930903429302>
- Maldonado, H., Klemmer, S., & Pea, R. (2009). When is collaborating with friends a good idea? Insights from design education. In C. O'Malley, D. Suthers, P. Reimann & A. Dimitracopoulou (Eds), *Computer Supported Collaborative Learning Practices: CSCL09 Conference Proceedings* (pp. 227-231). ISLS. <https://repository.isls.org/bitstream/1/2948/1/227-231.pdf>
- Manning, A. (2013). *EAP Teacher Assessment Literacy* [Unpublished doctoral dissertation]. University of Leicester. https://figshare.le.ac.uk/articles/thesis/EAP_Teacher_Assessment_Literacy/10170863
- Mathieu, J., Tannenbaum, S., Kukenberger, M., Donsbach, J. & Alliger, G. (2015). Team Role Experience and Orientation: A Measure and Tests of Construct Validity. *Group & Organisation Management*, 40(1), 6-34. <https://doi.org/10.1177/10596011145620>
- McDonald, J. (2003) Assessing online collaborative learning: process and product. *Computers & Education*, 40(4), 377-391. [https://doi.org/10.1016/S0360-1315\(02\)00168-9](https://doi.org/10.1016/S0360-1315(02)00168-9)

McLeod, J. (1999). *Practitioner research in counselling*. Sage Publications.

Meijer, H., Hoekstra, R., Brouwer, J., & Strijbos, J. W. (2020). Unfolding collaborative learning assessment literacy: a reflection on current assessment methods in higher education. *Assessment & Evaluation in Higher Education*, 45(8), 1222-1240.

Merriam-Webster. (n.d.). Orientations. In *Merriam-Webster.com dictionary*. Retrieved September 14, 2024. <https://www.merriam-webster.com/dictionary/orientations>

Messick, S. A. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13-103). Macmillan Publishing.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. <https://doi.org/10.1177/026553229601300302>

Michaelsen, L. K., & Sweet, M. (2008). The essential elements of team-based learning. *New Directions for Teaching and Learning*, 2008(116), 7-27. <https://doi.org/10.1002/tl.330>

Mockler, N. (2011). Beyond ‘what works’: Understanding teacher identity as a practical and political tool. *Teachers and Teaching*, 17(5), 517–528. <https://doi.org/10.1080/13540602.2011.602059>

Morse, J. M. (2010). Sampling in grounded theory. In A. Bryant & K. Charmaz (Eds.), *The SAGE handbook of grounded theory* (pp. 229-244). Sage Publications Ltd. <https://doi.org/10.4135/9781848607941>

Mueller, J. (2005). The authentic assessment toolbox: enhancing student learning through online faculty development. *Journal of Online Learning and Teaching*, 1(1), 1-7.

Mulligan, C., & Garofalo, R. (2011). A collaborative writing approach: Methodology and student assessment. *The Language Teacher*, 35(3), 5-10. <https://doi.org/10.37546/JALTTLT35.3-1>

Murphy, R. (2006). Evaluating new priorities for assessment in higher education. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp. 37-47). Routledge.

- Nepal, K.P. (2012). An approach to assign individual marks from a team mark: the case of Australian grading system at universities. *Assessment & Evaluation in Higher Education*, 37(5), 555-562. <https://doi.org/10.1080/02602938.2011.555815>
- Norton, L. (2004) Using assessment criteria as learning criteria: a case study in psychology. *Assessment & Evaluation in Higher Education*, 29(6), 687-702. <https://doi.org/10.1080/0260293042000227236>
- Norton, L., Tilley, A., Newstead, S., & Franklyn-Stokes, A. (2001). The pressures of assessment in undergraduate courses and their effect on student behaviours. *Assessment & Evaluation in Higher Education*, 26(3), 269-284. <https://doi.org/10.1080/02602930120052422>
- Nouri, J., Akerfeldt, A., Fors, U., Selander, S. (2017). Assessing Collaborative Problem Solving Skills in Technology-Enhanced Learning Environments – The PISA Framework and Modes of Communication. *International Journal of Emerging Technologies in Learning*, 12(4), 163-174. <https://doi.org/10.3991/ijet.v12i04.6737>
- Nusche, D. (2008). Assessment of learning outcomes in higher education: A comparative review of selected practices. *OECD Education Working Papers No. 15*. OECD. <https://dx.doi.org/10.1787/244257272573>
- Nyikos, M., & Hashimoto, R. (1997). Constructivist theory applied to collaborative learning in teacher education: In search of ZPD. *The Modern Language Journal*, 81(4), 506-517. <https://doi.org/10.2307/328893>
- O'Neill, G., & McMahon, T. (2005). Student-centred learning: What does it mean for students and lecturers. In G. O'Neill, S. Moore. & B. McMullin (Eds.), *Emerging Issues in the Practice of University Learning and Teaching* (pp. 27-36). AISHE.
- Ohta, A. (1995). Applying Sociocultural Theory to an Analysis of Learner Discourse: Learner-Learner Collaborative Interaction in the Zone of Proximal Development. *Issues in Applied Linguistics*, 6(2), 93–121. <https://doi.org/10.5070/L462005219>

- Ohta, A. (2013). 30 - Sociocultural theory and the zone of proximal development. In J. Herschensohn and M. Young-Scholten (Eds), *The Cambridge Handbook of Second Language Acquisition* (pp. 648-669). Cambridge University Press.
- Olson, G. J., Duffy, S.A., & Mack, R. L. (1984). Thinking-out-loud as a method for studying real time comprehension processes. In D.E. Kieras & M.A. Just (Eds.), *New methods in reading comprehension research* (pp. 253-286). Erlbaum.
- Otter.ai. (2024a). Otter.ai. [Computer software.] <https://otter.ai/>
- Otter.ai. (2024b). Otter.ai. [Mobile App]. Google Play Store. <https://play.google.com/store/apps/details?id=com.aisense.otter&hl=en>
- Oxford University Press. (n.d.). *Global Skills: Creating empowered 21st century learners*. Oxford University Press. <https://elt.oup.com/feature/global/expert/global-skills?cc=fi&selLanguage=en>
- Palmour, L. (2023). Assessing oral presentations: An analysis of score-reaching dialogue between EAP practitioners. *Journal of English for Academic Purposes*, 62, 1-12. <https://doi.org/10.1016/j.jeap.2022.101210>
- Panitz, T. (1999) Collaborative versus cooperative learning: a comparison of the two concepts which will help us understand the underlying nature of interactive learning. https://crlt.umich.edu/sites/default/files/resource_files/ED448443.pdf, accessed 29/07/23.
- Panitz, T., & Panitz, P. (1998). Encouraging the use of collaborative learning in higher education. *University Teaching: International Perspectives*, 161-201.
- Paran, A. (2010). More than language: The additional faces of testing and assessment in language learning and teaching. In A. Paran & L. Sercu (Eds.), *Testing the Untestable in Language Education* (pp. 1-13). Multilingual Matters.
- Paulus, T. (2005). Collaboration or cooperation? Analyzing small group interactions in educational environments. In T. Roberts (Ed.), *Computer-Supported Collaborative Learning in Higher Education* (pp. 100-124). IGI Global.

- Pitt, M. (2000). The application of games theory to group project assessment. *Teaching in Higher Education*, 5(2), 233-241. <https://doi.org/10.1080/135625100114876>
- Plake, B. (1993). Teacher assessment literacy: teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher*, 6(1), 21-27.
- Plastow, N., Spiliotopoulou, G. and Prior, S. (2010). Group assessment at first year and final degree level: a comparative evaluation. *Innovations in Education and Teaching International*, 47(4), 393–403. <https://doi.org/10.1080/14703297.2010.518427>
- Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9(3), 233-265.
- Price, M. (2005). Assessment standards: the role of communities of practice and the scholarship of assessment. *Assessment & Evaluation in Higher Education*, 30(3), 215-230. <https://doi.org/10.1080/02602930500063793>
- Project CETI. (2023). *Work Together*. <https://www.projectceti.org/about>
- Punch, K. & Oancea, A. (2014) *Introduction to Research Methods in Education* (2nd Edition). Sage Publications.
- Queen Mary Academy. (n.d.). *Intended learning outcomes*. Queen Mary Academy. <https://www.qmul.ac.uk/queenmaryacademy/educators/resources/curriculum-design/intended-learning-outcomes/>
- Race, P. (2001) *A Briefing on Self, Peer and Group Assessment*. LTSN Generic Centre.
- Ramsden, P. (2003) *Learning to Teach in Higher Education*, 2nd Edition. Routledge Falmer.
- Rasooli, A., Turner, J., Varga-Atkins, T., Pitt, E., Asgari, S., & Moindrot, W. (2024). Students' perceptions of fairness in groupwork assessment: validity evidence for peer assessment fairness instrument. *Assessment & Evaluation in Higher Education*, 1-16. <https://doi.org/10.1080/02602938.2024.2354510>
- Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21(3), 249-258. <https://doi.org/10.1191/0265532204lt283e>

- Rice, J. A. (2007). Bridging the gap: Contextualizing professional ethics in collaborative writing projects. *Business Communication Quarterly*, 70(4), 470–475. <https://doi.org/10.1177/10805699070700040306>.
- Richards, J. C. (2013). Curriculum approaches in language teaching: Forward, central, and backward design. *RELC Journal*, 44(1), 5-33. <https://doi.org/10.1177/0033688212473293>
- Rieber, R. & Carton, A. (Eds.) (1987). *The Collected Works of LS Vygotsky. Vol. 1: Problems of General Psychology, including the Volume Thinking and Speech*. Springer.
- Roschelle, J. & Teasley, S. (1995). The construction of shared knowledge in collaborative problem-solving. In C. E. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69-97). Springer-Verlag.
- Roulston, K. (2010). Theorising the researcher: The reflective interviewer. In K. Roulston (Ed.), *Reflective interviewing: A guide to theory and practice* (pp. 115-129). SAGE Publications. <https://doi.org/10.4135/9781446288009>
- Sadler, R. (1989). Formative assessment: Revising the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), p.77-84. <https://doi.org/10.1080/0969595980050104>
- Saldaña J. (2011). *Fundamentals of qualitative research*. Oxford University Press.
- Saldaña J. (2016). *The Coding Manual for Qualitative Researchers* (3rd Edition). Sage Publications Ltd.
- Saldaña J. (2021). *The Coding Manual for Qualitative Researchers* (4th Edition). Sage Publications Ltd.
- Salter-Dvorak, H. (2016). Learning to argue in EAP: Evaluating a curriculum innovation from the inside. *Journal of English for Academic Purposes*, 22, 19-31. <https://doi.org/10.1016/j.jeap.2015.12.005>
- Savignon, S. (1991). Communicative Language Teaching: State of the Art. *TESOL Quarterly*. 25(2), 261-277. <https://doi.org/10.2307/3587463>
- Schellekens, L, Bok, H., de Jong, L., van der Schaaf, M., Kremer, W. & van der Vlueten, C. (2021). A scoping review of the notions of Assessment as Learning (AaL), Assessment for Learning

- (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation*, 71, p. 1-15.
<https://doi.org/10.1016/j.stueduc.2021.101094>
- Schmitt, D., & Hamp-Lyons, L. (2015). The need for EAP teacher knowledge in assessment. *Journal of English for Academic Purposes*, 18, 3-8. <https://doi.org/10.1016/j.jeap.2015.04.003>
- Scoles, S. (2023, February, 9th). New Exascale Supercomputer Can Do a Quintillion Calculations a Second. *Scientific American*. <https://www.scientificamerican.com/article/new-exascale-supercomputer-can-do-a-quintillion-calculations-a-second/>
- Scriven, M. (1967). The methodology of evaluation. In R.W. Tyler (Ed.), *Perspectives of curriculum evaluation* (pp. 39–83). Rand McNally.
- Searle, M., & Poth, C. (2021). Collaborative evaluation designs as an authentic course assessment. *Canadian Journal of Program Evaluation*, 35(3), 422-436.
<https://doi.org/10.3138/cjpe.69691>
- Shapiro, D. (2003). Facilitating holistic curriculum development. *Assessment & Evaluation in Higher Education*, 28(4), 423-434. <https://doi.org/10.1080/0260293032000066245>
- Shaw, I. (2005). Practitioner research: evidence or critique? *The British Journal of Social Work*, 35(8), 1231-1248. <https://www.jstor.org/stable/23720554>
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20(4), 286-305. <https://doi.org/10.1016/j.jslw.2011.05.010>
- Şişman, E. P., & Büyükkarçi, K. (2019). A review of foreign language teachers' assessment literacy. *Sakarya University Journal of Education*, 9(3), 628-650.
- Smagorinsky, P. (2018). Deconflating the ZPD and instructional scaffolding: Retranslating and reconceiving the zone of proximal development as the zone of next development. *Learning, culture and social interaction*, 16, 70-75. <https://doi.org/10.1016/j.lcsi.2017.10.009>
- Smith, B. L., and MacGregor, J. T. (1992). What is collaborative learning? In S. Goodsell (Ed.), *Collaborative Learning: A Sourcebook for Higher Education* (pp. 11-30). National Center on Postsecondary Teacher, Learning and Assessment.

- Spooner, M. (2017). 'Qualitative Research and Global Audit Culture'. In N. Denzin & Y. Lincon (Eds.), *The SAGE handbook of qualitative research 5th edition* (pp. 894-914). SAGE Publications Inc.
- Stevick, E. W. (1980). *Teaching languages: A way and ways*. Newbury House.
- Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238–245.
- Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52, 119-158.
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14(3), 153-173. <https://doi.org/10.1016/j.jslw.2005.05.002>
- Storch, N. (2012). Collaborative language learning. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-6). Wiley. <https://doi.org/10.1002/9781405198431.wbeal0153>
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Multilingual Matters. <https://doi.org/10.21832/9781847699954>
- Storch, N. (2017). Implementing and assessing collaborative writing activities in EAP classes. In J. Bitchener, N. Storch & R. Wette (Eds.), *Teaching Writing for Academic Purposes to Multilingual Students* (pp. 130-142). Routledge. <http://dx.doi.org/10.4324/9781315269665-9>
- Storch, N. (2019). Collaborative writing. *Language Teaching*, 52(1), 40-59. <https://doi:10.1017/S0261444818000320>
- Storch, N. (2021). Collaborative writing: Promoting languaging among language learners. In M. P. García Mayo (Ed.), *Working Collaboratively in Second/Foreign Language Learning* (pp. 13–34). De Gruyter Mouton.
- Storch, N. & Knoch, U. (2023). Assessing collaborative writing: current challenges and future possibilities. In M. Li & M. Zhang (Eds.), *L2 Collaborative Writing in Diverse Learning Contexts* (pp. 229-247). John Benjamins.
- Storch, N. & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M. P. Garcia Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157-177). De Gruyter.

- Strauss, A. (1987) *Qualitative Analysis for Social Scientists*. Cambridge University Press.
- Strauss, P. (2001). 'I'd rather vomit up a live hedgehog' - L2 students and group assessment in mainstream university programs. *Prospect*, 16 (2): 55-66.
- Strauss, P., & U, A. (2007). Group assessments: dilemmas facing lecturers in multicultural tertiary classrooms. *High Education Research & Development*, 26(2), 147-161.
<https://doi.org/10.1080/07294360701310789>
- Srijbos, J. (2016). Assessment of Collaborative Learning. In G. T. L. Brown, & L. R. Harris (Eds.), *Handbook of Social and Human Conditions in Assessment* (pp. 302-318). Routledge.
- Swan, K., Shen, J. & Hiltz, S. (2006). Assessment and Collaboration in Online Learning. *Journal of Asynchronous Learning*, 10(1), 45-62.
- Swann, J. and Ecclestone, K. (1999a) Litigation and learning: tensions in improving university lecturers' assessment practice. *Assessment in Education: Principles, Policy & Practice*, 6(3), 377–389.
<https://doi.org/10.1080/09695949992801>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education*, 76, 467-481.
- Talib, T., & Cheung, Y. L. (2017). Collaborative Writing in Classroom Instruction: A Synthesis of Recent Research. *The English Teacher*, 46(2), 43-57.
- Thomas, T. A. (2014). Developing team skills through a collaborative writing assignment. *Assessment & Evaluation in Higher Education*, 39(4), 479-495.
<https://doi.org/10.1080/02602938.2013.850587>
- Thurlow, L. (2020). Grounded theory and the PhD—notes for novice researchers. *Journal of Humanities and Applied Social Sciences*, 2(4), 257-270.
- Tierney, R. D. (2014). Fairness as a multifaceted quality in classroom assessment. *Studies in Educational Evaluation*, 43, 55-69. <https://doi.org/10.1016/j.stueduc.2013.12.003>

- Timonen, V., Foley, G. & Conlon, C. (2018). Challenges When Using Grounded Theory: A Pragmatic Introduction to Doing GT Research. *International Journal of Research Methods*. 17, 1-10. <https://doi.org/10.1177/1609406918758086>
- Tomlinson, B. (2016). The Importance of Materials Development for Language Learning. In M. Azarnoosh, M. Zeraatpishe, A. Faravani, & H. Kargozari (Eds.), *Issues in Materials Development* (pp. 1-9). Brill
- Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 55-87). Kluwer Academic Publishers.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14(3), 281–294. <https://doi.org/10.1080/09695940701591867>
- Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education*, 38 (3), 323–342. <https://doi.org/10.1080/03054985.2012.689693>
- Turner, B. A. (1983). The use of grounded theory for the qualitative analysis of organizational behaviour. *Journal of Management Studies*, 20(3), 333-348. <https://doi.org/10.1111/j.1467-6486.1983.tb00211.x>
- Turner, C. (2012). Classroom Assessment. In C. Fulcher & F. Davidson (Eds.), *The Routledge Handbook of Language Testing* (pp. 65-78). Routledge.
- University of Macau. (2020 January 31st). *UM adopts distance education to keep students learning* [Press Release]. <https://www.um.edu.mo/news-and-press-releases/press-release/detail/49941/>

- Valentine, N., Durning, S. J., Shanahan, E. M., van der Vleuten, C., & Schuwirth, L. (2022). The pursuit of fairness in assessment: Looking beyond the objective. *Medical Teacher*, 44(4), 353-359. <https://doi.org/10.1080/0142159X.2022.2031943>
- Van Aalst, J. (2013). Assessment in collaborative learning. In C. Hmelo-Silver, C. Chinn, C. Chan & A O'Donnell (Eds.), *The International Handbook of Collaborative Learning* (pp. 280-296). Routledge.
- Van Daalen, M. (1999). Test usefulness in alternative assessment. *Dialog on Language Instruction*, 13(1-2), 1-26.
- Van Der Vleuten, C. P., Schuwirth, L. W. T., Driessen, E. W., Govaerts, M. J. B., & Heeneman, S. (2015). Twelve tips for programmatic assessment. *Medical Teacher*, 37(7), 641-646. <https://doi.org/10.3109/0142159X.2014.973388>
- Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2017). Authentic assessment: creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 43(5), 840–854. <https://doi.org/10.1080/02602938.2017.1412396>
- Villarroel, V., Bruna, D., Bruna, C., Brown, G., & Boud, D. (2024). Authentic assessment training for university teachers. *Assessment in Education: Principles, Policy & Practice*, 1–19. <https://doi.org/10.1080/0969594X.2024.2350395>
- Vygotsky, L. S. (1978). Interaction between learning and development. In Gauvainm M. & Cole, M. (Eds.), *Readings on the Development of Children*, (2nd ed., pp. 29-36). Scientific American Books.
- Wall, D. (2012). Washback. In C. Fulcher & F. Davidson (Eds.), *The Routledge Handbook of Language Testing* (179-92). Routledge.
- Walters, F. S. (2012). Fairness. In C. Fulcher & F. Davidson (Eds.), *The Routledge Handbook of Language Testing* (179-92). Routledge.
- Warford, M. K. (2011). The zone of proximal teacher development. *Teaching and Teacher Education*, 27(2), 252-258. <https://doi.org/10.1016/j.tate.2010.08.008>

- Widdowson, H. G. (1983). *Learning purpose and language use*. Oxford University Press.
- Wiggins, G.P., & McTighe, J. (2005). *Understanding by design (Expanded 2nd Edition)*. Pearson.
- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, 26(3), 445-466 <https://doi.org/10.1177/026553220910467>
- Williams, P. (2017). Assessing collaborative learning: big data, analytics and university futures. *Assessment & Evaluation in Higher Education*, 42(6), 978-989. <https://doi.org/10.1080/02602938.2016.1216084>
- Wilson, K. J., Brickman, P., & Brame, C. J. (2018). Group work. *Life Sciences Education*, 17(1), 1-5. <https://doi.org/10.1187/cbe.17-12-0258>
- Wingate, U., & Tribble, C. (2012). The best of both worlds? Towards an English for Academic Purposes/Academic Literacies writing pedagogy. *Studies in Higher Education*, 37(4), 481-495. <https://doi.org/10.1080/03075079.2010.525630>
- Wise, S. & Smith, L. (2016). The Validity of Assessment When Students Don't Give Good Effort. In G. T. L. Brown, & L. R. Harris (Eds.), *Handbook of Social and Human Conditions in Assessment* (pp. 204-220). (Educational Psychology Handbook Series). Routledge. <https://www.routledgehandbooks.com/doi/10.4324/9781315749136.ch12>
- Wong, B., Chiu, Y. L. T., Copsey-Blake, M., & Nikolopoulou, M. (2022). A mapping of graduate attributes: what can we expect from UK university students? *Higher Education Research & Development*, 41(4), 1340-1355. <https://doi.org/10.1080/07294360.2021.1882405>
- World Economic Forum. (2016). *New Vision for Education: Fostering Social and Emotional Learning through Technology*. World Economic Forum. https://www3.weforum.org/docs/WEF_New_Vision_for_Education.pdf
- Xi, X. (2010). How do we go about investigating fairness? *Language Testing*, 27(2), 147-170. <https://doi.org/10.1177/0265532209349465>

Yin, M. (2010). Understanding classroom language assessment through teacher thinking research.

Language Assessment Quarterly, 7(2), 175–194.

<https://doi.org/10.1080/15434300903447736>

Zhang, M., & Chen, W. (2022). Assessing collaborative writing in the digital age: An exploratory study.

Journal of Second Language Writing, 57, 1-17. <https://doi.org/10.1016/j.jslw.2022.100868>

Zheng, Y., Yu, S., & Lee, I. (2021). Implementing collaborative writing in Chinese EFL classrooms: Voices

from tertiary teachers. *Frontiers in Psychology*, 12, 1-9.

<https://doi.org/10.3389/fpsyg.2021.631561>

LIST OF APPENDICES

These appendices can be accessed via this secure UPV/EHU shared folder - Peter Levrai PhD Appendices

Appendix 1: Dissertation data collection tools (including ethic consent, surveys, pre-interview tasks & interview protocols)

Appendix 2: Survey data

Appendix 3: Pre-interview task data

Appendix 4: Interview transcripts

Appendix 5: Secondary Interview Data

Appendix 6: Data processing codebooks

Appendix 7: Data collection tools for original study (including ethic consent, student and teacher surveys, student and teacher interview protocols)

Appendix 8: Motivation and background studies

Appendix 9: First draft starter guide to assessing collaborative assessments

Appendix 10: Information sheet for starter guide evaluation

Appendix 11: Reviews of the starter guide

Appendix 12: Figure 4-9 - Designing a collaborative assignment assessment scheme

Appendix 13: Second draft starter guide to assessing collaborative assignments