The Role of Self-Regulation Feedback on Students' Self-Beliefs Systems in Physics

Kathryn Helen Bloom Milner

Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy, The University of Leeds, School of Education

January 2020

Acknowledgements

Firstly, I must thank my main supervisor Professor James Ryder, whose brilliant questions turned my head inside out to get to this point. I have learned so much about best practice supervision for my own students due to your expert modelling. Dr Jaume Ametller, I will never forget your injunction: ‘there is no substitute for data’, and I pass that on to my students too.

Thank you to the wonderful teachers and students who participated in the study, especially the two intervention teachers whose hard work and efforts with their classes, as well as with an intruding researcher, made them shine.

I am very grateful to my friends and colleagues at Leeds Trinity University and York St John University for their ongoing support over the years and am indebted to Helen Bromley who listened to and coded hours of physics lessons as my reliability rater. Quad Building Support Society (you know who you are): I finally got there.

Special thanks go to my lovely family. Ian, you are my rock, and thank you for taking the boys away countless times so that I could write in peace. Edward and William, noisy or not, I love you both dearly. Thank you for all your patience and support when I was being grumpy. Without you all, the hard work would be pointless.

Lastly to my teacher mother, who taught me a love of learning. This one was always for you.

Abstract

Teacher feedback is frequently described as one of the most important influences on student learning at many levels of education, yet few studies explore feedback from the student perspective. There is also a paucity of examination of the effects of continuing professional development (CPD) upon teacher feedback. Moreover, written feedback is often privileged over verbal feedback, both within the research literature and by students and teachers within classroom settings. This thesis reports on a longitudinal quasi-experimental study, exploring whether a CPD intervention emphasising the role of teachers' verbal self-regulation and process feedback impacted positively on students' self-belief systems (comprising self-concept, self-efficacy, anxiety, mindset) in the context of physics. The self-belief constructs were measured by pre- and post-intervention surveys and compared to a cohort whose teachers received no intervention. The study shows that the CPD intervention had a positive impact on changing teacher feedback for the two intervention classes, as well as improvement of self-concept, self-efficacy and anxiety (but not mindset) in the intervention classes when contrasted with the two comparison classes. However as shown in other studies, even highly committed teachers experience difficulties in transferring training content to real-life instructional practice. Analysis of the self-belief data provides support for both existing studies which operationalise these as separate constructs, and those which indicate instructor feedback has a mediating effect upon student self-beliefs. Using a three-level constructed typology of feedback, findings also suggest that verbal feedback has a much higher frequency within the classroom discourse than other studies have observed, and additionally that it has a role within the emotional space of a classroom in impacting self-beliefs. Additionally, qualitative insights into students' and teachers' perspectives of experiencing the intervention are presented as a contextualised case study of authentic classroom practice.

List of Contents

List of Figures

List of Tables

List of Graphs

List of Abbreviations

AfL	Assessment for Learning (formative assessment)
CG	Comparison Group
CPD	Continuing Professional Development
EEF	Education Endowment Fund
ET	Entity Theory
GCSE	General Certificate of Secondary Education
HE	Higher Education
IG	Intervention Group
IoP	Institute of Physics
IT	Incremental Theory
ITE	Initial Teacher Education
NEGD	Non-Equivalent Groups Design (Quasi-experiments)
NQT	Newly Qualified Teacher
Ofsted	Office for Standards in Education, Children's Services and Skills
SC	Self-concept
SE	Self-efficacy
SR	Self-regulation
VF	Verbal Feedback

Chapter 1 Introduction

This introductory chapter presents an overview of the study and seeks to explain why the research issue under study was important to investigate. The aims and rationale present an outline of the general research problem, including what is currently known, narrowing to a focus of the topical issue set within the context of the secondary physics classroom. The research questions are introduced, as well as a consideration of what potential outcomes may contribute to the field as new knowledge. The chapter concludes with a brief synopsis of the structure of the study.

1.1 Aims of the Study

The study arose from what is seen as an on-going issue with secondary physics education; students may attain highly within it, but many do not choose to take it at post-16 level, and females in particular have represented not much more than twenty percent of the post-16 cohorts for more than 25 years (Institute of Physics data, n.d.). There is a persistent and widely held belief amongst students and adults, documented by the Institute of Physics [IoP], amongst others that 'physics is difficult' (2006, 2008, 2012). Murphy and Whitelegg (2006) discuss the 'evidence that the perception of physics as difficult is widespread and part of the 'common-sense' knowledge of teachers and students'. This idea is perpetuated to young people who react to the increasing cognitive demand and conceptual challenge in different ways: for many it leads to a disconnect with physics through secondary education. DeWitt, Archer and Moote (2019) reveal this entrenched challenge to remain the case in their study of more than 13,000 Year 11 (15/16 years old in England) students who opted for other science choices post-16, citing 'cultural arbitraries' of physics as 'hard' and 'masculine'.

In seeking to address this cultural arbitrary, the IoP advocate an approach which involves avoiding the use of the words 'hard' or 'difficult' as student-received exculpatory reasons for not engaging fully with Physics, and instead employ the term ‘challenging’ as in: *‘Yes, it is challenging, but you can do it’*. Paradoxically, self-regulation feedback couched in this way, such as *‘You really worked hard at that, and you succeeded’* may be intended and imparted with the best of intentions, but that does not mean that it is received, internalised and acted upon in the same way. The recipient may well concentrate on the first part of the sentence rather than the latter and decide that they will stick to tasks that can be easily mastered instead, decisions that involve the domains to which they attribute their successes and failures. Hollins *et al*., (2006) indicated that student self-concept and the presence of a personally supportive physics teacher were key influences in successful student conversion to post-16 physics, and the use of language was also highlighted in the report. Given the importance of the teacher’s role, and the background of the author as a teacher of physics and a professional development professional, a case therefore emerged linking student self-beliefs in physics with the feedback that they receive from their teachers, and the need for professional development to enable teachers to deliberately use language that would build this supportive relationship and promote positive physics self-concept. That teachers need guidance and support in using the desired language of feedback in the immediate learning instance necessitates an intervention in the form of teacher professional development.

1.2 Effective Teacher Professional Development

Many studies on professional development assert its importance, and include details of key components to ensure its efficacy, such as being content-focused, using learning theory, modelling, and being of sustained duration (Joyce and Showers, 2002; Guskey, 2002a, 2000b; Kwakman, 2003; Darling-Hammond and colleagues, 2009, 2017, 2019; Yoon *et al*., 2007; Gabelica *et al*., 2012), yet it has been observed by Yoon *et al*. (2007) the *connection* between teacher professional development and student achievement seems intuitive, but that demonstrating it is difficult. Nevertheless, Continuing Professional Development [CPD] is regarded as a key method in enabling teachers to support students in learning knowledge, skills and competencies

(Darling-Hammond *et al*., 2017) at both in-service and newly qualified teachers [NQT] levels.

Although many studies indicate the effect that feedback has on developing self-beliefs (Butler and Winne, 1995; Chan and Lam, 2010; Brown, Peterson and Yao, 2016; Johnson, 2016; Ozan and Kincal, 2018), little is written on how a CPD feedback intervention on a *teacher* might impact their *students'* self-beliefs; perhaps the 'nested' nature of this as a study might be perceived as too distal a connection to claim impact, such as suggested by Yoon *et al.* (2007) above. Regarding feedback, the research has focused more dominantly on the role interventions have had on student *achievement* (Black and Wiliam, 1998; Wiliam *et al*., 2004; Hattie and Timperley, 2007; Shute, 2007; Hattie, 2009; Wiliam, 2011; Anderson and Palm, 2017). Hargreaves notes 'research on how *verbal* feedback can support learning in classrooms is currently sparse because this is a difficult area to pursue' (2014:296), and Svanes and Skagen (2017) who regard oral feedback as an important part of a teacher's repertoire, assert that it should be studied within the classroom context and further, that few studies link student learning *to* their teacher's repertoire. Having identified a gap in the existing literature, this study intends to firstly, observe what teacher oral feedback repertoires look like, and secondly, to use a feedback intervention with participant teachers to investigate whether changing their feedback repertoire to accommodate more supportive feedback language will impact on their students' self-beliefs in physics, having first ascertained a baseline of these.

1.3 The Role of Feedback

The influence of feedback has been identified as a powerful factor in improving students' learning (Black and Wiliam, 1998; Wiliam *et al*., 2004; Hattie and Timperley, 2007; Shute, 2007, 2008; Hattie, 2009; Higgins, Kokotsaki and Coe, 2011; Wiliam, 2011; EEF, 2017; Anderson and Palm, 2017, 2018), although it is important to note that 'feedback' is often an undifferentiated term within the literature, and effect-sizes for it as an intervention range from *d=0.40* to *d=0.96* (Kluger and DeNisi, 1996; Nyquist, 2003; Hattie and Timperley, 2007; Shute, 2008).

There is now increased recognition not only of the differential effects of feedback upon the learning, but on the students' own perceptions regarding the feedback (Gipps *et al*., 2000; Poulos and Mahoney, 2008; Voerman *et al*., 2012; Gamlem and Smith, 2013; Hargreaves, 2013). Additionally, the social element of such interactions, the effect upon students' cognitive (any mental skills that are used in the process of acquiring knowledge), metacognitive (awareness of their own knowledge and their ability to understand, control, and manipulate their own cognitive processes) and motivational variables, and the nature of the feedback itself are now emerging in the research as ways of identifying the variance in observed effects on learning and student response (Wiliam, 2018).

Another such emergent theme is that feedback should be purposefully *calibrated* rather than undifferentiated, and that it may take different forms to deliberately 'cue' the attention of the student to specifically, the learning task, task *processing* strategies (e.g., error-detection) or self-*regulating* strategies (e.g., monitoring learning goals), rather than being directed at the self or praise level (Hattie and Timperley, 2007, Hattie and Gan, 2010, Berg, Ros and Beijaard, 2013; Brooks *et al*., 2019). Self-regulated (SR) learning 'refers to students' skill in using a variety of learning functions and adapting this usage to the task demands at hand' (Vermunt and Verloop, 1999:276). Since SR feedback is itself the least utilised form of feedback in the classroom (Hattie and Masters, 2012, Berg *et al*, 2013, Gan, 2011), increasing its deployment is in itself an interesting endeavour.

There has been a shift in some literature from the teacher standpoint of what and how feedback is given to the student standpoint of what and how it is received, and what is its consequent adoption (Hattie and Masters, 2012), and this has obvious implications for a study on how students' self-beliefs may change in response to this calibrated teacher feedback.

1.4 Self-beliefs within the Situated Perspective

This study is framed within the social-cognitive paradigm, a psychological viewpoint of education that concerns the areas of individuals' motivation, and self, and personality. Bandura's (1997) perspective on social-cognitive theory concerns social persuasion, in that the self-belief construct of (e.g.) self-efficacy can be influenced, it can be strengthened, it could be changed, it could even be created and that it is socially dependent on actions which try to effect these sorts of changes. Dweck also operates within social-cognitive theory, but her work also encompasses developmental and personality psychology, in that it concerns the goals and beliefs of individuals, and how those goals and beliefs work together to create a meaning system in the person: 'This is what I think I am and it defines the way I operate.' (Dweck, Talk, York, 2010): an *identity*. These self-beliefs can therefore be emotionally influenced, by oneself, but also by others, and it is important to think of the self-belief constructs in process terms rather than possession terms – that is, from an interactionist perspective rather than an essentialist one. Both of these researchers' perspectives have been influential in shaping the theoretical approach to this study.

Teachers are well aware that their students self-judge their intelligence and abilities on a general basis, as well as within subject-specific contexts. It is striking that the three key influences on students' attitudes to physics identified by the Institute of Physics (IoP) are all positioned within the sociological paradigm of social learning theory. These were highlighted in the literature review by Murphy and Whitelegg (2006) as:

- self-concept (i.e. students' sense of themselves in relation to the subject: the value they place on the subject and their willingness to engage in it);
- views of physics (i.e. how students experience physics at school);
- teacher-student relationships (i.e. how personally supportive students find their physics teacher)

This has defined the theoretical framework for this study, and consequently I have decided to draw upon Wenger's (1998) conceptual framework of

Communities of Practice (CoP) as a backdrop for this study, where the situated perspective is the physics classroom.

Wenger (1998) describes learning and identity thus:

> *'Because learning transforms who we are and what we can do, it is an experience of identity. It is not just an accumulation of skills and information, but a process of becoming – to become a certain person or, conversely, to avoid becoming a certain person. Even the learning that we do entirely by ourselves contributes to making us into a specific kind of person. We accumulate skills and information, not in the abstract as ends in themselves, but in the service of an identity'* (p.215).

This passage raises three points: firstly, that a student may deliberately elect to disengage with physics precisely in order to exclude developing a physics identity; secondly, that there lies motivation behind the active acquisition of skills and knowledge rather than mere social compliance, and thirdly, that one chooses (or not) to expend valuable resource upon this acquisition process. Can feedback make a difference in the service of these choices? I hypothesise that an increase in the utilisation of *regulating* and *processing* teacher feedback will enable students to make internal adjustments to the relative position of physics in the individual's 'identity salience hierarchy' (Stryker and Serpe, 1982; Merolla *et al*., 2013; Brenner, Serpe and Stryker, 2014).

Current thinking about self-belief places it as a system comprised of constructs: self-concept, self-efficacy and anxiety, although the operational definitions, specificity and interrelation of these has been long debated (Bong and Clark, 1999; Bong and Skaalvik 2003; Ferla *et al*. 2009; J. Lee 2009; Morony *et al*. 2013), and confidence as an additional construct has been introduced in recent years (Stankov *et al.* 2012; Morony *et al*. 2013). These terms are further discussed in the literature review, but in brief, describe different aspects of an individual's relationship with their external world, and can vary with discrete aspects of that world. *Self-concept* refers to the cognitive judgements that individuals make of themselves, often based on their perception of self, compared to their peers: 'I'm not as good at maths as my friends'. *Self-efficacy* describes an individual's belief that they competently execute a specific action: 'I can calculate how many square metres of carpet I would need to cover a floor'. *Anxiety* is often linked to self-concept and describes the physio-emotional response triggered by the thought or onset of performing a specific task: 'I get very nervous doing mathematical problems'. *Confidence* has largely been viewed as a nonspecific term relating to the strengths of various beliefs and need not necessarily refer to successful outcomes: 'I can be supremely confident I will fail my spelling test'.

Individuals often form their self-beliefs based on past successes and failures, attributing the outcome in different ways to personal characteristics, and often influenced by the perception of the feedback they have been given. There has been extensive research, pioneered by Dweck (2000, 2006) on peoples' beliefs around their own 'intelligence', a term much contested in the literature. These beliefs, sometimes referred to in the literature as 'mindsets', are held to affect how individuals respond to challenges in learning, viewing them as tests of their intelligence or as an opportunity to learn more: performance goals versus learning goals. This is held to subsequently impact how enjoyment for learning is created and resilience promoted in so-called 'mastery learning', linked with self-efficacy (Bandura, 1997).

Endeavours to teach individuals who tend to blame their failures on their (perceived lack of) ability or intelligence to reinterpret their failures in terms of a different reason or attribute are known as 'attribution training' (Dweck, 2000, 2006). Such interventions seek to reposition the reason for success or failure away from attributes such as luck, ability and other people, and onto the effort expended by increased use of self-regulation strategies.

Ideally, feedback does not just correct or criticise, but also increases self-efficacy; to build students into more resilient and independent learners who are able to learn from their mistakes rather than become debilitated by them. Since 'self-regulation' feedback occupies such a small proportion of feedback generally given to students, generally less than 1% (see Chapter 2), it is difficult to correlate what outcome it may have to its use. It seems worthwhile to investigate increasing the deliberate deployment of such feedback to investigate the reaction to it.

Whilst self-regulation teacher feedback may encourage persistence, it does not compel it, and in fact can be helpful to the individual in deciding that another course of action is more suitable. In such situations, process feedback would be more helpful; how to set about the task. For this reason, although the main focus of the study is on the effect of increasing the frequency of self-regulation feedback, in those ‘failure’ situations where effort has indeed been expended, but not successfully, it is better for the student to receive feedback which enables them to change their strategy, rather than ‘try harder’ (Schunk, 2003).

1.5 Research Approach and Research Questions

This study takes the form of a quasi-experiment which aims to explore the extent to which a CPD intervention emphasising the role of teachers’ verbal self-regulation and process feedback impacted positively on students’ self-belief systems (comprising self-concept, self-efficacy, anxiety, mindset) in the context of physics. Few quasi-experiments incorporate qualitative aspects of participant data; thus, a secondary aim is to gain insights into the students’ and teachers’ perspectives during the experience of the intervention, adding to the corpus of feedback studies from the student viewpoint. Observations of teacher feedback within the situated classroom context will contribute to this ‘sparse’ area of feedback literature (Hargreaves, 2014, see also Shute, 2007; Svanes and Skagen, 2017) as the ‘social situation where feedback is given and received’ (Wiliam, 2018:1)

The teacher CPD intervention undertaken connected two research themes, introduced above. The first theme concerned the self-belief systems of students in physics in Year 10 and asked two research questions:

1. *What are the students’ self-belief systems in physics as they enter KS4?*

2. *Do the students' self-belief systems in physics change during the life of the study?*

The first question was concerned with identifying student self-belief constructs as outlined in the literature (Ahmed, 2012; Lee, 2009; Marsh et al., 2019a; Seon and Bong, 2019), and the second with investigating whether these changed over time, measured with a pre- and post-intervention survey. The second theme relating to the notion of the supportive teacher of physics (Hollins *et al*., 2006), encompassed the way in which teachers deployed feedback type orally, and whether the CPD had enabled the intervention teachers to modify their previously identified style to incorporate more self-regulation and process levels, which are less used in the classroom (Gan, 2011; Hattie and Masters, 2011; Bergh, Rose and Beijaard 2013, 2015; Brooks *et al*., 2019):

3. *What types and proportions of feedback do teachers use in their verbal interactions with students in physics?*
4. *Does the feedback pattern change during the life of the study in response to a CPD programme of flexing verbal interaction styles?*

The term 'flex' is used throughout this study to emphasise a distinction from a simple 'change' in that the teachers are being asked to not only increase the amount of self-regulation and process feedback, but to act within the learning instance to purposefully calibrate their feedback at that point, thus modifying it 'on the spot'.
A fifth research question therefore asked a process question:

5. *To what extent can data evidence an impact of RQ4 on RQ2?*

This was to be explored by combining data sets, but importantly, using the teacher voice to articulate aspects of their own learning journey through the process of engaging with the professional development intervention.

1.6 Overview and significance of the study

The study represents a significant contribution to our knowledge and understandings in several areas: additions to the literature on effective teacher CPD on feedback as well as impact of that feedback change on student self-belief change, a significantly under-researched context; additional findings regarding student perceptions of feedback are also not yet well understood at the secondary school age level; a reconceptualization of formative feedback as classroom dialogue in the form of learning conversations; the higher frequency of feedback given to students when observed during the teaching exceeding previous reports, and the role of emotion in mediating formative feedback messages.

This introduction outlined the research problem regarding students' perceptions of physics as 'difficult' and proposed that the language of teacher feedback used with students might impact their self-beliefs in physics. The literature review of Chapter 2 unites the two main theoretical frameworks, feedback and the self-belief system, and the complexities of these 'headline concepts' are delineated. It situates the polymorphous nature of feedback and its position within the accepted paradigm of assessment for learning feedback within assessment as a whole, outlining several aspects of feedback that emphasise both its use and complexity. Increasingly focusing on the particular aspects of feedback which relate to this study, the notion of self-regulation feedback is used as a bridge to a discussion of self-beliefs; the structure of self-belief systems, their role in personal identity, and how changes in self-belief can be mediated by locus of attribution of success and failure.

Chapter 3 describes and justifies the epistemological and methodological approaches taken within the study, the constructed typology (after Ramaprasad, 1983; Hattie and Timperley, 2007), and the coding decisions that enabled analysis of the observed lessons to be undertaken. Chapters 4, 5 and 6 identify how the data revealed the answers to the first four research questions and together with the teacher stories in Chapter 7, identifies the main finding that the CPD intervention on teacher feedback had been successful in improving student self-belief outcomes.

Other findings arising from the qualitative insights of the teacher and student perspectives include teacher challenges in transferring training content to their classroom practice, a contingent positive effect on student self-belief constructs with increased regulation, process and ‘next steps’ teacher feedback, teachers possess dialogic feedback characteristics, verbal feedback teacher-student interactions are more frequent than described in previous studies, and positive teacher-student relationships are an important contributor to students’ affective reception to teacher feedback. Chapter 8 collates these main findings with respect to the extant research literature, and frames them as new contributions to research, of interest to CPD and Initial Teacher Education [ITE] providers, school leaders and not least, teachers.

1.7 Chapter Summary

This chapter has presented an overview of the importance of the research issue and has begun to identify both existing literature and the gaps emerging through an exploration of such. A ‘research niche’ has been recognised and the study will now go on to explore the existing literature regarding the two main theoretical frameworks of feedback and self-beliefs in greater depth, and in discussing the importance of these concepts as contributors to student learning, enable a case to be developed for teacher professional development in feedback practices which highlights regulation and process feedback in developing student self-regulation practices and positive self-beliefs in physics.

Chapter 2. Literature Review

2.1 Introduction

This chapter examines and critically evaluates existing literature to enable a deeper theoretical and contextualised view of the research issue, incorporating aspects of teacher professional development, feedback and self-beliefs to provide an increasingly focused review of the complexities of these theoretical frameworks.

The review outlines the case for teacher professional development on feedback, drawing on evidence relating to both gains in student learning and self-beliefs, the roles of the participants within the classroom context, and aspects of professional development implementation which will impact on the subsequent research design.

The sections on feedback place it within the assessment for learning paradigm and demonstrate it to be an established and integral part of teachers' classroom verbal and pedagogic repertoire. However, teachers' implementation of it remains sporadic; the Education Endowment Fund [EEF] noted that 'it is challenging to improve the quality of feedback in the classroom'. (2017: Toolkit webpage). The multiple elements of the generic term 'feedback' are explored through firstly defining and situating it and considering research conceptualisations from differing viewpoints. A consideration of feedback models and typologies is presented to enable research design decisions to be made concerning classroom observations and classifications of teacher-to-student interactions. These oral interactions are then justified as classroom dialogue, and student perceptions as recipients of feedback are examined. Verbal feedback as a mediator of self-regulation then forms a bridge to a review of self-beliefs which may act as competency and contingency factors in forming regulatory habits

The section on students' self-beliefs is presented both generally, and with reference to physics education literature, rendering self-beliefs as a system comprised of defined constructs. The conceptual and operational distinctions between these to validate them as separate, identifiable constructs has implications for the research design to capture student data for these. A consideration of how these may also contribute to identity, and a student physics identity *per se* is offered, since this may service student subject choices away from, or towards the pursuit of physics.

2.2 The importance of feedback and the need for teacher professional development

There is a strong argument for teacher professional development on feedback from a range of themes within the research literature; a) the impact of (teacher) feedback on student achievement; b) the impact of feedback on students' self-beliefs and learning habits; c) the role of the teacher in providing feedback as the most powerful instrumental voice in the classroom, d) the role of the student in accepting and responding to the feedback, e) a consideration of the policy and leadership aspects of enabling access to teacher CPD on feedback as well as inherent stumbling blocks for effective teacher CPD.

2.2.1 The impact of teacher feedback on student achievement

Feedback has been hailed as a significant pre-requisite for student academic achievement (Black and Wiliam, 1998; Wiliam *et al*., 2004; Hattie and Timperley, 2007; Shute, 2007; Hattie, 2009; Wiliam, 2011; Anderson and Palm, 2017), typically correlating with effect sizes exceeding *d=0.73* (Hattie, 2009; Brooks *et al.*, 2019) in which however, the term 'feedback' is generic and undefined. The seminal meta-analysis of Kluger and De Nisi (1996) emphasised the variable nature of goal-related feedback on student attainment and response to feedback (see section 2.3.5 for student perceptions of feedback). The assessment meta-analysis by Hendriks, Scheerens and Sleegers (2014) concluded from the studies involved that *assessment* had little or no effect upon student achievement, attributed by Bennett (2011) to the likelihood of ineffective implementation by classroom teachers. However only two of the eleven studies involved focused on feedback, and charges of reductivism are occasionally aimed at meta-studies (Black and Wiliam, 1998a, Wiliam, 2019) in the lack of constituent comparability. Of the two feedback studies above, the Hattie and Timperley (2007) meta-analysis had an effect size d=0.96, more than twice that of the Kluger and De Nisi (1996, *d=0.41*) study (see also Nyquist, 2003 *d=0.40*; Shute, 2008 *d=0.4-0.8*).

In 2014, Hargreaves noted that ‘research on how *verbal* feedback can support learning in classrooms is currently sparse because this is a difficult area to pursue’ (page 296), and previously Shute had summarised this complexity as ‘the specific mechanisms relating feedback to learning are still murky, with very few general conclusions´ (2008:157). However, the influential Sutton Trust Toolkit authored by Higgins, Kokotsaki and Coe (2011) places feedback as the highest effect approach a teacher could employ, citing high impact for very low cost, and a potential gain on learning of +9 months. The updated Teaching and Learning Toolkit also place feedback as the highest impact intervention, with the caveat that providing effective feedback is challenging (EEF, 2017). In sum however, it is pertinent to recall that writing in 1996, Kluger and DeNisi cautioned that ‘research must focus on the *processes* induced by feedback interventions and not on the general question of whether feedback interventions improve performance’ (page 278), and Wiliam (2018) that more studies are concerned with *whether* feedback enhanced performance, rather than *how*, which leads on to section 2.2.2.

2.2.2 The impact of feedback on students’ self-beliefs

The role feedback performs in modifying students’ self-beliefs has also been explored; and although studies on CPD interventions often assess *teacher* self-belief (such as self-efficacy) as an outcome of the intervention, they do not often go on to assess the change in student self-belief. Butler and Winne’s seminal theoretical synthesis (1995) suggested that students’ beliefs and thinking (which fluctuate through a task) mediated the effect of feedback, impacting on their self-regulation as their own ‘internally generated feedback’ (page 259) and asserted even then, that traditional research on feedback 'had focused too narrowly on feedback's effect on achievement' (page 275). Boerkaerts and Corno (2005) applied existing conceptualisations of student self-regulation in the educational psychology literature to highlight volitional approaches, including dialogue and educators' understanding of process and regulatory strategies. Panadero, Jonsson and Botella's (2017) meta-analysis indicated a high effect size of (peer) assessment on self-efficacy (*d=0.73*) and a range of effect sizes for self-regulation (*d=0.23, 0.65 and 0.43*), however indicated that with the reductivist methodologies of meta-analyses, it was not possible to assert which assessment component, or combinations of components were the most effective in fostering these constructs. They did nonetheless recognise that most studies had used monitoring which relied upon immediate feedback.

A study from the students' perspective (Brown, Peterson and Yao, 2016) indicated that they believed feedback (over time) had contributed to their self-regulation and self-efficacy, but not their achievement, and two studies of the effect of evaluative feedback (Chan and Lam, 2010) indicated that both formative and self-referenced feedback had a higher impact on student self-efficacy than summative and norm-referenced feedback. However, few studies have explored the impact of a teacher feedback intervention on student self-beliefs. In her doctoral thesis, Johnson (2016) found that formative feedback (as a feedback intervention) positively impacted on students' attitudes, perceptions, willingness and motivation (in mathematics), however the students did not believe it had impacted their achievement. This was not however explored *as* the impact of teacher professional development; rather the association was assumed. Wiliam (2014) positioned formative assessment practices as vital for developing student self-regulation, but did not specify verbal feedback; Ozan and Kincal's (2018) quasi-experimental study highlighted verbal feedback as an integral part of a 28-week package of teacher professional development, and reported that feedback had increased students' learning, developed positive attitudes (towards the subject), but did not show a significant difference in the self-regulatory skills of the test group. The view overall is of formative feedback contributing to positive self-beliefs, but as argued above by the EEF (2017), the nature of the feedback is critical to its efficacy.

2.2.3 The role of the teacher

The teacher, as the recipient of feedback professional development, can also be regarded as the individual with the most powerful instrumental voice in the classroom; many studies explicitly refer to the requisite skills and knowledge that teachers must draw upon to diagnose, interpret and adapt instruction in the learning instance to formatively assess (Aschbacher and Alonzo, 2006 (written feedback); Feldman and Capobianco, 2008 (verbal feedback); Furtak, 2012 (verbal); Gamlem and Smith, 2013 (verbal); Gottheiner and Siegel, 2012 (verbal); Hargreaves, 2013, 2014 (verbal); Havnes, Smith, Dysthe, and Ludvigsen, 2012 (written and verbal); Lee *et al*., 2012 (verbal); Yin *et al*., 2014 (verbal); Elliott *et al*., 2016 (written)), which has obvious implications for teacher professional development. Additionally, many of these studies (labelled verbal, above) identified ‘dialogue’ and ‘discussion’ as a frequent formative assessment device involving teacher-to-student interactions, and this aspect is further explored in section 2.3.4. This also necessitates teachers’ ability to both interpret student learning and construct questions and prompts ‘on the spot’ (Feldman and Capobianco, 2008; Gottheiner and Siegel, 2012; Lee *et al*., 2012; Hargreaves, 2014), thus verbal feedback [VF] becomes a critical factor in achieving this. Moreover, teachers’ values and philosophies around teaching and learning were found to influence their implementation of formative assessment approaches (Havnes *et al*., 2012; Lee *et al*., 2012) as well as the extent to which learning goals, not just performance goals and ‘coverage’ of the curriculum were important to the teacher (Aschbacher and Alonzo, 2006).

2.2.4 The role of the student

Wiliam (2014) questioned whether students are always the beneficiary from the feedback process in that ‘next steps’ instruction may not always be ‘correct’, the student may not recognise feedback as regulation of their learning process, and indeed that learning is not always contingent upon the feedback (see also Nicol & Macfarlane-Dick, 2006; Hounsell, McCune, Hounsell, & Litjens, 2008; Wiliam, 2011; Brooks *et al*., 2019). Feedback studies on the student perspectives as recipients are less common (Gipps *et al*., 2000; Poulos and Mahoney, 2008; Voerman *et al*., 2012; Gamlem and Smith, 2013; Hargreaves, 2013) and section 2.3.5 below discusses student perceptions of feedback in more detail. Since the purpose of feedback is to promote learning, it follows that students may need guidance alongside the feedback message (Hounsell *et al*., 2008) and time to make an adjustment within the learning instance (Wiliam, 2011); gaining greater clarity on the student perspective of formative feedback is desirable since this is this is not as well explored in the literature (Poulos and Mahoney, 2008; Voerman *et al*., 2012; Hargreaves, 2013).

2.2.5 Teacher access to and barriers concerning professional development

Heitink *et al*. (2016) address the need for schools to adopt a culture of teacher autonomy and collaboration since school leaders play an important part in adopting and ensuring formative assessment policy (Aschbacher & Alonzo, 2006; Havnes *et* al., 2012; Feldman and Capobianco, 2008). Yoon *et al*. (2007) observe that the connection between teacher professional development and student achievement seems intuitive, but that demonstrating it is difficult. However, numerous studies on *effective* professional development (Darling-Hammond *et al*. 2017; *see also* Joyce and Showers, 2002; Darling Hammond and colleagues, 2009, 2019; Gabelica *et al*., 2012; Voerman *et al*., 2015) indicate the necessity of several inherent key features:

1. Is content focused
2. Incorporates active learning utilizing adult learning theory
3. Supports collaboration, typically in job-embedded contexts
4. Uses models and modelling of effective practice
5. Provides coaching and expert support
6. Offers opportunities for feedback and reflection
7. Is of sustained duration

and access to such professional development is variable, especially when factoring in the potential need to release teaching staff for a course of sustained duration, and cost: Yoon *et al*. (2007) assert for instance, that less than thirty hours of CPD is 'ineffective'. These aspects must be considered within the development of any CPD intervention (see Chapter 4).

As noted previously, much of the literature on teacher feedback CPD does not go on to map changes in student self-belief, but is more likely to map changes in teacher *frequency* of verbal feedback (Voerman *et al*., 2015; Schütze *et al*., 2017), amount of written feedback (Aschbacher and Alonzo, 2006), student conceptions of science (Gottheiner and Siegel, 2012) and achievement (Phelan *et al*., 2012; Andersson and Palm, 2017). Studies have identified the need outlined above for assessment professional development to be sustained (Aschbacher & Alonzo, 2006; Gottheiner & Siegel, 2012; Lee *et al*., 2012; Phelan *et al*., 2012; Voerman *et al*., 2015; Andersson and Palm, 2018), and Gottheiner and Seigel (2012) highlighted the need for teacher CPD in verbal feedback to be more effective in closing the gaps in student learning.

CPD studies on feedback and assessment professional development (Wiliam *et al*., 2004; Schneider and Randel, 2010; Wiliam, 2010; Berg, Ros and Beijaard, 2014, 2015; Heitink *et al*., 2016; Zaccarelli *et al*., 2018) as well as generalised teacher professional development (Guskey, 2002a, 2000b; Kwakman, 2003; Darling-Hammond *et al*., 2017, 2019) indicate that even teachers of some experience may find it challenging to change their established classroom practice, and that professional development must be sustained to effect a change on teacher knowledge and practice. As indicated by Feldman and Capobianco (2008), Gottheiner and Siegel (2012), Lee *et al*. (2012), and Hargreaves (2014) above, verbal interactions require immediate evaluation and response, that are simultaneous social and academic demands upon the educator (Hammerness *et al*., 2005; Korthagen, 2010). Eraut (2004) described how teachers develop intuitive routines over time to cope with this demand, which results in a 'routinisation' of practice, enabling their attention to be freed up for changing situations. This has implications therefore for the patterns of verbal feedback which over time they have become used to, since

Schütze *et al*. (2018) describe how even effective teachers may find this difficult to change.

In the next section, the theoretical notion of feedback will be considered in more depth, comprising its position with the assessment for learning paradigm, and indeed within assessment as a whole, and offer a definition for the purpose of this study. Literature relating to feedback will be examined to demonstrate how it has been interpreted, and consequently how the roles and perceptions of both teachers and students may impact on both the provision and reception of the feedback.

2.3 Teacher Feedback

Whilst the notion of feedback is positioned amongst other teaching strategies, it has assumed greater prominence in recent years due to its reported high average effect size, although this ranges for it as an intervention from *d=0.40* to *d=0.96* (Kluger and DeNisi, 1996; Nyquist, 2003; Hattie and Timperley, 2007; Shute, 2008). It is notable also that it had a position as an influence under the *teacher* domain (Hattie, 2009, 2012a; Hattie and Clark, 2019), as opposed to those of home, school and curriculum. In effect-size studies, control groups are taught with conventional methods and an experimental group with an intervention method. Both are given a pre- and post-test, and the results are directly compared. The shift, whether positive or negative, is measured in standard deviation units and is termed the effect size. An effect size of 0.5 is given as equivalent to the difference between two adjacent grades at GCSE, thus feedback (even generalised) at 0.79 is desirable indeed. It should be noted however, that assessment authors such as Wiliam have voiced concerns about the reductive nature of meta-analyses in the formative assessment field, citing issues such as inappropriate comparisons, variations in intervention quality and population variability, selection of studies and sensitivity of outcome measures (Black and Wiliam, 1998a; Wiliam, 2018, 2019)

Whilst feedback is regarded as a predominant tool in the teacher’s toolkit, the picture is not a clear one for several reasons: schools may have admirable feedback policies in place, but it is individual teachers who enact it, or not

(EEF, 2017). Further, it does presently tend to exist as a 'unidimensional notion' (Hattie and Masters, personal communication, 2012) where there is great assumption about what both teachers and students mean by this undifferentiated term, and how it is used. The question also needs to be asked 'is the feedback fit for the purpose intended?'; that is, what differences are there between the feedback we intend, what is perceived by the student, and how it is internalised? Hattie (1999) pronounced 'The simplest prescription for improving education must be 'dollops of feedback' - an attitude he has since moderated due to recognising the variable effects undifferentiated feedback can have. He now claims, 'there to be as much ineffective feedback as effective feedback' (2012b:19), perhaps in recognition of the Kluger and De Nisi (1996) research below, and other meta-analyses.

Price *et al*. (2010) regard feedback as a generic term disguising multiple purposes and organise it into five categories in a nested hierarchy, comprising correction, reinforcement, forensic diagnosis, benchmarking and longitudinal development. This rather strict view is challenged by an exploration of the myriad dimensions of education feedback; the individuals associated with the interaction, its nature and purpose, its form (written or verbal), model and learning conditions, all of which are deemed to impact the efficacy of the feedback. We turn first to a definition of what feedback is, and what it is for.

2.3.1 Situating and defining feedback

Dann (2019:353) asserts that '*feedback is typically seen as a key component of formative assessment, even when used to drive forward learning which will ultimately be part of summative assessment*'. Assessment, described by Broadfoot (1999:3) as 'a dangerously ambiguous concept' exists as a technical craft as well as a social technology (Madaus, 1994). Elbra-Ramsay (2019:37) extended this description to indicate the current assessment landscape as consisting of '*two* ambiguous and diametrically opposed concepts: assessment for learning (formative assessment), and performativity'. Authors have explored how these competing concepts may result in the mechanisation of formative pedagogic practice in the service of summative outcomes (Torrance, 2012; Dann, 2018; Wiliam, 2018) within the relational space of the classroom.

There is little doubt that 'feedback' is a disputed term; no one clear and uncontested scholarly definition exists in the literature. It is not the aim of this study to provide a definition fashioned from the output but rather to draw upon existing theoretical perspectives. In recounting the origination of the word feedback as a concept that came into everyday use from the field of communication engineering in the early twentieth century, Roos (2004) claimed its early nuances within that discipline were lost in the translocation to education parlance, and that feedback gained a more restricted meaning, becoming 'merely the return flow of information'. Shepard (2002) highlighted a major flaw in viewing it in these terms:

> *'The idea of feedback comes from electronics where the output of a system is reintroduced as input to moderate the strength of a signal. Correspondingly, both behaviourist and constructivist learning theories take for granted that providing information to the learner about performance will lead to self-correction and improvement.'* (Shepard, 2002:1091-2)

Although writing from a perspective in information management, Ramaprasad's definition of feedback as 'information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way' (1983:5) and further, that the information itself is not feedback; 'For feedback to exist, the information about the gap must be used to alter the gap. If the information on the gap is merely stored without being utilized to alter the gap, it is not feedback' (*ibid*). This formed the basis for the phrase 'feedback loop' which is in use today (Askew and Lodge, 2000; Mislevy, 2012). Sadler (1989) refined Ramaprasad's notion with respect to education applicability 'in terms of its effects, rather than its relational content' (1989:120), since he argued that in educational settings, student output cannot always be evaluated as correct or incorrect. Sadler's definition of the control loop outlined a prerequisite of 'indispensable conditions for improvement' in that 'the student...is able to monitor continuously the quality of what is being produced during the act of production itself' (1989:121). This argument proved to be seminal in later formative assessment discussions and literature, and Sadler became credited with the phrase ‘closing the gap’.

Kluger and DeNisi offered a definition in their feedback metastudy discussed below as ‘actions taken by (an) external agent (s) to provide information regarding some aspect(s) of one’s task performance’ (1996:255). Whilst this agrees with some aspects described above, it does not take into account the dynamics of *who* is involved in feedback interactions and precludes the possibility of self-assessment on the part of a learner. This definition is thus too narrow for the purpose of this study. Similarly, the suggestion by Nicol *et al.* (2006:208) of feedback as:

> *...good quality external feedback is information that helps students troubleshoot their own performance and self-correct; that is, it helps students take action to reduce the discrepancy between their intentions and the resulting effects'* (Nicol *et al*., 2006:208).

also focuses on external input, and thus is teacher-centric, although the notion that students are able to take action from the information is an important addition. The study, although applied in the higher education field, also explores the framework of cognition, motivation and belief which is central to the issues investigated in this thesis, which are presented below.

In conjunction with higher education students at an Australian university, Scott (2014) sought to discover what a student-centred definition would be, and offered ‘feedback is the means by which they can at each stage of a course gauge how they are going in terms of the knowledge, skills and understanding that will determine their overall result in the course’ as the result of the study (p.56). Whilst this characterisation celebrates knowledge, skills and understanding, its focus on the result of undertaking a university course renders it less serviceable to this thesis.

It is necessary to view a definition of feedback which does not necessarily identify both donors and receivers, and which the nature of the information and what is to be done with it is sufficiently broad. Although an older source, Butler and Winnie depict feedback as:

> *'information with which a learner can confirm, add to, overwrite, tune, or restructure information in memory, whether that information is domain knowledge, metacognitive knowledge, beliefs about self and tasks, or cognitive tactics and strategies'* (1995:275).

The use of Butler and Winne's feedback definition for the purpose of this study is not intended to discount or discredit other authors' characterisations but merely to underpin the theoretical exploration of student learning in terms of cognition, metacognition and particularly the students' self-belief systems in the physics classroom over time.

Feedback cannot exist in a vacuum; it is both contextualised and personal. It is but one component of a formative assessment repertoire and should be part of a classroom environment in which 'students see constructive criticism as a good thing and understand that learning cannot take place without practice' (Brookhart, 2008). Assessment is held to function formatively when it fulfils three key processes, as described in Black and Wiliam's seminal study *'Inside the black box'* (1998b): to ascertain where the learner is going; where the learner is now; how the learner is going to get there, colloquially known as 'closing the gap', from Sadler (1989). It is clear that this can take place in a number of ways, within different timeframes, and between different participants, notably; teacher, student and peer. There has been a shift from viewing this as a purely teacher process or 'teacher regulation' to a process in which the students' input is required as well.

Leahy *et al.,* (2005) instigated a number of different teacher-involved projects during 2004-5, including three-day workshops introducing them to relevant research literature, monthly meetings, and follow-up classroom observations that led the researchers to define the 'territory' of assessment for learning as five broad approaches:

1. Clarifying and sharing learning intentions and criteria for success;
2. Engineering effective classroom discussions, questions, and learning tasks;
3. **Providing feedback that moves learners forward** *[my emphasis];*

4. Activating students as the owners of their own learning;
5. Activating students as instructional resources for one another.

Wiliam (2011) highlights the difficulty of creating a definition of formative assessment which adequately covers all these points, and offers the following:

> *'an assessment functions formatively to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have made in the absence of that evidence.'* (p.43)

Formative feedback is often described as minimal and over-shadowed by the prevalent use of summative, criteria-referenced grades. In 1996, Broadfoot highlighted that this dominant role of criteria-referenced assessment as policy has led to educational situations of competition and control, both between individuals and institutions, rather than an educational situation of motivation and enabling learning to take place, a situation still commented upon more recently (Torrance, 2012; Dann, 2018, 2019).

A number of authors emphasise the relationship between learning goals (which may be characterised and shared as success criteria or learning intentions) and feedback, which are the first and third approaches in the list of Leahy *et al.* (2005) above (Nicol and Macfarlane-Dick 2004; Liu and Carless 2006; Hattie and Timperley, 2007; Brookhart, 2012; Hattie, 2012b; EEF, 2017; Hattie and Gan, 2017;). Hattie and Timperley's feedback model (2007), which forms a theoretical underpinning for this study is goal-oriented, and calls attention to the importance of having clear goals; goals without clarity are too vague to enhance learning. It must be recognised that these sources focus predominantly on written feedback to students, however the EEF (2017) and Elliott *et al*. (2016) do indicate that feedback may take place in written, verbal *and* peer-reviewed forms.

For the purposes of this study however, it is the teacher verbal feedback [VF] which will be investigated. Hargreaves (2013, 2014) conceptualises VF as part of classroom dialogue (see also Ajjawi and Boud 2017), and Elliott *et al*. (2016), although reporting predominantly on written feedback (marking), do list re-teaching concepts in class, guidance from the teacher during class time, pupil-teacher dialogue, and questioning all as forms of VF. Indeed, Svanes and Skagen (2017:334) argue that ‘the focus of some of the feedback literature appears too narrow to understand what is going on in a classroom’, and that ‘feedback has to be studied in a classroom context and as a part of the teaching process to become more useful for teachers and pupils’. They too assert verbal feedback to be an important part of a teacher’s repertoire.

These aspects of feedback as both a process embedded within other classroom strategies and a form of communication that is dependent upon the individual deploying it give rise to a consideration of how researchers, and teachers themselves conceptualise feedback, and this will be explored in the next section.

2.3.2 The analogous and metaphorical faces of feedback: research conceptions of feedback

Definitions of feedback are often influenced and characterised by different views of learning: ‘feedback is a complex notion, often embedded in a common-sense and simplistic dominant discourse’ (Askew and Lodge, 2000:2). The model of teaching (receptive-transmission, constructivist etc.) and view of learning, from both teacher and student viewpoints can shape how feedback is described and intended. Analogies and metaphors are becoming a common tool within the feedback discourse; these have their origins in the researchers’ observations of classrooms, where teachers and lecturers enact these forms both verbally and in written form:

- feedback as a loop (Askew and Lodge, 2000; Mislevy, 2012)
- ‘feedback as telling’ moving to ‘appreciating’ (Boud and Malloy, 2012, 2013)
- feedback as a consequence (Hattie and Timperley, 2007)
- feedback as talk (Ajjawi and Boud 2018);
- feedback as dialogue (Ajjawi and Boud 2017, 2018; Carless *et al.*, 2011; Hargreaves, 2-13, 2014; Svanes and Skagen, 2017)
- ‘feedback as a gift’ (Askew and Lodge, 2000; Hargreaves, 2005)

- the feedback triangle (Yang and Carless, 2013)

An in-depth exploration of eight alternative faces of feedback was proposed in Elbra-Ramsay's thesis (2019), drawn from her own literature review. She outlined feedback conceptualisations as: monologism (transmission), dialogism (dialogue), identity, self-determination (motivation), self-regulation (empowerment), relationships, performance (compliance) and agency. In separating motivation and self-regulation in this way, Elbra-Ramsay cites Ryan and Deci's assertion (2000:69) that in self-determination theory, motivation is not one single construct. However, for the theoretical framework of this study, self-regulation feedback *is* defined as including motivational aspects, as in section 2.3.6 below (Butler, 1987; Hattie and Timperley, 2007; Locke and Latham, 1984; Nicol and Macfarlane-Dick 2004).

Given that educators at all levels will have their own philosophy of what feedback both is and should achieve as reflected in the list above, it is unsurprising that students may not always correctly identify that they have received feedback when they have, nor does it follow that they have the time and inclination to act upon it (Glover and Brown, 2006). For the purposes of this study, the conceptualisation of feedback as both 'talk' and 'dialogue' (Ajjawi and Boud 2017, 2018; Carless *et al.*, 2011; Hargreaves, 2-13, 2014; Svanes and Skagen, 2017) is most relevant for the intention for verbal teacher feedback to attempt to affect a change in students' self-beliefs in the learning instance.

2.3.3 Feedback models and typologies

'*The main purpose of feedback is to reduce discrepancies between current understandings and performance and a goal'* (Hattie and Timperley, 2007:86).

Feedback can thus be considered from the student perspective, but it is also important to consider what the feedback is *about*. Amongst others, the aims of feedback are to offer specific task-related assistance, clarification or challenge of alternate conceptions, focus (or re-focus) on the task, motivation towards a goal, and so on. Clearly one point of feedback cannot achieve all these things, and a number of sources emphasise the importance of feedback specificity (Sadler, 1989; Tunstall and Gipps, 1996; Gipps *et al*., 2000; Hattie and Timperley, 2007; Shute, 2007, 2008; Wiliam, 2011). As Hattie says: ‘feedback thrives on error’ (2012:115); feedback is most effective when it used to address deficits in knowledge, skills and application and supply information to enable the student to move between their current level to the desired goal (Ramaprasad, 1983; Sadler, 1989; Kluger and DeNisi, 1996; Tunstall and Gipps, 1996; Black and Wiliam, 1998b; Hattie and Timperley, 2007; Shute, 2007, 2008; Voerman *et al*., 2012, 2015), although these authors have presented different feedback *models*.

In their seminal study, Kluger and DeNisi (1996) described three generalised, hierarchical levels of feedback comprising task learning (at the base), task motivation and meta-tasks, the latter including self-related processes, such as reflection. They highlighted that the feedback intervention effectiveness appeared to decrease as the focus moved up from task level to the self, which has implications for this study, since SR feedback is posited as desirable.

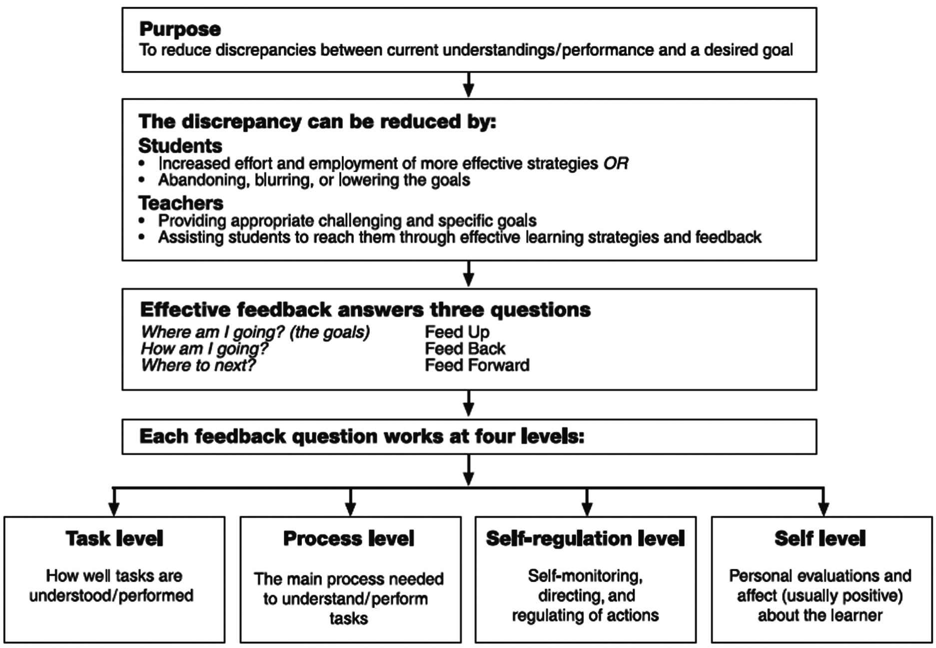


Figure 2.1 A model of feedback to enhance learning (Hattie and Timperley, 2007).

Hattie and Timperley (2007), amplified in Hattie (2012a) proposed a model incorporating three feedback questions, corresponding to the terms ‘feed up’ (where am I going? What is the goal?), ‘feed back’ (how am I going? What progress is being made towards the goal?) and ‘feed forward’ (where to next? What activities need to be undertaken next to make better progress?), which is shown in Figure 2.1.

This model also distinguished four *levels* of feedback, differing to Kluger and DeNisi’s (1996) levels: introducing the typology of Task, Process, Self-regulation, and Self, as shown in Figure 2.1 also. They regarded feedback on the self as the least effective form of feedback, which did concur with Kluger and DeNisi’s (ibid) as it did not appear to enhance learning and was occasionally detrimental to learning (see also Brooks *et al*., 2019; Hattie and Timperley, 2007; Kluger and DeNisi, 1996; Mueller and Dweck, 1998; Shute, 2008), although Voerman *et al.* (2014) disagreed. Hattie (2012a) gave approximate frequencies of incidence (through moderated observations in 18 secondary Australian classrooms) as:

• Task (59%) – how well someone performs on a task, or product, usually information-based, for example, ‘*how their answer meets the success criteria, and how they might elaborate on the answer’.* It can be very powerful, especially when correcting faulty interpretations, but often does not generalise to other tasks.

• Process (25%) – aimed at the processes used to complete the task, for example, strategies for error-detection, or effective data-gathering, for example, ‘*what is wrong, and why? What are the relationships between this-and-this part of the task? Can you look for similarities and differences?’*

• Self-regulation (2%) – enabling students to monitor and ultimately self-direct their learning, improve confidence and be willing to invest in making the effort, for example, ‘*what learning goals have you achieved? How did your ideas change during the task, and why?’*

• Self or praise orientated (14%) – praise, often non-specific, which can often direct attention away from the task or strategy, such as *'You're a great student! Well done!'*.

Recent similar audit studies of types of classroom feedback interactions are summarised in Table 2.1 below, and show comparable percentage amounts:

	Hattie and Masters (2011)	Gan (2011)	Bergh, Ros and Beijaard (2013)	Brooks *et al*., (2019)
Level	**18 secondary classes (*given* to students by teacher)**	**235 peers (post-16 Chemistry students)**	**32 teachers in 13 Dutch primary schools**	**28 students, 1 teacher in Australian primary school**
Task	59%	70%	51%	77.8%
Process	25%	25%	42%	15.9%
Self-regulation	2%	1%	2%	0.6%
Self	14%	4%	5%	*Omitted from study*

Table 2.1 Frequencies and types of feedback interaction from a range of studies.

Hattie asserts that the first three of these levels correspond to phases of learning: from novice, through proficient, to competent:

'*Novices mostly need task feedback; those who are somewhat proficient mostly need process feedback; and competent students mostly need regulation or conceptual feedback'* Hattie (2012b:21)

It would appear that Hattie views ‘proficient’ as ‘progressing’ towards competent, and in need of process guidance to reach it. In contrast, Schunk (2003) advises that what he termed ‘effort’ [self-regulation] feedback to students in the early stages of learning is more beneficial, but adds the caveat that as they develop skills, feedback should move to markers about effective strategy use. If a student is expending effort but not succeeding, feedback on their effort is akin to salt in the wound; here a shift to strategy/process feedback to enable them to progress is what is required. Hattie and Timperley (2007) describe seeking help as a learner proficiency and discuss the distinction between *instrumental* help (seeking hints on how to work something out) versus *executive* help (seeking answers).

Influenced by both the Hattie and Timperley (2007) paper which highlighted the role of specific feedback in providing evaluative information from the teacher to the student, as well as Shute (2008), Voerman *et al*. (2012) contributed to the literature by an analysis of specific feedback that could be characterised as either *progress* or *discrepancy* feedback. Historically, and in line with feedback as a deficit model (Elbra-Ramsay, 2019), discrepancy feedback or ‘feed forward’ (Hattie and Timperley, 2007) is deemed to provide greater learning gains than progress feedback (Kluger and DeNisi, 1996; Shute, 2007, 2008).

Unsurprisingly in a teacher feedback study, Voerman *et al*. (2012:1108) provide a teacher-centric definition of feedback ‘as, information provided by the teacher concerning the performance or understanding of the student, with reference to a goal and aimed at improving learning’. However, the authors position *progress* feedback as relating to the ‘where am I now? How am I going? How am I doing in relation to the learning goal?’ described by Hattie and Timperley (2007) as ‘feedback’, which agrees with other authors’ notions of *progress* feedback (Schunk and Swartz, 1993; Schunk and Ertmer, 1999; Gipps *et al*., 2000; Svanes and Skagen, 2017). Discrepancy feedback should provide specific information about the deficit between the current position and the learning goal, and so this equates to Hattie and Timperley’s notion of the feed forward ‘where to next? What activities need to be undertaken next to make better progress?’ (*ibid*). This is shown in Figure 2.2 below.

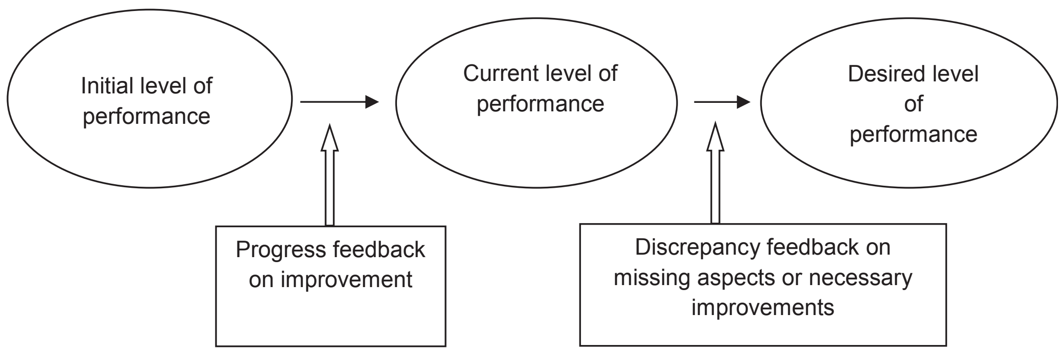


Figure 2.2 Voerman, Meijer, Korthagen and Simons (2012, 2015) progress and discrepancy feedback model

By re-imagining this in terms of Ramaprasad's (1983) notion of the feedback loop, the flow in Figure 2.2 above can be represented as a cycle (Figure 2.3) in which a progress feedback statement of 'where you are now' may be followed by a discrepancy feedback prompt of 'this is where you go next, and this is how you get there'. It is envisaged that reframing it in this way may enable swifter identification of VF type during classroom observations.

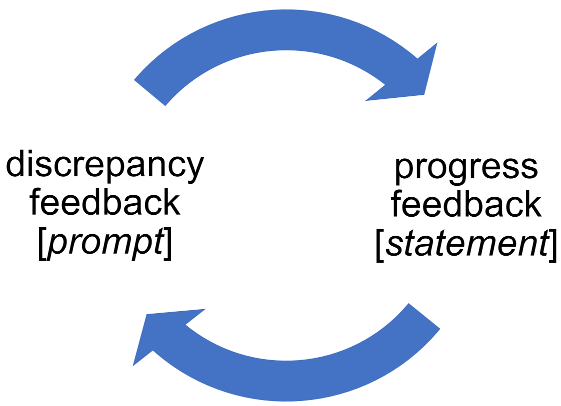


Figure 2.3 Voerman *et al.* (2012, 2015) progress and discrepancy feedback model re-imagined as a feedback loop (after Ramaprasad, 1983)

The Tunstall and Gipps, (1996) and Gipps *et al*. (2000) typology leads into both positive versus negative feedback, and achievement versus improvement feedback as shown in Figure 2.4:

	Evaluative		**Descriptive**		
Positive Feedback	**A1 Rewarding** *Giving rewards*	**B1 Approving** *Expressing approval*	**C1 Specifying attainment** *Telling students they are right/wrong* *Describing why an answer is correct* *Telling students what they have/have not achieved*	**D1 Mutual construction of achievement** *Discussing with students the features of a piece of work*	*Achievement Feedback*
Negative Feedback	**A2 Punishing** *Giving punishments*	**B2 Disapproving** *Expressing disapproval* *Reprimands; negative generalisations* *Negative non-verbal feedback*	**C2 Specifying improvement** *Specifying or implying a better way of doing something*	**D2 Mutual construction of achievement** *Getting students to suggest ways they could improve*	*Improvement Feedback*
	Evaluative		**Descriptive**		

Figure 2.4 Tunstall and Gipps (1996) and Gipps *et al*. (2000) feedback typology

Feedback types A and B are more aligned to *summative* assessment practices, and thus enhance performance goals, whereas C and D types are more associated with learning goals, although discrepancy feedback C2 and D2 are framed as negative. This typology was drawn from a study of very young children, not secondary students, and it is unclear whether the focus was on verbal or written feedback, or both, nor did it encompass every type of oral interaction. Gamlem and Smith (2013) drew upon Tunstall and Gipps' typology (Figure 2.5), and applied it to lower secondary students in Norway although the sample size was small (11 students).

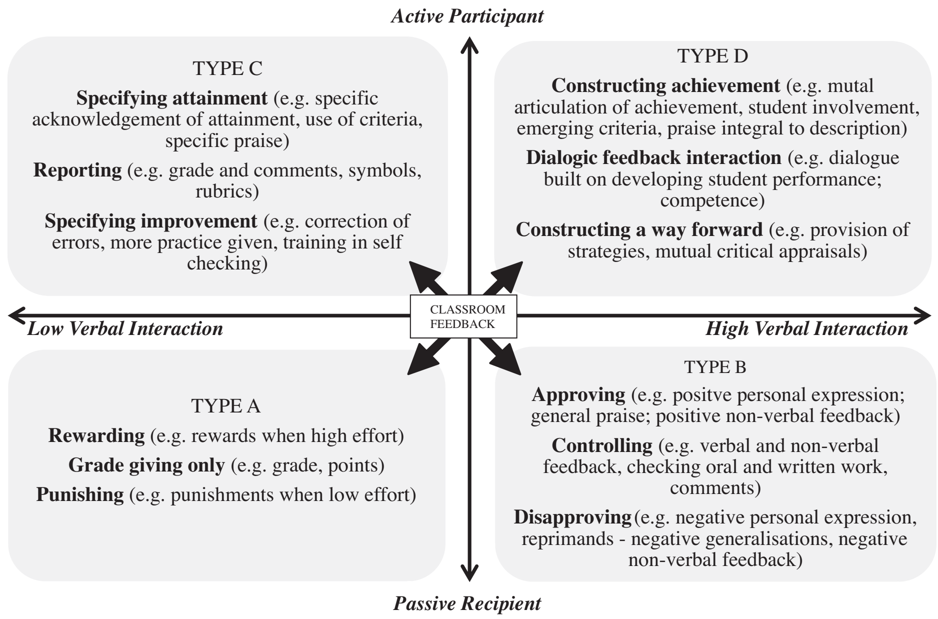


Figure 2.5 Typology on classroom feedback in lower secondary school (Gamlem and Smith (2013), modified after Tunstall and Gipps (1996) and Gipps *et al.* (2000).

Since the focus of Gamlem and Smith’s study was student perceptions, the continuum of low to high verbal interaction is a useful one. However, despite Knight (2003) successfully using the Tunstall and Gipps typology to evaluate the quality of teacher feedback to students, there are placings within this model which may be perceived differently, depending on the student perception. For instance, in C1, ‘telling pupils what they have done wrong’ is in the positive feedback section but may be comprehended as negative by the student. Equally, in D2, ‘getting the students to suggest ways to improve’ is in the negative half yet could be executed in a positive learning manner. Additionally, this typology may be overly prescriptive to map the nature of teacher-to-student verbal interactions in the moment of classroom dialogue, to which we now turn.

2.3.4 Teacher Verbal Feedback as a Contributor to Classroom Discourse

Unless it is the focus of the feedback study, what tends to be less well highlighted in is whether feedback itself is written, or verbal, or both; on deeper inspection, many studies seem to be concerned only with written feedback, perhaps influenced by the enduring criticism of the predominantly monologic way feedback is utilised in higher education (Ajjawi and Boud, 2018; Boud and

Malloy, 2013; Nicol, 2010). Verbal feedback [VF] however, is more likely to be given at the learning instance, which characterises more effective assessment for learning principles, rather than at endpoint, and yet it is not as well discussed in the literature (Ruiz-Primo and Furtak, 2007; Hargreaves, 2013, 2014; Svanes and Skagen, 2017). Kerr (2017) indicated in her study that secondary-level students perceived verbal feedback as a 'form of focused conversation, different from normal classroom dialogue', however the sample size was very small (*n*=4), and 'normal dialogue' was not defined. Svanes and Skagen (2017:336) emphasised that feedback was an important part of a teacher's repertoire and assert that 'Many studies within traditional classroom research do not link pupils' learning to teachers' teaching repertoires.'

Torrance and Pryor (1998) and subsequently Pryor and Croussard (2008) developed a formative assessment model in which constructivist-to-socio-cultural perspectives were embedded, describing formative assessment as a discursive social practice (2006:1,8), involving dialectical and sometimes conflictive processes. Both Dann (2019) and Elbra-Ramsay (2019) discuss how feedback can be viewed as a relational concept with links to socio-cultural theory, in which language is a key construct. Both Vygotsky's notion of the zone of proximal development, and his assertion that learning is socially constructed (1978) can be used to explore how language can close the learning gap (Sadler, 1989):

> *'It is the distance between the actual level developmental level as determined by independent problem solving and the level of potential development as determined by independent problem solving under adult guidance or in collaboration with more capable peers.'* (Vygotsky, 1978:86).

Dann (2019) also acknowledges the dichotomy all despite(verbal) formative assessment being 'in the moment of production' (p.361), the learning and teaching oppositions towards explicit performativity standards. She further argued a caution in assuming that students are intentional learners and wish to 'play the game' that school projects onto them (Dan, 2016, 2018, 2019:362). Wiliam (2011) highlighted that 'feedback should cause thinking', and this is central to Dann's (2019) relational concept of feedback in creating agency for the students (Bandura, 1997)

Within the relational space of the classroom, the juxtaposition of power becomes crucial. Gamlem and Smith's (2013) study on student perceptions of feedback indicated that 'trust' of the teacher was paramount, and affected attention to feedback given. Dan (2019) regards this as a form of relationship where students accept power differences yet retain autonomy within the relationship. Using a lens of metaphorical economies and a ribbon model to represent the teacher and student 'sides' of the feedback situation, Elbra-Ramsay (2019), proposed the tensions between triadic factors of morality, pedagogy and relationships as being constantly in motion, shown in Figure 2.6.

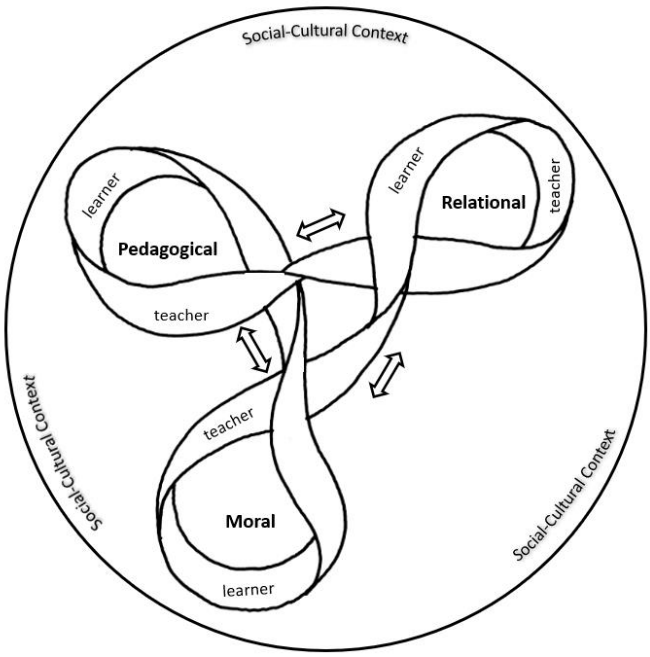


Figure 2.6 The Feedback Ribbon model (Elbra-Ramsay, 2019)

Given that within the learning environment, verbal feedback is most likely to take place within classroom dialogue, Howe and Abedin (2013) conducted an extensive systematic review on four decades of classroom dialogue research, with a focus on primary and secondary classrooms (rather than higher education). They drew upon 225 studies to define the nature of dialogue as '*all* verbal exchanges (and series of such exchanges) where one individual addresses another individual or individuals and at least one addressed individual replies' (p.326), and add that verbal exchanges may be textual, and not restricted to talk. Kerr (2017) notes that verbal interactions within classes may encompass a range of non-specific communication, and other forms may comprise argumentation, conversation, instruction, behaviour for learning, and even 'banter' as well as feedback.

As a prominent author in the dialogic field, Alexander (2006, 2017) has been influential in helping teachers to understand the nature of their interactions, and define dialogic teaching 'collective, reciprocal, supportive, cumulative and purposeful' (Alexander, 2006:28). Commenting on the value of *talk*:

> *'vitally mediates the cognitive and cultural spaces between adult and child, among children themselves, between teacher and learner, between society and the individual, between what the child knows and understands and what he or she has yet to know and understand'* (Alexander, 2008:92).

Scott and colleagues (Mortimer and Scott, 2003; Scott, Mortimer, and Aguiar, 2006; Scott and Ametller, 2007) have influenced notions of dialogue in science education for their studies on four classifications resulting from the combinations of two dimensions; authoritative – dialogic and interactive – non-interactive. Howe and Abedin (2013:327) thus describe these as 'authoritative interaction' (discussion amongst several individuals about a single, received idea), 'dialogic interaction' (discussion amongst several individuals about contrasting ideas), 'authoritative non-interaction' (one individual's assertion of a single, received idea) and 'dialogic non-interaction' (one individual's commentary on contrasting ideas), and consider their definition to exclude the 'dialogic non-interaction' category. Nonetheless, when considering different pedagogical approaches within the classroom, there are occasions where the teacher adopts a monologue in conveying ideas.

Mercer (2007) identified both Alexander (2006) and Scott (2003) as noted influences on the conceptualisation of dialogic teaching. Although each author employs the term to mean a different entity; Alexander describes a pedagogic approach and provides 47 indicators, whereas Mortimer and Scott (2003) use dialogic to define the nature of interactions. Mercer (*ibid*) rationalised that these two conceptions of ‘dialogic’, whilst similarly rooted in pedagogic efficacy, ‘having developed independently, are difficult to reconcile into one analytic scheme’ (p.3). His study draws on both, using Alexander’s (*ibid*) 47 indicators to serve in the classroom, and Mortimer and Scott’s (*ibid*) framework for the analysis of the interactions, the nature of which is the one of the foci of this verbal feedback thesis.

It was notable in Howe and Abedin’s review (2013) that more was recognised about how classroom dialogue was organised than in considering the efficacy of different modes: 158 of the 225 studies characterised the dialogue *in situ* but did not remark upon its consequences. Yet the role of verbal feedback as interaction between teachers, learners and peers is vital in enabling progress in learning (Kerr, 20107). It is also important, from the student perspective, that they are able to isolate verbal feedback from (for example) more instructional interactions in order to be able to act on it more effectively. Sinclair and Coulthard (1975, cited in Howe and Abedin, 2013) identified the three-step pattern of Initiate – Respond – Feedback or – Follow-up, now known as ‘IRF’ that Howe and Abedin (*ibid*) describe as ‘the paradigm participation structure’ and note its high visibility within classroom dialogue. The acronym IRE is also used for Initiate – Respond – Evaluate, although many authors consider this to be a sub-type of IRF (Howe and Abedin, 2013). In this way, VF can be expected to be a central part of teacher dialogue, however the substance and amount would be a critical part of the individual’s teaching and talk styles.

Teachers’ pedagogic styles have been a predominant focus of the dialogue literature since 2002, and there has been a shift towards more balance of teacher – student interactions, both in terms of extending response to teacher questions, and in discussing peers’ contributions. This appears to have been achieved by more ‘direct encouragement for target practices as opposed to provision of activities that are expected to draw the practices out’ (Howe and Abedin, 2013:339). Other (less well-researched) predictors of teacher variability included age, experience, expectation, gender and personality.

Howe and Abedin (2013) concluded that much classroom dialogue centred around teacher-to-student IRF, and that the teachers’ pedagogic style was critical in shaping further ‘embellishment’ of interaction, whether that be in extending the student response to the teacher initiation or increasing the opportunity for student-student responses. Much may also depend upon the learning environment being such that the student is enabled to ask questions and seek feedback as a response, and this too may depend upon the teacher’s pedagogic style.

This section has described the importance of situating verbal feedback as an ‘in the moment’ interaction where teacher and student identity, language and pedagogy collide in a complex fashion, leaving the student to untangle and make sense of the feedback in order to move their learning forward, which will be discussed in what the literature can disclose about student reception and perception of feedback.

2.3.5 Student perception of feedback

In Glover and Brown’s study (2006), undergraduate students contended strongly that they attended to feedback, even if they failed to act upon it. Where this occurred, it was more likely that the feedback had been task-focused, an inadequate or vague pointer to improvement, and to the student, in the past and consequently too late, or that it did not relate to future work. They further indicated that they found feedback from their tutors to be neither plentiful nor particularly helpful, despite the tutors themselves believing that they were providing feedback of quality. This was similarly illustrated by Carless (2006) who through a large scale questionnaire across eight universities showed that tutors believed that they were providing more detailed feedback than the students themselves believed they received, and that the nature of their feedback was more useful than the students perceived it to be. Both Carless (2006) and Handley, Price and Millar (2008) highlight confusion about the purpose of feedback from both tutor and student viewpoints, as it is often linked to grades rather than to improve learning. Teachers can have a ‘here and now’ viewpoint of feedback to give students; comment, clarification, correction or criticism that are almost summative statements (Hattie and Masters, 2012), whereas students want ‘where next?’: an indication of next steps in the form of cues or prompts. Berthold *et al.* (2007:566) describe prompts as questions or hints that ‘induce productive learning processes’ and may be useful in both scaffolding and activation.

Students can also view feedback as too vague to be of any use, too cryptic to understand what is meant by the intended comment, and inconsistent (from the same and between different teachers), negatively focused, and unrelated to the assessment criteria (Glover and Brown, 2006, Weaver 2006). Gamlem and Smith (2013) showed that students perceived corrective feedback as negative if they did not have time to work on the feedback received, and positive when they did, indicating that they could understand feedback as enhancing learning when integrated *into* the learning process. However, Nuthall (2007) showed that students may claim to understand the feedback, even when they did not, and when they did understand it, they did not know how to apply it to their future learning (also Scott, 2019). In Poulos and Mahoney’s study (2008), students were able to recognise that the feedback was ‘good’, but perceived it differently to the tutor’s intentions, thus there were issues regarding both impact and credibility. Students often privilege written feedback over verbal (Price, Handley and Millar, 2011), and may not even recognise some verbal commentaries as feedback (Yang and Carless, 2013).

For their seminal 1996 study on the effects of feedback in schools, colleges and the workplace, Kluger and DeNisi looked at more than 3000 research studies on the impact of feedback on performance, resulting in a meta-analysis yielding an average effect size of feedback on performance of 0.41. They applied some very stringent acceptance criteria (minimum sample size, existence of a control group, and calculation of size of impact found). Only 131 of the 3000 met these requirements and of these, 50 studies (38%) showed that the feedback given had actually lowered performance; that is, the control group had out-performed the experimental groups. Upon analysing the output, they determined that how an individual perceives the feedback related to the ‘gap’ [in learning] depended on the relative positions of their current performance and the goal of the learning as described below.

When the feedback indicates that the performance has exceeded the goal, the teacher may hope that the individual will increase their aspiration by raising their standards (goal) upwards. Kluger and DeNisi (*ibid*) found that the student could also respond by ‘easing off’ or even abandoning the goal or ignore the feedback as irrelevant. Where the feedback indicates that the performance has not met the goal, the teacher’s intention may be to stimulate an increase in effort to meet the goal. What may happen is that the student decides that the goal is now unrealistic, and either abandons or reduces the goal, or who again rejects the feedback as not applying to them, a finding confirmed by Dann (2015, 2018). Wiliam (2011) summarised these responses thus in Table 2.2:

When feedback indicates performance *exceeds* goal, the student:	**When feedback indicates performance *falls short* of goal, the student:**	**Student response to feedback typified as:**
exerts less effort	*increases effort*	Student changes their behaviour
increases aspiration	reduces aspiration	Student changes their goal
decides goal is too easy	decides goal is too hard	Student abandons the goal
ignores feedback	ignores feedback	Student rejects the feedback

Table 2.2 Summary of student response to feedback (*table modified from Wiliam, 2011:115)*

The effect of the feedback had depended on the nature of the feedback. Only two of these responses (*italicised*) are likely to improve performance; the other actions are likely to have no, or even a detrimental effect on performance (Wiliam, 2011). This has implications for this study, as the students’ internalisation response will have an impact on their resilience and self-belief in physics.

There is thus a large range in how feedback about ‘how students are doing’ is received and internalised. This in turn can impact on how they feel about work that provides a challenge to them individually. Generally, an increase in challenge should correspond with an increase in feedback, yet in practice, teachers may sometimes design the challenge out of lessons so that students ‘can all achieve’, fostered by a belief that success, praise and confidence are inextricably entwined.

2.3.6 Teacher Feedback as a mediator of self-regulation

> ‘*For all self-regulated activities, feedback is an inherent catalyst*.’ Butler and Winne (1995:246)

It is interesting to note that ‘feedback’ when used as Sadler (1989) described as ‘closing the gap’ is also in effect, a ‘deficit’ model to be using with individuals. Its focus, however well intended, is on ‘the gap’. In contrast, ipsative formative assessment compares existing performance with previous performance (Hughes, 2011). An example of this would be a ‘personal best’ in sport. In contrast to academic learning, many informal or practical learning experiences are assessed in this way, such as playing music or computer games. Academically, the focus is on norm-referenced or criterion-referenced judgments where little or no credit is given for how an individual may have progressed. Consequently, many children may find their feedback demotivational. ‘Ipsative assessment focuses on learner progress rather than a performance gap’ (Hughes, 2011:354), and as such is more suited to aspects of self-regulation that teachers would like to build in their learners. Hughes also notes that whilst feedback on tasks is necessary, it is feedback about learning processes and self-assessment that are more likely to elicit change.

A recurrent and serious issue with current (English) high-stakes assessment regimes is students’ ‘differential dispositions to view the testing [system and] process as a definitive statement about the sort of learner they are’ (Reay and Wiliam, 1999:343). Especially of concern are the students who do not perform well and who experience demotivation and a decreasing sense of self-worth as a consequence of what, to them, has been perceived as negative feedback. In contrast, some learners within the same system do not feel threatened by ‘negative’ feedback. They use it as an opportunity to respond to the challenge of learning: that is, they value *learning* goals over *performance* goals, and experience motivation differently.

A key area of interest in this study is that of self-regulation feedback; much of the research literature on this has been completed at the higher education level (Boud, 2015; Boud and Malloy, 2012a, 2012b, 2013; Carless, 2015; Nicol and MacFarlane-Dick, 2004, 2006a, 2006b; Vermunt and Verloop, 1999; Yang and Carless, 2013). Vermunt and Verloop (1999) define self-regulation as ‘for example, the flexible employment of different processing activities, depending on circumstances and on interim learning outcomes’ (p.262), and ‘Self-regulated learning refers to students’ skill in using a variety of learning functions and adapting this usage to the task demands at hand’ (p.276). self-regulation feedback can thus be described as comment or advice which is intended to support or enhance skills in using such learning functions, even such as helping students recognise, seek, and accept feedback. It is intended to enhance confidence and willingness to expend effort in and practise the learning. Writing from both a school and HE perspective, Elbra-Ramsay (2019:65) stated her opinion of its importance for *teachers*:

> *‘If self- regulation is the ‘pivot upon which students’ achievement turns’ and is essential to learning and outcomes (Zumbrunn, Tadlock and Roberts 2011), one questions why there is so little attention paid it to it within the standards discourse in schools and Higher Education. In addition, if as Carless and Boud (2018) suggest, the development of self-regulatory feedback behaviour ‘emerges through observation, imitation, participation and dialogue’ (p.1316) focus should be given to this within the core Teacher Standards (DfE 2012).*

Nicol and MacFarlane-Dick (2004, 2006a) reviewed literature on feedback which purported to build self-regulation through assessment. Their proposed model has become well-established in HE circles and consists of seven principles, the aim of which enabling the learner to close the performance gap, although they do not distinguish between undifferentiated and task-specific feedback (Hughes, 2011):

1. helps clarify what good performance is (goals, criteria, expected standards);
2. facilitates the development of self-assessment (reflection) in learning;
3. delivers high quality information to students about their learning;
4. encourages teacher and peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape teaching.

In a non-higher education review, Hattie and Timperley (2007) described how self-regulation is key to effective learning. When feedback is given to a student, both their ability and motivation to utilise the feedback must be activated (thus affecting their self-beliefs), else they will not be able to accept directional support in moving their learning forward. In their opinion, task feedback becomes most effective when *combined* with self-regulation feedback, so that it is instrumental, and not executive. Conversely, in a higher education setting, Glover and Brown (2006) demonstrated that this does not often occur, in a study which highlighted the dichotomy between tutors' beliefs about the (written) feedback they were providing, and the student reception of it, and the use to which it was put. However, this thesis will be exploring the verbal feedback that teachers use in the classroom with their students, and investigating the processes induced in the students by the feedback intervention (Kluger and DeNisi, 1996); we turn therefore to a more in-depth consideration of different student self-beliefs, and how their complex interplay may affect the students' reception of that feedback.

2.4 The Self-belief system

Self-belief is a system comprised of constructs: self-concept, self-efficacy and anxiety, and recently, confidence, although the operational definitions, specificity and interrelation of these has been long debated (Bong and Skaalvik 2003; Ferla *et al.*. 2009; J. Lee 2009; Morony *et al.*. 2013). There can often be a lack of clarity in the literature; some terms seem to be used interchangeably, and similar survey items used to measure different constructs. It is also necessary to have an understanding of how self-belief systems sit with current identity theory research, since 'identity as an embodied and a performed construction that is both produced by individuals and shaped by their specific structural locations.' Archer *et al.* (2010).

The physics *self-concept*, 'students' sense of themselves in relation to the subject, both currently and in the future' (Murphy and Whitelegg, 2006) has received much attention under the gender lens, in this country and abroad. Physics or science *identity* has been explored to some extent, in terms of attitude and belief and to a lesser extent on self-efficacy in (mainly US) tertiary physics education (Fencl and Scheel (2005), Shaw (2004)) and in chemistry, by Dalgety and Coll, (2006). Low self-efficacy clearly has implications for retention in the subject at any level. Cleaves (2005) identified self-concept as a powerful factor in post-16 science choices and subsequent career paths, exacerbated by limited science-related careers information and the negative school-science experience. Cleaves (*ibid*) highlighted the complex dynamics of students comparing a skill and ability self-evaluation of 'now' with a potential future self in the context of that subject. This sets up an interesting dichotomy of feeling able to 'do' physics, but not feeling able to 'be' a physicist, an outcome reported in other STEM areas as well. The definitions for the purpose of this discussion are below.

2.4.1 Self-concept

The idea of self-concept pre-dates modern psychology; Hattie (1992) noted that the foundations can be traced to Socrates and Plato in the fourth century BC, thus making it one of the oldest constructs in psychology. In 1976, following some years of uncritical findings, Shavelson *et al*. proposed a construct definition and model of self-concept, and shaped it as multidimensional and hierarchical in nature (Marsh *et al*., 2019b):

> *'a person's perception of himself ... formed through his experience with his environment ... and influenced especially by environmental reinforcements and significant others'* (Shavelson *et al*., 1976:411).

This multidimensional model of self-concept Shavelson *et al*., (1976) comprised 17 different conceptual features, and can be seen in Figure 2.7a. This model consisted of general self-concept (then also called 'self-esteem') as the pinnacle, above second-order layers of academic self-concept, and non-academic self-concept. This suggested compartmentalised ways of thinking about the self in terms of academic subject, peers, emotional state, *etc*. This was subsequently adapted by Marsh and Shavelson (1985) through the adoption of validated measurement instruments of self-concept, and the Marsh-Shavelson model is shown in Figure 2.7b. A critical feature of both models is the domain specificity, particularly in the *academic* self-concept, relating to (e.g.) subject areas.

Estimated to begin in early childhood from an age range of 4-7 (Marsh *et al*, 2019b), self-concept is not a fixed trait, and will change over time; Marsh *et al* (2005) debated the difficulties of obtaining accurate results with young children due to a lack of an appropriately modified measurement instrument, and the potential incapacity for the children to reliably convey their self-perceptions. The formation of self-concept embodies:

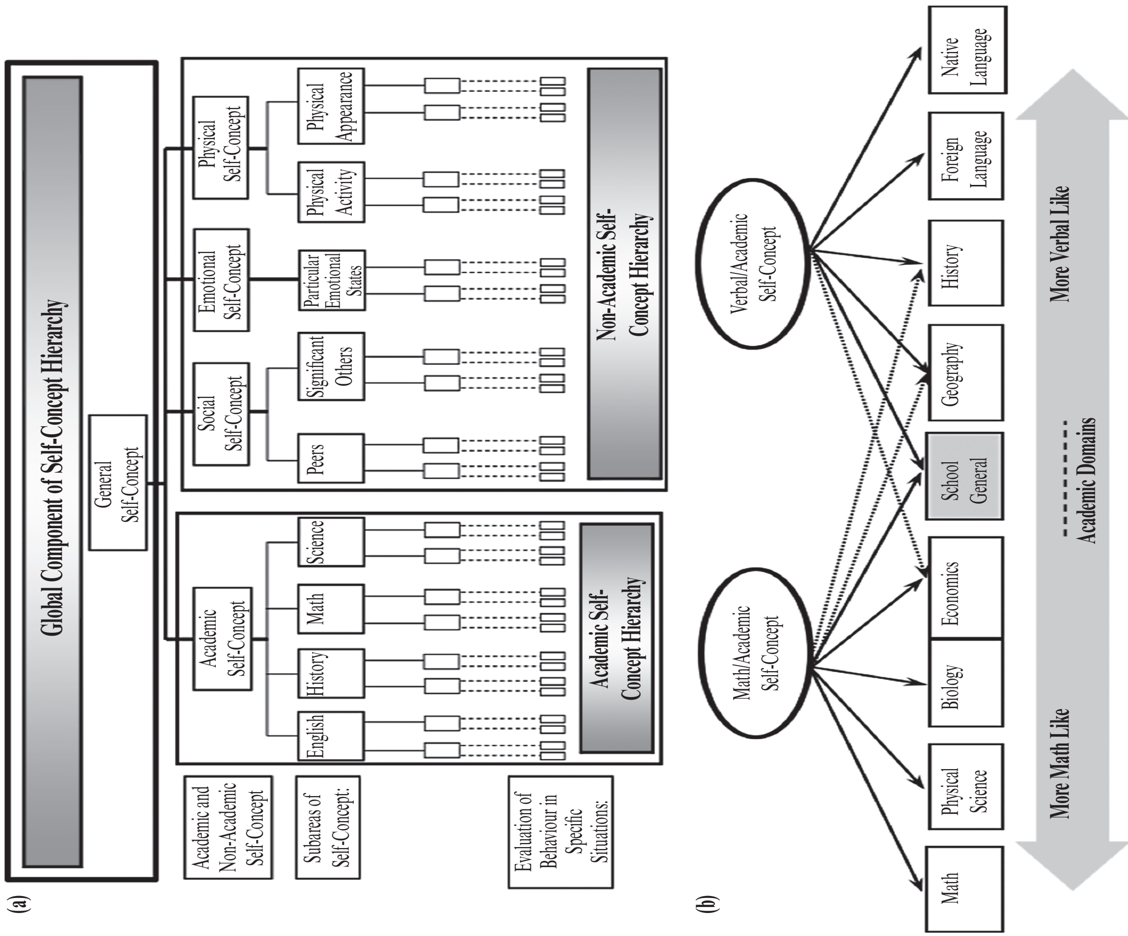


Figure 2.7 The structure of academic self-concept. (a) The original Shavelson Model. (b) The Marsh/Shavelson revision of the academic component of the self-concept structure. Adapted from Self-Concept Theory, Measurement, and Research into Practice: the Role of Self-Concept in Educational Psychology, by H. W. Marsh, 2007 (Marsh *et al*, 2019b).

> *'a basic psychological need that has a pervasive impact on daily life, cognition and behaviour, across age and culture ... an ideal cornerstone on which to rest the achievement motivation literature but also a foundational building block for any theory of personality, development and well-being'* (Elliot and Dweck, 2005, p. 8)

The notion of self-perception has evolved over time, with researchers recognising the role of social processes in its construction and modification, as in Shavelson *et al.'*s definition above (1976). Arguably, self-concept and behaviour are mutually influencing factors; how one views oneself is a function of one's interactions with others (Marsh *et al*., 2019) in the sense that 'others are a mirror through which one catches glimpses of who and what one is' (p.38). This agrees with Tice and Wallace (2003:91) who define 'the reflected self' as 'the idea that people come to see themselves as they believe others to see them', and in turn, 'their perceptions of how they are viewed are influenced by their own prior beliefs about themselves' (p.103). As Dweck says: 'Our language tells people what we value' (Talk, York, 2010). This has implications for this study as both peer and teacher appraisal, as well as language *as* classroom dialogue can affect this self-image.

Marsh *et al.* (2009) describe a self-concept as consisting 'of cognitive appraisals and judgements that individuals attribute to themselves' (p.320), and may operate in a general sense, or to specific facets of the individual, and thus a typical survey item is '*I get good grades in maths*', based on a Likert-type response. Since measures of self-concept are often based on an individual's perception of 'self' compared to his or her peers, and may vary widely between domains, a single academic self-concept is thus insufficient to cover the whole school province (Marsh *et al.*, 2009, Marsh *et al*., 2019a).

A correlation between self-concept and academic achievement was first established in 1982 by Hansford and Hattie, supported by a study by Marsh *et al*., (2006) in a large German secondary level study (n= 4,475, 45% males), in which student self-concepts in mathematics, English and German were substantially positively related to their outcomes. This introduced the question of whether a causal link existed between the two; Marsh (1990) had argued for a dynamic reciprocal effects model incorporating both skills development and self-enhancement, which has subsequently been supported by several studies (Chen *et al*., 2013; Huang, 2011; Pinxten *et al*., 2010), and has implications for this thesis.

Craven *et al*. (1991) noted (at the time) that self-concept researchers generally accepted that teacher feedback resulted in positive outcomes for self-concept, and instead proposed an 'internal mediating model' whereby the student moderated feedback on an activity into generalised feelings towards a subject. Burnett and Mandel's (2010) exploration of teachers' and students' perceptions of praise and feedback supported some aspects of this, however making a distinction between ability and effort feedback, and noting that more effort feedback should be utilised with increasing age of the student. A meta-study (O'Mara *et al*., 2006) on interventions aimed at enhancing self-concept found that the largest effect-sizes were for feedback and praise interventions ($d=1.13$), higher than for other inventions such as cooperative learning, group and individual counselling, skills training and peer-tutoring. The researchers do not define praise within this work; however, given the outcome, they conducted additional analyses in which the praise/feedback interventions were allocated into contingent praise, noncontingent praise, attribution feedback and goal feedback groups. Again, the nature of the praise was not defined, but the authors found that interventions administering noncontingent praise were less effective than the other sub-types. They linked this to similar findings by Mueller and Dweck, 1998 and Kamins and Dweck, 1999, who do make more formal explanations of different types of praise; the effect of noncontingent praise as person- or trait-related, or more generic forms was asserted to be unhelpful to learning, and potentially detrimental to self-worth. Next we turn to a self-belief construct whose difference from self-concept has long been the subject of debate; self-efficacy.

2.4.2 Self-efficacy

Self-efficacy was first introduced by Bandura (1977a, 1977b) as ‘an individual’s subjective conviction in her or his capabilities to perform a specific task successfully to achieve a desired outcome’; he later defined self-efficacy as the belief ‘in one’s capabilities to organise and execute the course of action required to produce given attainments’ (1997:3). Marsh, Hau, and Artelt (2009:320) added ‘particularly in one’s ability to overcome barriers through one’s own efforts’. Expectations of an individual’s self-efficacy are said by Bandura (1997) to be derived and internalised from four sources of information, in order of strongest influence first:

- ‘personal mastery experiences’, where one’s past achievements, whether successful or unsuccessful, will strengthen or weaken the belief accordingly (Britner and Pajares, 2006; Usher and Pajares, 2008; Schunk and Usher, 2013);
- ‘vicarious learning experiences’, acquired through observing others (Bandura, 1977a), such as the teacher modelling activities in the classroom;
- social persuasion experiences’, where positive feedback from significant others augments the efficacy beliefs of students (Bandura, 1997; Zeldin and Pajares, 2000), while little or negative feedback under- mines their efficacy beliefs (Seon and Bong, 2019);
- ‘physiological state’ such as pulse rate, nervousness, deemed by Bandura (1977a, 1977b) to be the weakest source of information.

It is also context dependent in that an individual may have a high self-efficacy for the performance of a task in one context, but a lower one for the same task performed in a different context, for example, in private, small groups or publicly. Self-efficacy measures tend to be domain-specific (Bandura 2006) and typical survey items tend to ascertain the confidence of being able to perform discrete tasks: ‘I can calculate the mean’, ‘I can solve equations containing square roots’.

Seon and Bong (2019:63) highlight the functional advantage of possessing strong self-efficacy:

> *'self- efficacy plays a critical role in academic contexts, predicting choice of activities and courses, interest, persistence, effort expenditure, use of learning strategies, self-regulation, and eventual achievement of students' (Bandura and Schunk, 1981; Pajares, 1996, 1997; Schunk, 1995).*

Bandura's (1997) social cognitive theory offered a model of 'triadic reciprocity' in which personal, behavioural and environmental factors interact to explain human behaviour (Seon and Bong, 2019). Seon and Bong hail self-efficacy as a critical exponent in *agency* (Bandura, 1997, 2018), a cornerstone of social cognitive theory and define agentic perspective as referring:

> *'to the view that people are self-organizing, proactive, self- reflective, and self-regulating toward their own development, adaptation, and change in given situations. In social cognitive theory, agentic perspective includes four core properties of human agency, namely intentionality, forethought, self-reactiveness, and self-reflectiveness'* (p. 64)

Self-efficacy is one of the primary dimensions of students' overall science identity and contributes to their persistence in physics (Hazari *et al.*, 2010, Sawtelle, 2011). In a gender study in Physics, Sawtelle (2011) made the distinction that: 'self-efficacy is an important component of success for all students in introductory physics [tertiary level], but that the particular way self-efficacy develops varies with gender' (p.13). She advocated that should the intention be to increase retention rates of women in Physics, then an arena of opportunities to attend to self-efficacy through the mode of instruction should be modelled in the classroom, to improve cognitive judgements of one's own capabilities.

Kunkel (2012) describes self-efficacy as a sense of personal competence rather than a comparison to others, and this makes it distinct from how self-concept operates. Flammer (1995:78) asserts that 'there is evidence that a higher sense of self-efficacy supports motivation, even when the efficacy is an overestimation.... but there are dangers in continually overestimating performance as well'. Chan and Lam (2010) demonstrated that students receiving summative feedback showed a decrease in self-efficacy compared to those receiving formative feedback, and that self-referenced feedback was more beneficial to self-efficacy than norm-referenced feedback (see also Arslan, 2012). Since many studies have found that mastery experience is a powerful predictor of strong self-efficacy (Lopez and Lent, 1992; Britner and Pajares, 2006; Joët *et al*., 2011) and that self-efficacy is a powerful predictor of academic achievement (Bandura, 1997; Alivernini and Lucidi, 2011; Joët *et al*., 2011; see also Honicke and Broadbent, 2016 for a systematic review), Kunkel (*ibid*) suggests the following strategies to help regulate this:

- Emphasize past success and review progress
- Set learning goals with your students, and model goal-achieving behaviour
- Make specific suggestions for improvement and revise grades accordingly
- Stress the connection between past efforts and past accomplishments

Finally, self-efficacy has also been shown to negatively correlate with anxiety (e.g., Bong *et al*., 2012) and this is considered below.

2.4.3 Anxiety

Anxiety can be described as one's physio-emotional reactions when she/he thinks about performing a task, such as 'I get very nervous doing mathematical problems' (Lee, J. 2009; Stankov *et al.* 2012), and thus is also regarded as a domain-specific construct. Morony *et al.* (2013) offer several reasons for the negative relationship between anxiety and the other two self-belief constructs. They suggest that students try to avoid such feelings by avoiding the subject arousing it (Moeller *et al*, 2014), and/or find it less enjoyable: both actions will lead to a failure to master it. Additionally, struggling with a subject can 'use up' working memory capacity, reduce available resources, and consequently lead to impaired performance.

Mallow (2006) who claims to have coined the term 'science anxiety' in 1977, described it resulting from several contributing factors such as negative science ability messages received from teachers in earlier education, a lack of role models, a lack of physics specialists teaching physics and a lack of training in analytical methods. This has resulted in science (especially physics) being perceived as a ‘difficult’ subject to learn, ultimately hindering their success in science (Mallow, 2006; Ucak and Say, 2019). Other studies suggest that anxiety and negative attitudes both hinder student learning, and ultimately reduce performance and attainment (Osborne, Simon and Collins, 2003; Udo, Ramsey and Mallow, 2004). In an American middle school study, Atwater, Gardner and Wiggins (1995) showed that students with *low* levels of science anxiety and stress had higher achievement levels and positive attitudes towards science. With respect to physics, Gungor, Eryilmaz and Fakioglue (2007) found that physics anxiety affected their achievement (see also Sahin, 2014). In 2015, Sahin, Caliskan and Dilek developed the Physics Anxiety Rating Scale [PARS] and reported that physics anxiety was higher in females than in males (n=495), supported by Agra, Fischer and Beilock (2017), contradicting earlier findings by Brownlow, Jacobi, and Rogers, (2000) of similar amounts of physics anxiety in a sample of American university students.

Udo, Ramsey and Mallow (2004) offered their thoughts on reducing science anxiety through firstly science skills learning, which notably saw a difference in females coming to recognise aspects of their learning was dependent on their ‘hands-on’ experience in science; and secondly, cognitive restructuring, an interesting and insightful notion based on Ellis (1962, cited *op. cit.*) that entities such as laboratory equipment or physics problems are not in themselves intrinsically anxiety-provoking. They are however:

> *the stimuli for negative self-statements which students tell themselves (usually unconsciously), such as, "No matter how hard I study, I'll never understand science;" "Science is not for girls;" "Everyone understands it but me." These self-statements are the real stimuli provoking the anxiety responses’* (Udo, Ramsey and Mallow, 2004:443).

They position cognitive restructuring as a technique for getting students to acknowledge their negative self-statements, recognise the emotions and beliefs which underpin them, and then replace them with ‘objective, emotionally neutral coping statements’ (Udo, Ramsey and Mallow, 2004).

More recent studies on science anxiety have shown that secondary school students’ anxiety levels vary significantly depending on a range of variables, but unlike Udo, Ramsey and Mallow (2004) above, gender was not one of these. Both Avci and Kirbaclar (2017) and Ucak and Say (2019), two secondary-level studies comprising large sample sizes, show that whilst liking for science, and a liking for their science teachers did have a significant effect on lowering anxiety, there was no significant difference between students’ anxiety scores in terms of gender. It could be hypothesised that there may be both cultural and generational differences between the results of these studies. Ucak and Say (2019) also identified an incremental decrease in anxiety with an increase in positive grade outcomes for their students across five different levels of grades. When the negatively correlating relationship suggested between anxiety and self-concept described above is considered, this appears to be supporting evidence.

Lastly, Udo, Ramsey and Mallow (2004) offer some advice on interventions which might combat anxiety development, based on the American Association of Physics Teachers (AAPT) workshop for teachers, Developing Student Confidence in Physics by Fuller *et al*., 1985, cited *op. cit*.). These include:

a) the classroom learning environment
b) information transfer between teacher and student
c) teacher-student interactions
d) teachers’ evaluations of student performance, including how commentary on work may diminish or enhance student confidence.

2.4.4 Conceptual and methodological distinctions between self-belief constructs

Are these separate constructs? It has been postulated that researchers may not have been able to operationalise definitions of these self-constructs sufficiently, or that there may be some underlying factor to all of them, which is not being directly observed (Bong 1996, Bong and Clark, 1999). In early research, Pajares (1996) suggested that the different self-constructs may share overlapping sub-dimensions; for example, the 'cognitive dimension of self-evaluation' is utilised in forming both a self-concept (judgement of oneself as a person), and self-efficacy (judgement of how one will perform at a task). Furthermore, are these various self-constructs more or less likely to be distinguishable when investigated within subject-specific domains, rather than in a generalised sense?

This is the background that Lee (2009) sought to explore, from a mathematics-domain basis, to ascertain whether they existed as discrete constructs, and if so, could this be demonstrated across different countries and cultures? He used the data from PISA 2003, which had a mathematics theme, to identify whether (mathematics) self-concept, (mathematics) self-efficacy and (mathematics) anxiety were indeed distinguishable empirically, as theoretically differentiated previously. Although these earlier findings had suggested that these concepts were too closely conceptually related, Lee showed them to be separate constructs both across and within the 41 countries supplying the data. He did concede their close relation to each other, but part of his conclusion supported the argument [of other papers] that maths self-efficacy is a better predictor of academic performance than maths self- concept. These findings are supported by Ferla *et al.* (2009) who showed that self-concept was a better predictor for more motivational variables, and self-efficacy was the better predictor for academic achievement. Additionally, Ahmed *et al*., (2012) have noted the negative correlation between anxiety and self-concept.

Seon and Bong (2019) summarise the differences between self-concept and self-efficacy, and this is shown in Table 2.3 (see also Marsh *et al*., (2019a). Seon and Bong (ibid) also highlight that self-concept relies more upon the individual's past success or failures than does self-efficacy, thus making it the more stable construct and resistant to change. They assert however (p.65) that 'self-efficacy is relatively more malleable compared to self-concept or self-esteem, which renders itself a desirable target of instructional interventions', an important issue for this study.

Comparison dimension	Self-esteem	Self-concept	Self-efficacy
Conceptual definition	Evaluative judgments of oneself, which include one's feelings of and satisfaction toward oneself	Knowledge and perceptions about one's competencies and attributes, along with resultant emotional reactions	Subjective convictions for successfully executing a course of action to achieve a desired outcome
Judgment specificity	Global	Domain-specific	Domain- and context-specific
Dominant reference point	Past experiences	Past experiences	Future possibilities
Temporal stability	Stable	Stable	Malleable
	Academic self-esteem	Academic self-concept	Academic self-efficacy
Conceptual definition	Evaluative judgments of oneself in achievement situations, which include one's feelings of and satisfaction toward oneself	Knowledge and perceptions about one's competencies and attributes in achievement situations, along with resultant emotional reactions	Subjective convictions for successfully performing given academic tasks to a desired level
Levels of formation	Overall judgments formed at the global levels of functioning	Formed at the global as well as domain- and subject-specific levels	Formed in reference to specific domains, subjects, or tasks
Relevant constructs	Academic self-esteem	Academic self-concept, subject-specific self-concept	Academic self-efficacy, subject-specific self-efficacy, task-specific self-efficacy
Example constructs	Academic self-esteem	Academic self-concept, English self-concept, mathematics self-concept, etc.	Academic self-efficacy, mathematics self-efficacy, self-efficacy for solving subtraction problems, etc.
Example statements	I like myself in school I am satisfied with the way I am at school	I have always done well in English I am a good student in mathematics	I am confident that I can receive a grade of B or better in English I am confident that I can perform well in mathematics I am confident that I can successfully solve the subtraction problems

Table 2.3 Comparison between the self-constructs in achievement situations (Seon and Bong 2019)

2.4.5 Confidence

There is an emerging argument for including ‘confidence’ as another, discrete, construct within the self-belief system. Stankov (1999) initially positioned ‘confidence’ as a trait in the ‘no man’s land’ between cognitive abilities and personality, and pursued research in this area, describing its conceptual links to several areas of educational psychology: studies in metacognition (amount of knowledge); assessing metacognitive monitoring processes; assessing the ‘feeling of knowledge’ (FOK) paradigm and it is frequently used in studies of self-efficacy (Pajares, 1997). However, writing with other authors, Stankov later argued for confidence as a separate self-belief factor that encompassed the three prior constructs (Stankov, Lee, Luo and Hogan, 2012). They highlight its position as a psychological variable in behaviour and extend its strong relationship to achievement, claiming it to be an important predictor of achievement.

The main conceptual distinction between them arises from the domain-specificity of self-concept, self-efficacy and anxiety as subject-situated (Seon and Bong, 2019), unlike the construct ‘confidence’ which is presented as much broader in scope. Stankov *et al.* (2012) claim that their findings demonstrate confidence as a trait, set in the context of describing the level of uncertainty about the accuracy of an item response. They assert it encapsulates cognitive aspects (for example, high vs. low ability, easy vs. difficult test), personality aspects, and motivation towards the task. Kröner and Biermann (2007) showed confidence to have a moderate correlation with self-concept and concluded that people may use their self-concept construct as an indicator for their confidence level if adequate task-located cues are not available.

For the purpose of this study, confidence cannot be sufficiently operationalised to monitor, remaining as it does a non-specific term generally describing the strength of belief. Should the focus of the study include measures of achievement, then confidence could better be operationalised as a construct with the other self-beliefs, however, it does not.

2.4.6. The relation between Self-beliefs Constructs and Identity

Wenger (1998) states that theories of identity *'are concerned with the social formation of person….They address issues of gender, class, ethnicity, age, and other forms of categorization, association, and differentiation in an attempt to understand the person as formed through complex relations of mutual constitution between individuals and groups'* (p.13). Identity theory is most concerned with social networks that involve interpersonal relationships, and how such relationships (and perception of them) shape role-related identities in an on-going process. Merolla *et al.* (2013) state that 'virtually all persons in modern differentiated societies have multiple identities organised in a hierarchical manner' (p. 151) and describe *identity salience* as the relative positioning of a particular identity within this hierarchy, after Stryker and Serpe, (1982); see also Brenner, Serpe and Stryker, (2014:232) who state:

> 'An identity's salience indicates its relative position in a hierarchy of salience ranked by its propensity of being called up: identity theory predicts that a highly salient identity is likely to be enacted or to define a situation to promote its own enactment.'

Thus, an identity higher in the hierarchy has more *relevance* to the individual and is more likely to be called upon.

In her doctoral thesis, Li (2011) suggests that 'identity is the way we know how to be a member in a specific community' and asserts that 'it is necessary to be specific about what kind of identity because it is context-specific' (p.32). As secondary school is a crucial time for identity construction, it is necessary to explore what identity means, and how it may affect educational choices, since it is in continual development. Schreiner and Sjøberg (2007) argue that a more important question for today's youth is not 'what do you want to be when you grow up?' but rather, 'who do you want to be when you grow up?'

Provenances of influencing elements are multiple, and complex (such as gender, curricula, teachers, culture), and often interactional. Shanahan (2009) positions the communities of practice (such as the classroom) as the dominant theoretical framework for identity studies, against which to examine the three levels of identity analysis considered by researchers: personality, interaction and social structure, and argues that many science identity studies have 'focused their attention on aspects of identity related to individual agency to the exclusion of issues of social structure'. In their 2006 review of the research on the participation of girls in physics, Murphy and Whitelegg drew attention to the way in which social identities in relation to science have come to influence the way in which the teaching and learning process is understood, and station 'the physics classroom' as a community of learners.

Boaler (2002) offers a situated perspective on learning: 'The students who were learning in these traditional [maths] classrooms were generally successful, but we found that many students experienced an important conflict between the practices in which they engaged, and their developing identities as people' (p.5). In her study of contrasting classrooms, Boaler (*ibid*) found that there were differences in the student identities that developed in relation to mathematics, depending on the type of pedagogical approach employed, and the learning environment it created.

In an online survey of [1881] undergraduate students in STEM subjects at university, differences in science identity and self-efficacy were studied by STEM speciality as well as gender (Williams and George-Jackson, 2014). The results indicate that science identity is positively impacted by students using and doing science, rather than by their self-efficacy, but also that there are differences between male and females in science identity and their perceived self-efficacy. Their survey consisted of 14 questions relating to science identity, with two distinct factors of 'Identifying as a Scientist' and 'Using and Doing Science', and 17 self-efficacy survey items. However, the latter included a number of questions such as 'I feel good about myself' and 'I feel I do not have much to be proud of', which appear to relate more to the construct of self-concept than self-efficacy, and perhaps there has not been sufficient differentiation in operationalizing these constructs, as suggested by Bong (1996). Whilst self-belief constructs are important contributors to identity, and not least, beliefs in physics a contributor to a physics identity, the concept of identity is not the focus of the research questions in this study, although there may be implications for future identity studies.

2.4.7 Attribution as a contributor to Self-beliefs

Dweck's early work involved studies of student 'learned helplessness' and their own beliefs about their academic abilities. Her work moved somewhat into the social realm as she became intrigued by why some people functioned well, and others not so well from the same circumstantial stimuli. In summarising research studies spanning three decades, with students ranging in age from pre-school to undergraduate level, Dweck and her colleagues (2000) found three themes running through students' own responses to their successes and failures in response to simple questions such as 'when you get an A grade, why is that?' or 'if you got an F grade, why do you think that is?':

1. that success or failure was due to internal or external factors, e.g. how it was *personalised*. (e.g., *internal*: I got a good grade because it was a good piece of work; *external*: I got a good grade because the teacher likes me)

2. that success or failure was due to factors that were transient or long-lasting, i.e. the *stability* (e.g., in the case of failure, *stable*: I got a bad grade because I'm no good at that subject; *unstable*: I got a bad grade because I hadn't reviewed the material before the test)

3. that success or failure was due to the *specificity* of the attribution (I'm good at that, but that's the only thing I'm good at) or how some students over generalise their success or failure = *global*: I'm good at that, so I'll be good at everything).

The best learners consistently attribute both success and failure to internal unstable causes: 'it's up to me, and I can do something about it' (Wiliam, 2011). Attribution theory describes how people make sense of their world, and in particular, to what they ascribe their successes and failures, pioneered by Weiner (1984).

An example of the negative version of these attributes is described in the 'Bright Girls' Helpless Response' first documented by Licht and Dweck (1984), in which very able girls were debilitated by failure, and subsequently disengaged. This does not mean however that any student, not just the 'Bright Girls', is doomed to debilitation. Dweck attempted 'success training' in 1975, and there have been many attribution-retraining studies since. In Dweck's 1975 study, half of a group of students identified as exhibiting an extreme helpless response to failure were subjected to training in how to interpret their failures, called attribution (re)training, as they were being taught a new explanation for failure (rather than their self-perceived lack of ability). Feedback was given in terms of the effort that had been expended, and during the 25 sessions, their decline in performance when failure occurred decreased. By the end of the testing period, the improvement was noticeable to the point of some students showing a better performance after failure than before, having apparently independently used the same feedback on themselves. Interestingly, their teachers (who did not know who had been in which groups) were able to perceive differences in attitude in the classroom in general.

Yeager and Dweck (2012) reviewed research detailing the impact of students' mindset on their resilience in the face of academic and social challenges. They assert that often, mindset-type is not exhibited until challenges are encountered, such as difficult school transitions, though the challenge does not have to be severe at all. They showed that students who believe – or who come to believe through 'attribution training' – that intellectual or personal characteristics are qualities that can be developed, as opposed to qualities that are fixed, tended to show higher achievement across such transitions, as well as (in this review) higher achievement in a challenging maths course.

Resilience has become a byword in UK education (Sterling, 2010). It refers to the 'capacity of individuals to prosper despite encountering adverse circumstances' (Agasisti *et al*., 2018:4) and perhaps even more strongly that 'being resilient is the 'capacity to endure ongoing hardship in every conceivable way' (Walker, Gleaves and Grey 2006:251). The effect is that similarly to feedback, resilience can be regarded as a deficit model in which the individual is seen as lacking it; ‘not thriving within this climate is therefore the fault of the individual, not an indicator of the climate they operate in’ (Elbra-Ramsay, 2019). This has implications for this study since some feedback literature indicates that feedback can make the recipient feel emotional (Daniels *et al*., 2009; Eva *et al*. 2012; Pekrun *et al*. 2002), even vulnerable (Elbra-Ramsay, 2019).

Further, varied examples of attribution training include weekly mentoring emails based on incremental theory (Good *et al*.. 2003), writing ‘pen pal’ letters of incremental theory-based support (Aronson *et al*., 2002), study skill intervention (Blackwell *et al.,* 2007), and resilience training in middle-school maths for girls facing stereotype threat (Spencer *et al*.,1999) all claim to have shown convincing evidence that mindset can be changed, and that it is not a transitory effect (Yeager and Dweck, 2012, Dweck and Yeager, 2019). A contrasting response can be seen in Rattan *et al*., (2012), in which comfort praise from the (fixed) mindset of maths instructors served to reinforce students’ (and in particular, female students) self-concept that they were no good at maths and that it gave them permission to mentally ‘walk away’ from the subject; surely the antithesis of self-regulation in learning ideals. Consideration of attribution theory has a research design implication for the construction of the professional development intervention as well as the langue deployed by the teacher during classroom discourse.

2.4.8 Is ‘Mindset’ a self-belief construct?

Dweck (2000, 2006, 2017) has been at the forefront of establishing a modern framework for understanding intelligence and achievement. She has proposed contrasting self-theories that people hold about their own ‘intelligence’, which gives valuable insight into learner motivation. The definition of intelligence is a much debated one, from measures of pure intelligence factoring out all personality and motivational factors, to ones that include them, and regard them as integral parts of intelligence. There are cultural differences also in the way that it can be viewed, for example Asian cultures believing that effort plays a large part in intelligence, whereas, for example, American culture is less likely to (*ibid*).

How *individuals* themselves view intelligence has been the focus of extensive research by Dweck. ‘Entity theorists’ view intelligence as fixed; an innate ability that one is born with and cannot change, or not very much. These individuals tend to view any task as an opportunity to either reaffirm their ability or to fail (there is no in-between). They tend to be goal-oriented and interpret failures as indicative of lack of (natural) ability. Consequently, they may give up when faced with a challenge, even if they have previously enjoyed successes. This can lead to increased helplessness and a desire to avoid ‘difficult’ subjects (maths and science are often cited). They may even reject support and can interpret critical feedback as ‘evidence’ of their lack of ability. They often define intelligence as ‘inborn ability’, being ‘very smart’, ‘brilliant’ or ‘bright’ (Dweck, 2000:61).

‘Incremental theorists’ are more likely to define intelligence in terms of knowledge, skills and effort and think that motivation is a key part of it: ‘studying hard’, ‘to try your best’ (Dweck, 2000:62). They think that intelligence is developmental and that it can be grown. To them, failure is an opportunity to learn how to do something better, rather than evidence of low ability, and have learning goals rather than performance goals. Whereas entity theorists want to look good at what they do, and to be seen to be achieving it, incremental theorists are mastery-oriented and relish the process, even when it is ‘difficult’. There has been a widespread uptake of the notion within schools that such a ‘growth mindset’ approach can be used to encourage learners.

In more recent publications, Dweck (2017) has sought to clarify that there is not a ‘state of mind’ called growth mindset; rather it is a way of thinking in a certain circumstance (DeWitt, 2017). Clearly individuals holding each of these views, even in different subjects, or aspects of their lives will respond differently to feedback relating to success and failure in different circumstances, with consequent impact on the development of their self-beliefs regarding both ability and enjoyment:

> *‘Assessment readily perpetuates a myth that ability is innate for entity theorists, if feedback is strongly task outcome orientated, and these outcomes are grades in comparison with others. It is easy to interpret critical feedback as evidence for an inability to meet the goals, when performance on task is relatively low and the process of learning is undervalued.’* (Hughes, 2011:358).

If entity theory students are confident in their ability to achieve within the task, they will attempt it. If not, they very often ‘can’t be bothered’ and disengage. Feedback in these situations is then often viewed as rubbing salt in the wounds. This therefore has implications for shifting the *focus* of feedback. Whilst reporting in grades is here to stay, there remain many opportunities for low-stakes formative assessments where the process of learning is valued and reported back on, rather than the product, and in that process, build more self-regulated learners.

The current mindset discourse has however become one of contention. Amid widespread adoption by schools in both the UK and USA (EEF, 2015, 2019; Dweck and Yaeger, 2019), it is important to note that some studies do not fully support Dweck’s theories, and others contradict it. In a study of 182 Irish university-level students, O’Shea, Cleary and Breen (2010) showed weaker correlations between the theory of intelligence [TOI] and the goal orientation that Dweck’s theory would predict, and no evidence that TOI beliefs are related to gender. Dupreyrat and Mariné (2005) asserted that TOI did not seem to influence goal orientation, but findings were otherwise consistent with Dweck’s theoretical predictions, in a study of adult learners of age range 20-49 of mean age 31 years. Carmichael and Taylor (2005) studied a group of tertiary students (n=129, median age 29) and found that most subscribed to an incremental theory of intelligence. They concluded that the issue of TOI is perhaps not as relevant in adult learners, as it may be in younger learners, such as those that Dweck and her collaborators tend to study. O’Shea *et al.* do concede that their correspondents may not want to label themselves as other than desiring learning goals, and it is possible that their maturity and/or self-selection onto such tertiary courses has affected the sample.

Lynch (2018) criticised the disparate nature of Dweck and colleagues’ investigations, the increasingly small effects reported over time form the original, influential Mueller and Dweck study in 1998, and highlighted serious methodological and statistical issues. Brown (2017) challenged inconsistencies in findings data published by Mueller and Dweck (1998), to which the authors responded acknowledging some errors in the reporting of results. Other authors have tried to replicate Dweck's findings in several 'mindset' studies without success and in a three-study trial on praise intervention, challenge of material, and analysis of links between mindset and education attainment, Li and Bates (2017) found no support for mindset effects on the above conditions. These were, however, short-term interventions.

In 2012, Burnette *et al*. undertook a meta-study of 113 research projects, finding that incremental theories [of participants] significantly predicted goal setting, learning goals and mastery-oriented strategies but described these effects as weak, and the correlation small. They did not support Dweck and colleagues' claims that growth mindset interventions have a profound effect on academic achievement. The claims were further contradicted by two meta-analyses by Sisk *et al.* (2018), which found that correlation of growth mindsets with achievement was weak (correlation coefficient r =.1) and that the effect of growth mindset interventions on academic achievement was very small (Cohen's d = 0.08); they did concede that some academically at-risk students might benefit. In England, the Education Endowment Fund conducted two randomised control mindset efficacy trials, the first in 2013 with 36 schools and 1505 primary pupils, who showed an additional two months' progress in English, but no additional progress in any other measure (EEF, 2015). In 2015, a second trial was conducted with 101 primary schools and 5018 pupils and reported a finding of high security as no additional gains measured on academic outcomes, or an examination of four measures of non-cognitive skills: intrinsic value, self-efficacy, test anxiety, and self-regulation. (EEF, 2019). The evaluation report concluded that 'teachers should be cautious about using growth mindsets alone as a way of boosting pupil attainment' (EEF website summary, 2019).

In response to these refutations, Dweck has made several points. In 2016, she opened dialogue on a 'false growth mindset', a term apparently covering both those who say they possess a growth mindset when they do not, and those who do not understand what it really is:

> *'Many people understood growth mindset deeply and implemented it in a very sophisticated and effective way. However, there were many others who understood it in a way that wasn't quite accurate or distilled it down to something that wasn't quite effective, or assimilated it into something they already knew'* (Dweck, 2016).

Additionally, Dweck asserts that there are subtleties in its implementation, and denies that failure to replicate Dweck and colleagues' studies is an indication that the results are insecure, arguing attempts to replicate can fail because the scientists had not created the right conditions:

> *'Not anyone can do a replication. We put so much thought into creating an environment; we spend hours and days on each question, on creating a context in which the phenomenon could plausibly emerge. Replication is very important, but they have to be genuine replications and thoughtful replications done by skilled people' (Dweck, cited in Chivers, 2017).*

Brown (2017), who challenged the Mueller and Dweck (1998) study has signalled scepticism at this, asking 'If your effect is so fragile that it can only be reproduced [under strictly controlled conditions], then why do you think it can be reproduced by schoolteachers?' (Chivers, 2017). This argument suggests several areas for caution during the study; that mindset interventions may not have the impact previously claimed, students (and their teachers) may present 'false growth mindset', and/or the previously validated mindset questions included in the survey (section 3.8.1 and Appendix 2) may not meet the context and environment required by Dweck above.

2.5 Chapter Summary

The literature review has drawn together a synopsis of literature on teacher professional development, and highlighted how feedback is a complex concept, despite the inclination of the education system to treat it as a unidimensional notion. The concept of student self-beliefs has been presented as a system in which constituent parts can be distinguished and operationalised to gather student data. There has been a deeper consideration of how apposite studies have revealed gaps which form a research niche, in terms of lack of study of feedback interventions on changing student self-beliefs; few studies which show how feedback can modify and strengthen different self-belief constructs; how written feedback is preferenced to verbal, yet verbal feedback is an integral part of classroom discourse, but considered by some authors to be low frequency within other classroom dialogue; how the student perspective on feedback is not often considered; and how self-belief studies tend to have narrow foci on one or two constructs, and do not necessarily map self-belief change over time.

The chapter has thus explored studies which have influenced the research design with regard to the design of the professional development intervention, and the choice of feedback model and typology which will yield useful classroom feedback observations; since such typologies and model tend to sit behind institutional academic paywalls, teachers are often unaware of such research, yet gaining knowledge of them may impact their feedback repertoire, which may in turn impact on their students' self-belief systems. The thesis will now explore the decisions framing the research design in more detail, so that the research questions may be addressed, and subsequently data analysed and discussed with reference to the literature outlined in this chapter.

Chapter 3 Research Design

3.1 Overview

The investigation took the form of a quasi-experimental semi-longitudinal study of four classes of Y10 Physics students, of differing abilities, in two North Yorkshire secondary schools in the United Kingdom. The study involved pre-intervention and post-intervention measures, complemented with semi-structured interviews with selected students, and on-going observations of classroom feedback interactions. The intervention was a Continuing Professional Development [CPD] trajectory (Joyce and Showers, 2002; Gabelica *et al*., 2012) designed to inform teachers about different types of feedback and use a coaching approach to encourage the teachers to use different forms of feedback with their students. The mixed-method approach for data collection enabled triangulation and corroboration. There is an additional case study, presented as a descriptive narrative that will document the intervention teachers' response to the process of intended change within their feedback typology in their physics lessons. A survey tool was used pre-and post-intervention to yield quantitative aspects of students' beliefs about physics, and their belief in their ability to learn physics. The survey was also be undertaken by a similar number of comparison classes in each institution to strengthen internal validity. The pre-intervention survey provided baseline information from which defined subsets of the treatment classes were followed in greater focus. Qualitative methods of pre-, during, and post-intervention teacher interviews, as well as pre- and post-intervention student focus groups and a continuous schedule of classroom teacher observations were included to enable both a fine-grained analysis of any change in students' perceptions and beliefs, as well as changes in the teachers' feedback styles. A schematic overview of these constituent parts is given in Figure 3.1.

3.2 Research Focus

This thesis has thus far outlined current concerns regarding lack of further study of physics by young people, and that significant contributors to positive student attitudes to physics include socio-cognitive aspects such as physics self-concept and the notion of a supportive physics teacher. Consequently, existing literature on self-belief constructs has been explored, along with feedback in line with the premise that the language that the physics teacher uses with the class could impact those self-beliefs. Additionally, works on teacher development have been examined, since the role of the teacher is deemed critical in this endeavour.

This chapter now explains how the research was designed; the term methodology has deliberately not been used for two reasons. Firstly, it is variously defined as ‘a body of methods, rules, and postulates employed by a discipline: a particular procedure or set of procedures’ and ‘the analysis of the principles or procedures of inquiry in a particular field’ (Merriam-Webster, n.d.), and ‘the methods and principles used for doing academic research’ (MacMillan Dictionary, n.d.). Many sources of educational research design do not actually define it, and it is seen to be conflated with terms such as epistemology, research method and research approach. Waring (2012) describes the multiplicity of labels resulting in confusion over both meaning and conceptual level of terms; ‘Sometimes it is difficult to distinguish clearly labels that denote an epistemological stance and those that refer to method’ (Tesch, 1990: 58, see also Grix, 2002). Secondly, this thesis comprises two theoretical frameworks as themes, each within a semi-longitudinal, quasi-experimental approach, in which a teacher CPD intervention connects the two, and utilises mixed data collection strategies that straddle the quantitative-qualitative dual discourse. The complexities are thus better unfolded as a large ‘research design’ (Figure 3.1).

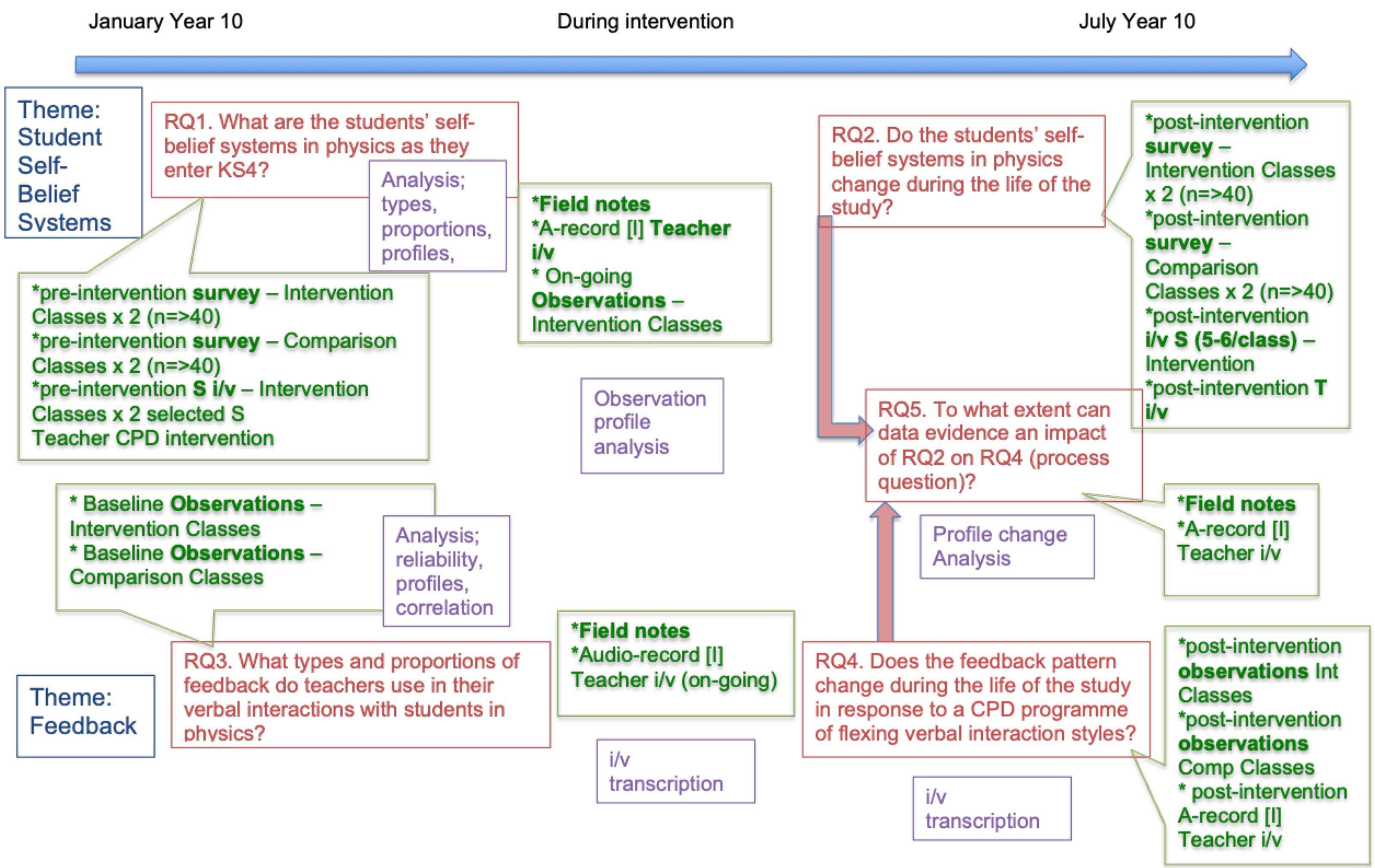


Figure 3.1 A schematic showing how the research questions relate to the data collection and analysis.

[Key: S= Student; T=Teacher; i/v = interview]

The first theme of the study concerned the self-belief systems of students in physics in Year 10 and asked two research questions:

1. *What are the students' self-belief systems in physics as they enter KS4?*
2. *Do the students' self-belief systems in physics change during the life of the study?*

The first question was concerned with identifying self-belief constructs as outlined in the literature (Lee, 2009; Ahmed, 2012; Marsh *et al*., 2019a; Seon and Bong, 2019), and the second with investigating whether these changed over time. This study used a pre-intervention survey to explore the range and depth of a moderate sample (n=66) of Year 10 students' self-belief systems in physics, from both intervention and comparison classes. The output from this survey was used to create subsets of students of differing beliefs within the test groups, and who were interviewed pre-intervention to provide rich contextualised baseline data in addition to the survey. A post-intervention survey was used to track any change in the self-belief system of the intervention cohort groups, comprising self-concept, self-efficacy and anxiety, against a sample of Year 10 students who did not receive the treatment (comparison group).

The second theme of the study, relating to the notion of the supportive teacher of physics (Hollins *et al*., 2006), encompassed the way in which teachers deployed feedback type orally, and whether the CPD had enabled the intervention teachers to modify their previously identified style to incorporate more self-regulation and process levels, which are less used in the classroom (Gan, 2011; Hattie and Masters, 2011; Bergh, Rose and Beijaard 2013; Brooks *et al*., 2019):

3. *What types and proportions of feedback do teachers use in their verbal interactions with students in physics?*
4. *Does the feedback pattern change during the life of the study in response to a CPD programme of flexing verbal interaction styles?*

This study will seek to examine how aspects of formative feedback in the literature map to the real-life situation of the physics classroom. Through a series of observations of teacher-to-student interactions, a baseline sense of what feedback types were presented in these teachers' practices, and in what proportion, were obtained. The interactions were verbal in nature, and an observation schedule (Appendix 1) was used to systematically capture categories of feedback events from the teacher to the students, using a constructed typology (see section 3.8.2.1). A tally of these could then be compared through time to ascertain whether the amount of self-regulation and process feedback used by the intervention teachers increased, through a professional development episode undertaken by the teachers involved.

A fifth research question therefore asked a process question:

5. *To what extent can data evidence an impact of RQ4 on RQ2?*

The overarching aim of this study is to investigate whether and how attribution training in the form of formative self-regulation, and where necessary process feedback impact the students' self-belief system in physics. Through pre- and post-intervention surveys, teacher interviews (coaching conversations), selected student pre- and post-intervention interviews and an extended series of classroom interaction observations, this study will seek to ascertain (on the hypothesising assumption that the intervention teachers *had* been able to flex [purposeful modification] their verbal feedback), whether there has been an impact from increasing the self-regulation feedback, and where necessary, process feedback, and also whether it had affected how the students viewed physics, and their perceived ability to 'do' physics, in their own vernacular.

It was not a certainty that such an intervention would have an increase on the amount of self-regulation and process feedback employed by the teachers; however, the professional development experience of the researcher suggested that it was a likelihood, as such changes can and do occur with varying success. This intervention question will explore how and indeed if such CPD can result in a change in feedback practice, but this study exceeds the normal time range over which more 'traditional' professional development tends to take place, such as in-service courses of varying length (Joyce and Showers, 2002; Darling Hammond and colleagues, 2009, 2017, 2019; Yoon *et al*., 2007; Gabelica *et al*., 2012). After an initial professional development session, there were follow-up teacher interviews at approximately two and four months, the semi-structured character of which served as a coaching conversation (Whitmore, 2002) to keep the intervention 'on track'. Through classroom observation and bi-monthly teacher interviews, the embeddedness of the practice was followed. The baseline content of RQs 1 and 3 are necessary reference points against which to judge whether such a positive change in either self-belief constructs and/or feedback styles occur. Students' views on feedback practices could also indicate whether they have perceived a change in their teacher's practice.

Extant literature explored in Chapter 2 has revealed a paucity of research regarding several key foci of the role of feedback within this study; there is a lack of studies on the impact of a feedback intervention on student self-beliefs (Johnson, 2016) and within the general feedback field, a lack of a) verbal feedback studies since written feedback appears to be preferenced; b) a consideration of feedback studies incorporating the perspectives of both teacher and student; c) some limited use of typologies to help teachers understand differences between feedback levels; and finally d) studies on the use of teacher language on student self-belief systems. However, examination of the research approach within apposite studies can provide a useful starting point for this investigation.

3.3 Philosophical underpinnings

This study proposes to investigate the amount and nature of self-regulation feedback within the situated experience of a physics classroom, and whether increasing the use of formative self-regulation and process feedback as 'attribution training' has consequences for the students' self-belief systems in Physics, and for the growth of a 'Physics Learning Identity'.

Waring (2012) outlines the four related assumptions through which researchers approach their work described by Grix (2002) as the four building blocks of research; ontology, epistemology, methodology (here, research design) and methods. My professional practice shapes the rationale for the adopted approach, proceeding from analysing my own assumptions about the nature of social science: an ontological perspective that social reality is a ‘product of individual consciousness’ (Cohen, Manion and Morrison, 2017 p.7); an epistemological perspective that acquired knowledge is subjective, unique and personalised (*ibid*); and a view on human nature whereby humans possess free will, and initiate and change their environments. This initially led to thoughts about a research design that was predominantly idiographic in nature, in seeking to understand individual behaviours and what may influence or change them, set within the interpretive (sometimes called social constructivism in the literature (Denzin and Lincoln, 2011)) paradigm.

The ontological perspective asks, ‘what is the form and nature of the natural world?’, and this is shaped by the objectivist-constructivist continuum. In objectivism, the assumption is that reality exists independently of individuals’ perceptions of it, whereas constructivism posits that ‘reality is neither objective nor singular, but multiple realities constructed by individuals’ (Waring, 2012:16). It is the latter description which holds more resonance for the researcher, based on years of teaching in educational settings.

Epistemology considers what should pass as acceptable knowledge, and here the relationship between theory and research is critical; a study in which the theory *guides* the research is deductive, yet the research questions are also asking for an inductive theory *derived* from the research, creating an epistemological dichotomy. A deductive approach, consisting of drawing upon extant knowledge to produce a hypothesis that is subjected to empirical scrutiny, can be operationalised through data collection. The hypothesis can thus be confirmed or rejected, and theory revised, thereby inductively contributing to apposite literature (Bryman, 2016). This may reflect a positivist, objective paradigm, the ‘so-called scientific method’ (Anderson and Arseneault, 1999). However, Bryman (2016:24) questions whether the social sciences should be ‘studied according to the same principles, procedures and ethos as the natural sciences.’ Cohen, Manion and Morrison (2017) also highlight the difficulty of applying positivism to the complexity of human behaviour, and certainly within the school classroom. There is a tension between a stance which asserts that ‘things are only meaningful if observable and verifiable’ (Anderson and Arsenault, 1999) and the fact that observations in the physics classroom are not value-free. Testing a hypothesis in social science research has both objectivist realities and subjectivist economies; ‘the search for meaningful relationships and the discovery of their consequences for action’ (Cohen, Manion and Morrison (2017:7). A more nuanced, reasoned approach may be to embrace the notion of ‘empirical realism’ arising from this dissatisfaction with positivism within social science (Slater, 2001) in which the reality is understood simply through appropriate methods (Bryman, 2016), thus avoiding charges of incommensurability between these dichotomies (*Coe et al*., 2017).

Research questions 1-3 are described as ‘descriptive’ by Anderson and Arsenault (1999), in which ‘what is happening’ is answered through (for example) survey or observation, given that methods are regarded to be free of ontological and epistemological assumptions (Grix, 2002). Research questions 4 and 5 are ‘explanatory’ (Anderson and Arsenault, 1999) in seeking to explore the impact of the professional development intervention upon the teachers’ feedback (should this happen), and ultimately did this cause students’ self-beliefs to alter? Overall, this study is evaluation research, in that the question being asked is, ‘has the intervention achieved its desired goals?’ Being a ‘nested’ design (the teachers are nested in the schools and the students are nested in the classes), proving impact distally is recognised to be challenging. Since it is a layered study in including both quantitative evaluations and qualitative narratives, this is not a wholly experimental approach. Greene (2000) recognises the differences in opinion about how qualitative evaluation research ‘should’ be carried out and discusses an in-depth understanding of the context in which the intervention occurs, the viewpoints of the stakeholders, and an examination of the range of outcomes of the intervention. The questions being researched are of a ‘what’, ‘how much’ and ‘how’ nature, within the multifaceted context of the physics classroom. It is essential to develop a rigorous methodological path, with attention paid to validity and reliability of evidence gathered by different methods in the ambiguous non-laboratory setting of a school classroom. Bryman (2016:2) writes of the ‘messiness of social research; it often does not conform to a neat, linear process’. This study has challenges of layered complexity and apparent methodological dichotomy, however, in adopting a pragmatic perspective, it is possible to reject traditional philosophical oppositions (Coe *et al*., 2017). The objective of this research will rely on the participants’ views and responses to the intervention, and the researcher’s intent to interpret the meaning of these views and responses.

3.4 Research design

In the literature review in Chapter 2, it was established that there was a paucity of sources explicitly combining the interacting areas of interest considered here. Yet it is useful to examine the research approaches of some of these studies to investigate what was investigated and how. Inclusion criteria therefore consisted of those studies which had looked at using or identifying feedback typologies and/or feedback frequencies (Gipps et al., 2000; Gan, 2011; Hattie and Masters, 2011; Bergh, Rose and Beijaard, 2013; Voerman *et al*., 2012, 2015; Adie *et al*., 2018; Brooks *et al*., 2019); the application of a feedback intervention (Voerman *et al*., 2015, a manipulated treatment of interrupted time series); a verbal feedback focus (Knight, 2003; Gamlem and Smith, 2013; Gamlem and Munthe, 2014; Voerman *et al*., 2012, 2015); those including a teacher perspective (Knight, 2003; Vercauteren, 2009); a student perspective (Vercauteren, 2009; Hargreaves, 2013; Dann, 2015; Kerr, 2018); or both (Vercauteren, 2009). This review also revealed that the data collection methods were also variable and included classroom observation, sometimes through a pre-devised observation schedule or thematic analysis (Hattie and Masters, 2011; Hargreaves, 2013; Voerman et al., 2012, 2015; Bergh, Rose and Beijaard 2013; Kerr, 2018; Brooks *et al*., 2019;); classroom audio/video recordings with transcript analysis (Vercauteren, 2009; Voerman *et al*., 2012; Adie *et al*., 2018; Skovholt, 2018); and student and/or teacher interviews or focus groups (Knight, 2003; Vercauteren, 2009; Gamlem and Smith, 2013; Hargreaves, 2013; Dann, 2015; Kerr, 2018). That methodological disparity existed despite some of these authors investigating similar themes may indicate that some studies' assertions may not be as rigourously defended as others which used more data collection sources for triangulation, as the 'major safeguard on validity is to obtain confirmation from as many data sources as possible' (Anderson and Arsenault, 1999). Additionally, gaining a teacher or student perspective might have contributed richer detail to findings, for example in Voerman *et al*.'s feedback CPD intervention (2015). Although some of these studies post-date the researcher's design approach, using findings from their analyses has proved useful in analytical decisions made here, such as the decisions by Adie *et al.* (2018), Kerr (2018), and Skovholt (2018) to only use limited amounts of transcript, and by Brooks *et al*.'s (2019) decision to exclude praise as a feedback category from their observations.

As explored above, the deductive approach necessitates an experimental design, whilst the reciprocating inductive theorising draws upon the 'messiness' of the teacher and student perspectives to infer implications for existing knowledge. Consequently, the main research design is one of quasi-experimentation, utilising data collection methods described above. This is a form of intervention research (Coe *et al*., 2017) which purports to effect some change in an educational setting and report the reaction to it. Additionally, the intervention teachers' perceptions of the intervention will be included to offer complementarity and convergence as a situated case study (Yin, 2014) offering rich and contextualised granular detail.

3.4.1. Quasi-experimental design

The hypothesis of this investigation is that an increase in self-regulation and process type feedback will have a positive impact on student self-belief systems in their physics education. This necessitates a research design whereby the independent variable [the type of feedback] is changed and the dependent variable [the student self-belief in physics] is measured. The physics classroom is, however, a non-laboratory setting in which other variables cannot be rigidly controlled; since these designs are close to being an experiment, but do not fully meet the requirements, their internal validity may be compromised (Bryman, 2016) [see later section on validity]. Quasi-experimentation ensues when, although there is an experimental group and a control group, the participants to each group have not been randomly assigned, often because there are practical difficulties with implementing it (*ibid: 50*). In social science research, these are therefore more commonly referred to as the treatment or *intervention* group, and the *comparison* group, rather than experiment and control. Consequently, quasi-experimentation is used to manipulate a treatment, often on multiple units, in order to infer a causal effect (Hedges, 2012).Thyer (2012:78) uses the term control, indicating that in his view, a comparison group refers to individuals who receive some sort of alternative treatment (which may be treatment as usual), which is not the intended meaning for this study, and comparison group [CG] will be used throughout.

A quasi-experiment is defined by Grant and Wall (2009:655) thus:

> *'… a study that takes place in a field setting and involves a change in a key independent variable of interest but relaxes one or both of the defining criteria of laboratory and field experiments: random assignment to treatment conditions and controlled manipulation of the independent variable.'*

One of the more commonly used forms of the quasi-experimental method is the non-equivalent groups design [NEGD], which in its simplest form requires a pre-test and post-test for treated and comparison groups (Trochrim, 2007) or the pre-test, post-test, no-treatment control group design (Thyer, 2012). This design is used 'to help determine if a given intervention produces any effects above and beyond those attributable to the passage of time, concurrent history, or the experience of being assessed' (Thyer, 2012:95). As random assignment to two GCSE groups is not possible within educational settings, it was desirable to use two intact groups (in each school) that were deemed to be as similar as possible; they are therefore *not* known to be equivalent on all possible factors, however using two groups per cohort rather than one does attempt to remediate this somewhat. Although these two Triple Science groups in each school are considered comparable and can be used for treatment and comparison purposes, it is wise to remember that they are not equivalent

(Trochrim, 2007), and to explore the susceptibility to all internal threats to validity.

With this non-equivalent group design, the members of both cohorts undergo a pre-intervention test to establish a baseline measurement. The *teachers* of the intervention groups receive a professional development trajectory over the duration of the study, whereas the teachers of the comparison group do not. At the end of the intervention, all the students in both groups undertake the same test at the same time (though not necessarily physically together) as a post-intervention measure. If the post-intervention measure of the intervention group demonstrates a change in what is being measured whilst the comparison group do not, there is held to be 'some modest logical justification' that it was indeed the CPD intervention that produced these changes (Thyer, 2012:95). Note here that there is no assumption that this will be a positive change, though indeed this is to be hoped for.

Early exponents of quasi-experimentation Cook and Shadish (1994:566) claimed that:

> 'the most frequently employed quasi-experiment still involves only two (non-equivalent) groups and two measurement waves, one a pre-test and the other a post-test measures on the same instrument.'

By incorporating two cohorts per group, it is hoped that this will decrease not only the internal threat to validity for selection, but also for mortality (see below), in that a greater initial number of participants may insulate against student 'drop-out'; should a student take a pre-intervention test but not a post-intervention test, they would have to be removed from the entire experiment. This may be an issue if there were, for example, a greater proportion of low-self-concept students leaving one group preferentially.

Whilst quasi-experimental studies are strong in ecological validity, and results of such studies can be compelling because they are not artificial interventions in social life (Bryman, 2016), there remains criticism of single group studies. Threats to internal validity are discussed in section 3.10.1, however the significant investment in a second group serves to eliminate many of these, provided the multiple-group threat to internal validity shows consideration of the selection bias. Grant and Wall (2009) mount a thorough exploration of the benefits and challenges of quasi-experimental approaches and offer suggestions to mitigate some of the concerns. Where lack of random assignments to groups and limited control to experimental variables occur, they suggest causal inferences can be strengthened by seeking further information that might discount rival explanation. The study is a double NEGD, in two schools, where additional data is being collected from the participants in the form of interviews and field notes during observations, thus strengthening findings in a contextualised setting.

3.4.2 Case Study design

Research question five is concerned with how the treatment teachers responded to the professional development relating to increasing their use and proportion of self-regulation and process feedback, and requires an in-depth understanding of events, processes, or situations, set in their real-world contexts. Although there is some dispute and even overlap between the methodological diversity within this interpretive approach, an often-cited typology is offered by Creswell (2013), who distinguishes the following: narrative research, phenomenological research, grounded theory, ethnography, and case study. A case study is the most appropriate procedure to explore an issue in which the case(s) are a specific illustration.

In rejecting other methodologies then, it becomes necessary not only to explain why, but to recognise if there are inherent aspects within them which might complement or strengthen a case study approach.

Narrative studies draw in personal experiences, as stories, usually gained through interview, and fashioned into chronology (Creswell, 2013). Although sharing a longitudinal nature, the intent of this study is to go beyond the narrative in crucially, intervening, and in not only developing an in-depth description but an analysis of the case(s). Phenomenological studies focus on shared participant commonalities as they undergo the same phenomenon, however this study goes beyond an exhaustive account of that experience in drawing upon the student perspective as well.

The key intent of a grounded theory approach is to go beyond description and allow theory to emerge that has been grounded in data from the participant within the process or experience. The premise of this study is that self-regulation and process feedback impact positively on self-belief, and as such the study is characterised as deductive rather than theory-seeking.

In an ethnographic study, the researcher is interested in those shared patterns however, the focus is on the whole culture sharing group, and the intent is to look at how the culture works. An ethnographic approach would entail becoming immersed in the culture of the classrooms as a participant observer.

Creswell's typology should not however be taken to imply that the boundaries between method choices are sharp: although each method has distinct characteristics, there are large overlaps between them: '*the goal is to avoid gross misfits'* (Yin, 2015:9). On consideration of the issues to be explored, I have adopted a case study approach for the 'best fit' for exploring a '*real-life, contemporary bounded system…through detailed, in-depth data collection involving multiple sources of information'* (Creswell, 2009:97)

There are inherent strengths in the case study approach, providing as it does a rich and authentic account within a real-life, complex, non-laboratory context: '*They catch unique features that may otherwise be lost in larger scale data (e.g., surveys); these unique features might hold the key to understanding the situation'* (Cohen *et al*, 2017:379) and in: '*…being strong in realism, so that it recognises the tensions and ambiguities of social truth, and that it allows tentative generalisations about or from the case to wider education'* (Johnston, 2012:196).

They do have their disadvantages however: case studies are sometimes disdained as a strategy over concerns such as a lack of rigour, perhaps in the lack of systematic procedure, or presence of bias. Another frequent argument is that they provide little basis for scientific generalisation: 'case studies, like experiments, are generalizable to theoretical propositions and not to populations or universes.' (Yin, 2015:21).

There is also a concern about the ability of a case study to establish causal relationships, as opposed to using 'true experiments' such as randomised field trials. However, this is to not fully appreciate the complexities and ambiguities of non-laboratory settings, and an essential tactic is to use multiple sources of evidence, with data needing to converge in a triangulating fashion. This challenge is but one of the ways that makes case study research 'hard', although it has classically been considered a 'soft' form of research (Yin, 2015:2).

Yin (2015) identifies 3 types of case study: a) Exploratory, which may be a pilot study in preparation for future research; b) Descriptive, which may provide narrative accounts; and c) Explanatory, which may test theories and hypotheses. Bassey (1999) describes three different categories of empirical research (albeit with mobility between them): a) Theoretical carried out to understand; b) Evaluative carried out to understand and evaluate; and c) Action research carried out to understand evaluate and change. Consequently, I would classify this case study as a blend of descriptive and evaluative.

Creswell (2013) points out the tension between a study design of single or multiple cases. Having only one case will dilute the analysis, whereas having more than one case will inevitably reduce the depth spent on each one. It is recognised however, that a larger number of cases enables generalisations (even if the fuzziness is not refined):

> *'It [a fuzzy generalisation] is not an admission of frailty in the way that the research was conducted. It is a firm reminder that there are many variables which determine whether learning takes place. And it is an invitation to teachers to enter into discourse about it: to read the evidence in support of this statement, to discuss it with anyone else who engages in x, to reflect on the issue, to test out in their own classrooms the efficacy of y and to report the outcomes to whatever group will listen*' (Bassey, 1999:51-52).

Case study designs need to maximise their quality through four critical conditions related to design quality: construct validity, internal validity, external validity and reliability. (Yin 2015:45), and some of these considerations are addressed below.

3.5 The Pilot Phase

Numerous authors advise of the necessity to conduct a pilot study (Glasne, 2006; Ashley, 2012; Bryman, 2016; Cohen, Manion and Morrison, 2017), in order to test the procedures and techniques to ensure that they work as they are intended to (Anderson and Arsenault, 1999). These authors offer several reasons to expand upon this; Glasne (2006) recommends that the pilot be conducted in as similar a setting to the intended one to bring one from the drawing board to the realities of the test situation; to refine the research instruments being used, such as survey and interview questions (Bryman, 2016) and increase the reliability, validity and practicability of a survey (Cohen, Manion and Morrison, 2017); identify if data are missing (Bryman, 2016); and perhaps use the opportunity to pilot intended data analysis methods on a small sample (Ashley, 2012).

In identifying the main data collection tools that would need to be deployed to investigate these research questions, a selection of survey items were drawn from existing validated self-belief surveys (Lee, 2009; Marsh *et al*., 2009; PISA, 2006), since Bryman (2016) indicates, that in a sense, these have already been piloted and validated. A range of interview questions were collated for both the students and the participating teacher, and this enabled ethical and logistical considerations to be rehearsed also.

3.5.1 The Pre-Pilot

A pre-pilot was conducted in a North Yorkshire secondary school [not a school in the main study] with Year 10 students, aged 14-15, in the first year of their two-year GCSE course. A focus group comprising of seven of these students, (four females, and three males) was interviewed. A survey had previously been administered to two Physics classes, from whom these focus group students were drawn. The classes had been described as one ‘high ability’ and one ‘middle ability’ by the classroom teacher, who was known to the researcher.

‘Verbally expressive’ students had been selected for the focus group by their Physics teacher after consideration of their ability and confidence at answering in class. The interview questions sought to obtain a collected view of their wider perceptions and experiences of physics, and of their impressions of feedback in and resulting from physics lessons. It produced a large amount of data on attitudes, values and opinions:

They clearly indicated that they liked their teacher and found her personally supportive. Indeed, they were occasionally defensive of her, when discussing the feedback policies of the school. The students reported that they quite enjoyed physics; saw the importance of it as a subject, but not its relevance to 'real-life'; ‘*when am I ever going to use the parallax angle?’* They did not see themselves 'being' physicists (DeWitt *et al.*, 2019), although two said that they would need it for their (then current) career choices. They did not on the whole regard or value the feedback they received, which seemed mainly task related. They perceived feedback as 'meaningless questions' that were more oriented towards neat books rather than learning (a school marking policy). Feedback tended to occur at endpoint, rather than at the learning instance.

The student perceptions of feedback as based on one episode with one focus group were informative but inconclusive. To these students, feedback existed as questions written at the end of submitted work, subsequently returned. Usually, these questions related to grammatical and/or presentational issues rather than seeking to extend their learning further. Instructional, procedural or motivational aspects were ignored or not utilised. Feedback as a currency to these students seemed depreciated: they did not value it or seek it out. They appeared wearied by the school 'policy' of feedback questions, and largely ignored it.

When questioned explicitly on feedback about learning, students ceded that if they had a problem, 'Miss' would help them, but they described these as isolated incidents. From the professional development experience of the researcher, it was difficult to believe that these were the only types of feedback that existed, and it was likely that there were other, multiple, interactions which were not being characterised by the students as feedback, but which may have been occurring all the time. It seemed likely that in observation, instances of feedback of different types would be documented, and therefore a method of measuring a baseline indication of types and frequency became desirable.

A pilot survey had been prepared and administered which constituted items seeking to establish whether the instrumentation could reliably capture student attitudes to self-concept, self-efficacy, anxiety to physics, and mindset. Each area had multiple questions to enable cross-checking of responses, including some reverse coding. 51 students completed the survey, although some respondents did not fully complete each question. An analysis of the results was undertaken, and some findings are given below:

1. Self-concept showed a somewhat mixed picture. Most agreed that they learned physics quickly and got good grades in the subject (63.7% and 68.6% respectively), and approximately half agreed that they understood even the most difficult work, more than half then described themselves as 'just not that good' at physics, and more than 78% did not believe that physics was one of their better subjects, including 85% of the girls.
2. The self-efficacy items of content-specific tasks tended to show a higher proportion of agreement from the students, such as judging their competence to perform speed calculations, or evaluate methods of generating electricity. The combination of the self-concept and self-efficacy results supports literature which finds that students can 'do' physics, but don't view themselves as being a physicist or being good at physics. Also, since self-efficacy evaluations tend to be made against specific challenges, the student judges whether they are able to do it, often against a background of having previously done it, thus experiencing mastery to some degree. Self-concept challenges may be more linked to how the student views themselves against not only their peers, but also against a (possibly nebulous) perception of a mastery standard.

3. In questions concerning anxiety about physics, 70% of girls reported agreeing to strongly agreeing, in four of the five items. In contrast, the boys exhibited less anxiety about physics.
4. Although the girls were a smaller sample, they outnumbered the males in subscribing to fixed intelligence theory and seemed to hold those views more strongly.

On reflection, the survey showed an imbalance of distribution of item types across the self-belief constructs. Additionally, the internal consistency of the survey had to be calculated so that the reliability of the instrument to measure the same construct in different items and yield the same result could be known (see the evaluation study below).

3.5.2 The Pilot

It had become apparent through the preliminary testing that classroom measures of oral interaction types would need to be obtained through observation. Since the types of feedback had been influenced by existing typologies (Hattie and Timperley, 2007; Gan, 2009; Hattie and Masters, 2009; Bergh, Rose and Beijaard, 2013), a structured observation schedule was created using support guidance from Cohen, Manion and Morrison (2017), and decisions taken to create a constructed typology for the purpose of this study: what was the first level of feedback as described by the authors above, but also, was it a statement feedback, or a prompt for future action on learning as a second level? The Tunstall and Gipps typology (1996) was rejected for this study as being too complex to make on-the-spot evaluations in the lesson, and also due to the researcher's unease in the categorisation (see page 51 of this thesis). It was essential that such a potentially complex schedule be piloted to ensure no overlap of categories, and practice was required to achieve dexterity in using the pre-determined codes within the table. This was trialled in a York (England) secondary science classroom over several lessons with a second physics teacher known to the researcher [again, not a school used in the main study]. The survey, having been re-balanced for split-half reliability was also piloted with this class (n=26). It had additionally been made simpler to look at, Arial font used for readability (a preferred font for dyslexia), the Likert scale made into four sections so that students were persuaded not to take a ‘middle ground’, and a clear instruction at the base to turn over to the second page of the survey.

3.5.3 Pilot Evaluation Study

The purpose of evaluating the pilot studies was to ensure the utility and credibility of the instrumentation (Patton, 2014). The utility concerns the usefulness of the study: does it allow its value to be decided and enable assessment of how justified conclusions made from it are. Its credibility is concerned with the rigour of the objectivity, validity and reliability of the procedures and instruments used (*ibid*).

The validation of the re-worked survey, and calculations of its inter-item and split-half reliabilities (Cohen, Manion and Morrison, 2017) are examples of such an evaluation. Due to the inconsistencies and imbalances of survey items raised by the pre-pilot, it was necessary to further refine the survey, and undertake some evaluation procedures on the pilot material, which required trialling on another class. Specifically, the survey was re-balanced and validated, and a structured observation schedule trialled to capture incidence and type of feedback. The finalised version of the survey tool and observation schedule are given in Appendix 2.

The initial schedule began as a simple tally of ‘events’, but its trial indicated that those initial categorisations proved to be insufficient or not clear enough, for example when more complex combinations of feedback occur. Through the pilot work the observation schedule underwent a number of iterations until the final version could be used in the main study, and this can be seen in Appendix 1.

3.6 Selection of research participants

There were four physics teachers taking part in the study; two in the intervention groups and two in the comparison groups. None of the teachers had a personal or professional connection to the researcher, however the researcher, through her professional education background had familiarity with the schools, which were both in North Yorkshire, England. The head of science of one institution asked one of her physics teachers if they would be interested in taking part. Another physics teacher was recruited via a pilot study teacher who was not able to participate in the full study due to a changed timetable. These two designated intervention teachers then recruited their ‘equivalent’ physics teacher on the Triple Science programme in their school to serve as the comparison classes. The comparison groups and their teachers were asked to take part to serve in parallel within the same institutions, in case any existing feedback, or other policies had been institutionally structured and enacted. The first teacher [T1] was in their second decade of teaching physics, as was their partner comparison Teacher 3 [T3]. In the second school, both teachers were relatively new to teaching though not newly qualified; Teacher 2 [T2] was in the intervention group teacher in this setting, and Teacher 4 [T4] the comparison teacher. This then reflected a range of physics classroom experience.

These groups were typically 20+ students per group, with one intervention group and one comparison in each school (pre-test n=84). There were both males and females in both the intervention and comparison groups, at an approximately 2:1 ratio. All four sets were mixed ability Year 10 (age 14-15), and all were classes termed ‘Triple Physics’; that is, all were taking the three sciences as independent GCSEs. Many schools enable higher ability students to take this option in Key Stage 4, and this was the case in each of the partner schools. However, due to the size of each school, these students were drawn from across two sides of the student body and were deemed by their science departments to be ‘comparable’, though Key Stage 3 data was not made available to the researcher.

3.7 The Intervention

Drawing on the CPD research literature as well as the professional development background of the researcher, an intervention was produced in the form of a 90-minute training session with accompanying PowerPoint slides and tuition activities. An abridged form of this can be seen in Appendix 3. Its intended learning outcomes were for the intervention teacher to understand the different levels of feedback (according to the Hattie and Timperley (2007) typology); to become aware of the different self-beliefs students possess; to consider what the effects of these different levels are and to identify how students receive, comprehend and use feedback in their learning. The intervention teachers were asked for their own definition of feedback, so that their current views could be established. Using several sources as a thinking activity, this definition was teased out, and the teachers indicated that they positioned feedback as 'where the student is now' and 'next-steps' advice to help them move on in their learning. They also shared what they wanted to happen to student learning as a consequence of taking part in the research, which was for their students to take more ownership of their learning.

The position of feedback within the five areas of assessment for learning was shared, which moved on to a discussion of how students view their successes and failures in learning. Both teachers were familiar with the notion of 'mindset' but less knowledgeable about the different self-belief constructs. There was some discussion of how feedback could impact on student self-belief, and both teachers shared illustrative anecdotes about cohort members. Using the notion of calibrating the learning need within the learning instance, the teachers were challenged to 'flex' their verbal feedback response during a training activity. This enabled them to realise the challenge of having to consciously think about what they were going to say to promote a more regulatory response on the part of the learner. At this point they were made aware of the frequency with which the different levels of feedback (Gan, 2011, Hattie and Masters, 2011; Bergh, Ros and Beijaard, 2013) were used by teachers, and that the self-regulatory level was the least used. The teacher conception of feedback was compared to the students', as well as the view that verbal feedback was not always recognised as feedback, occurring within the dialogue of the classroom (Yang and Carless, 2013). Practical strategies for involving the students in understanding learning criteria, recognising & encouraging effort and persistent behaviours, not shying away from challenge and modelling beliefs in the potential for success were all examined. A ‘bookmark’ of sentence starters was shared to enable the intervention teachers to practise their feedback language was shared (Appendix 4).

There followed successive phases of teacher continuing professional development, facilitated by the researcher as ‘teacher interviews’, on the use of self-regulated feedback planned at months 1, 3 and 5 of the life of the study, for the intervention groups only. The teacher interviews took the form of coaching conversations using the GROW model (Whitmore, 2002), and provided a basis for continuing professional development as well as data gathering (see 3.8.3.2). Extracts from both the teacher intervention and coaching conversations are discussed in Chapter 7.

3.8 An overview of the data sources and collection procedures

The design of the study required data collection that was extensive, and employed mixed methods not always confined to a quantitative approach. Figure 3.2 outlines how a method of data collection was identified to answer each research question and enable triangulation to answer the overarching goal of the study. A multi-tool approach is less susceptible to criticism of rigour than had a single approach been adopted (Bryman, 2016)

In this study, research questions one and two asked:

1. *What are the students’ self-belief systems in physics as they enter KS4?*
2. *Do the students’ self-belief systems in physics change during the life of the study?*

It was therefore determined that survey instrumentation was the appropriate tool to gather data on Year 10 students’ self-belief constructs which could not be directly observed from both intervention and comparison group classes to the extent required. Observation schedules and survey composition were developed from existing literature (Lee, 2009; Marsh *et al*., 2009; PISA, 2006), reviewed and modified following the pilot study. This would enable identification of individuals' self-belief constructs, and the same test, applied at the end of the intervention study would indicate whether these were stable, or had changed.

Analysis of the pre-intervention surveys yielded sub-sets of students with particular suites of constructs from both intervention and comparison groups, for example: high self-concept – low anxiety – high efficacy – 'growth' mindset; low self-concept – high anxiety – low efficacy – 'fixed' mindset. In discussion with the intervention teachers, 4-6 students in each treatment class were identified that were suitably 'polarised' enough in their beliefs to follow and consented to take part in a focus group to provide further information on their notions of self-belief that were rich, granular and contextualised. These students were therefore interviewed pre-intervention to gather data on their sense of the situation at that point in time. The change in student self-belief was also tracked, though in a less structured way, through on-going classroom observations, conversations with the host teacher and researcher field notes, which provided additional rich context.

For the second theme, that of teacher verbal feedback, research questions three and four asked:

3. *What types and proportions of feedback do teachers use in their verbal interactions with students in physics?*
4. *Does the feedback pattern change during the life of the study in response to a CPD programme of flexing verbal interaction styles?*

Using an observation schedule refined through a series of iterations developed and described in the pilot study above, a baseline frequency and type of feedback was established for all intervention and comparison groups through several audit observations conducted during the physics classes of each teacher, as one observation was insufficient to yield reliable baseline data. Each lesson was additionally audio-recorded so that data was not lost as the teacher moved around the room, and to enable later transcription if required. Mercer (2007:3) notes recorded lessons as 'a cohesive temporal discourse' and that 'talk functions in a temporal context… which even researchers involved in longitudinal studies can only sample' (*ibid*, 2010:10), thus highlighting the importance of capturing in-the-moment dialogue. Analysis of the observation data would enable any change in teacher feedback frequency and type to be verified; the use of a quasi-experimental approach having been chosen to allow the ‘modest justification’ (Thyer, 2012) that the intervention had produced any change.

Observations and audio-recordings of these intervention groups continued on a regular basis through the duration of the intervention months, focusing on the frequency and type of feedback (Figure 3.3). The observation schedule had been modified and amplified during piloting to enable concurrent field notes to be added *in situ*.

Finally, the post-intervention survey was administered, again to intervention and comparison groups, and post-intervention interviews of the sub-set students held (see Figure 3.1 for a schematic representation).

To gain an answer for the fifth research question:

5. *To what extent can data evidence an impact of RQ4 on RQ2?*

That is, to what extent could the impact of the intervention on the teachers’ feedback practices be implied to have impacted on student self-beliefs in physics, evidence had to be drawn from multiple sources, not least from the opinions of the teachers involved themselves. As this could be seen to be a two-step intervention, thus potentially claiming an impact further downstream, it was critical here to act to reduce researcher bias, and this is discussed further below. The study thus incorporates survey instrumentation, structured classroom observations, teacher audio recordings, teacher interview and student focus groups, as well as field notes.

This section will now describe the data collection tools in greater detail, outlining with reference to research design literature how they were developed and modified for the purpose of the study through the pilot phase above, and

Research Question	Survey		Observations		Interviews				Field notes
	Pre-intervention survey	Post-intervention survey	Pre-intervention Audit Feedback Observations	During intervention Audit Feedback Observations	Pre-intervention Student interviews	Post-intervention Student interviews	During-intervention Teacher coaching conversations	Post-intervention Teacher interviews	
RQ1	✓				✓				✓
RQ2		✓		✓		✓		✓	✓
RQ3			✓	✓	✓	✓	✓		✓
RQ4			✓	✓	✓	✓	✓	✓	✓
RQ5						✓	✓	✓	

Figure 3.2. An overview of data sources to gather evidence for the research questions

<table>
<tr><td rowspan="10">Pre-intervention survey, analyse and create subsets (intervention focus groups)</td><td rowspan="2"></td><td colspan="6">Intervention period</td><td></td></tr>
<tr><td>Month 1</td><td>Month 2</td><td>Month 3</td><td>Month 4</td><td>Month 5</td><td>Month 6</td><td></td></tr>
<tr><td rowspan="6">Pre-intervention audit observations per class</td><td colspan="6">Intervention Class 1</td><td rowspan="6">Post intervention surveys and interviews of sub-set students</td></tr>
<tr><td>CPD intervention (Teacher)</td><td></td><td>Coaching conversation (Teacher Interview)</td><td></td><td>Coaching conversation (Teacher Interview)</td><td>Post-intervention Teacher interview</td></tr>
<tr><td>Student sub-set pre-intervention interview</td><td>Oral Interaction Observations</td><td>Oral Interaction Observations</td><td>Oral Interaction Observations</td><td>Oral Interaction Observations</td><td></td></tr>
<tr><td colspan="6">Intervention Class 2</td></tr>
<tr><td>CPD intervention (Teacher)</td><td></td><td>CPD intervention (Teacher)</td><td></td><td>CPD intervention (Teacher)</td><td>Post-intervention Teacher interview</td></tr>
<tr><td>Student sub-set pre-intervention interview</td><td>Oral Interaction Observations</td><td>Oral Interaction Observations</td><td>Oral Interaction Observations</td><td>Oral Interaction Observations</td><td></td></tr>
<tr><td rowspan="2">Pre-intervention audit observations per class</td><td colspan="6">Comparison Class 1 Comparison Class 2</td><td rowspan="2"></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Post-intervention audit observations per class</td></tr>
</table>

Figure 3.3. A temporal overview of data sources for all groups in the study

how disadvantages of each collection tool were minimised through both careful design and meticulous researcher behaviour.

3.8.1 Survey instrumentation

In this study, survey instrumentation gathered data on Year 10 students' self-belief constructs which cannot be directly observed from both intervention and comparison group cohorts. These groups were typically 20+ students per group, with two intervention groups in each school (pre-test, n=84); the survey can be seen in Appendix 2.

The survey is of a cross-sectional type to gather data from individuals at one point in time (Vignoles, 2012) although this has been used in a semi-longitudinal sense at two points, since asking recipients about something that happened previously is unreliable (*ibid*). There are certain feasibility issues controlling the use of survey instruments, (such as number of respondents, response rate, what facilities are there to process and manage the study?) which have to be balanced against the desired outcomes (Cohen, Manion and Morrison, 2017). The survey design was influenced from pre-existing validated surveys from the literature; most notably Dweck's various 'mindset' surveys (2010), and the self-belief testing surveys of Lee (2009), Marsh *et al*. (2009) and the PISA surveys, all previously validated by their authors. The literature review indicates that self-belief constructs can be distinguished (Seon and Bong, 2019) so that the instrument had to satisfactorily target separate, operationalised constructs.

On reflection of the pilot evaluation, the survey showed an imbalance of distribution of item types. Additionally, the internal consistency of the survey was calculated so that the reliability of the instrument to measure the same construct in different items and yield the same result could be known (see section 4.1).

Bryman (2016) remarks that the survey and the structured interview are similar methods in social research; this does however mean that in the absence of the interviewer, the survey should be easy to follow and the questions easy to answer (*ibid*:222). Sellitz *et al*., 1976 cited in Cohen, Manion and Morrison (2017) provided a useful guide for construction of questionnaires, and since the purpose of this survey was to identify self-belief constructs, yet also obtain frequency counts, a Likert-scale response for 20 self-belief items was created, since the use of ordinal scales of data yield tallies that can be analytically manipulated. They ‘build in a degree of sensitivity and differentiation of response whilst still generating numbers’ (Cohen, Manion and Morrison, 2017:480). The decision was made to have four ordinal points [‘Strongly disagree’, ‘Disagree’, ‘Agree’, ‘Strongly agree’], thereby avoiding a participant tendency to choose the neutral midpoint. Both Bryman (2016) and Cohen, Manion and Morrison (2017) highlight disadvantages to surveys, including a lack of a check on truthful responses, respondee avoidance of extreme choices, subjective interpretation of an ordinal point such as ‘agree’, and a lack of equal scales in using Likert; ‘Strongly agree’ is not twice the measure of ‘Agree’. It is important therefore not to infer greater distinction or subtlety from the responses than they bear (Cohen, Manion and Morrison, 2017).

This data was collected from students at the same time both before and after the classroom intervention, and the same data collected from a comparison group of students who did not receive the intervention. The surveys were completed in class time to maximise the number completed, however, any student who did not complete both the pre- and post-intervention survey had to be excluded from the survey for continuity of data (experiment mortality Campbell and Stanley (1966, cited in Grant and Wall, 2009)); the total sample taking part was n=66 (33 from each group).

The pre-intervention survey output underwent preliminary exploratory data analysis, in order to identify if students were presenting with particular combinations of constructs, for example, high self-concept, high self-efficacy; low self-concept, low self-efficacy student subsets, which together with student information from the intervention teachers might indicate students to approach to form a focus group to gain thicker descriptions of their physics self-beliefs.

3.8.2 Observations

Angrosino (2012) notes that observation-based research is a well-established method in educational research whereby the observer is able to gather temporal and in situ information on day-to-day events, processes and behaviours. Naturalistic observation of this kind, identified here as unobtrusive or non-reactive observation, holds value for its ecological validity, in which the researcher has not taken part in the action they are observing (*ibid.*). However, as Cohen *et al*. (2017:542) note, 'observation is more than just looking'. It is the systematic checking of people, events, dialogue and behaviours within a setting.

The observation in this study was of the structured type (Cohen *et al.*, 2017) which allows numerical data to be generated from the frequency of interaction occurrence. The event, in this study the type of teacher-to-student interaction used, was simply written down, and contextualised both temporally and *in situ* with field notes. To enable this, it is necessary for the factors under study to be both clearly identified and discretely categorised prior to undertaking observations so that their incidence can be recorded, and these, as well as practising to be able to capture the factors, were part of the purpose of the pilot study. As the observations were to be analysed, it was likely that the data would be aggregated, and finer categorisation (e.g., of single students) would not be available. The schedule needed to remain open to developing this finer categorisation and coding, as sub-sets of students become identified, and this was therefore an issue for the pilot evaluation study to balance a robust enough schedule that was not too open to become unmanageable.

The observation schedule (Appendix 1) was developed from the pilot study that captured the levels and types of teacher-to-student interactions and identified to whom the interaction was addressed. Drawing from Hargreaves' (2013, 2014) conceptualisations of feedback as 'all the comments made by the teacher as a reaction to any activity or behaviour by pupils' (2014:295) enabled all teacher verbal interactions to be recorded and coded. Observations continued through the study as shown in Figure 3.3, comprising 45 lessons in total. The classroom observations were also audio-recorded through use of a tie-clip microphone and were subsequently coded for feedback content. Some of the classroom interaction was transcribed. There were several initial (baseline) observations, as the assumption cannot be made that one audit is characteristic of current practice: the teacher may significantly vary their interaction style for differing abilities, for example, and the type of feedback can also be context dependent.

3.8.2.1 Development of the Teacher-to-Student Interaction Typology

Observation data was initially used to help answer RQ3, to assess the nature and frequency of feedback, however the *duration* of each sort was not to be recorded. Although some authors deem VF to be of low frequency in the classroom (Hattie,1999; Bond *et al*., 2000; Campbell-Mapplebeck, 2019), pilot work indicated that it was not, although it must be noted that some of these studies used a different typology. Observation is a systematic and ultimately quantitative data collection method in this instance: the schedule had to be multifaceted enough to cope with the complexities of the classroom, but not so that it was difficult or onerous to complete in the moment. The schedule so produced was designed to minimise observer bias - being effectively a structured event sampling record in the form of a modified tally sheet. Fitting the teacher with a microphone also enabled non-directly observed or quiet interactions to be included.

A simple tally schedule to provide a frequency count of Teacher to Student verbal feedback events did not meet the criteria of more complex interactions. The schedule needed to unpick more than the occurrence of the event, and so has been modified to capture whether the feedback was a statement and/or prompt, and at whom this was directed to see whether there was variation in this too, within classes, and between classes, or whether more random or undifferentiated. Ultimately, there were three categories characterised by mutually exclusive variables: the *type* of interaction (6 variable of which 4 variables were feedback); the *mode* of feedback (2 variables); and the *recipient* (4 variables) to obtain a finer-grained view of events (see Table 3.1).The timing aspect of this version enabled multiple and complex combinations to be recorded. Should the timing aspect prove to be irrelevant, the table could be turned into a tally sheet of different possible combinations, though this might have provided demands of a different complexity.

Interaction category or type (first level)		**Feedback mode** (second level)		**Recipient** (third level)	
Task/product	**T**	Statement	**s**	Girl	**G**
Process/strategy	**P**	Prompt	**p**	Boy	**B**
Self-regulation	**R**			Class	**C**
Praise/self	**S**			Team	**T**
Instructional	**I**				
Other	**O**				

Table 3.1 A constructed feedback typology at three levels.

As the observations were to be analysed, it was likely that the data would be aggregated, and finer categorisation (e.g., of single students) would not be available. The schedule needed to remain open to developing this finer categorisation and coding, as sub-sets of students become identified, and this was therefore an issue for the pilot evaluation study to balance a robust enough schedule that was not too open to become unmanageable.

Each type of code, and some combinations of code was then tallied and summed for each lesson, and each teacher. From this operation, it is possible to show the percentage frequencies of all *types* of oral interaction, as summarised for all teachers in Table 3.1. The coding employed three semantic sampling levels; the first related to the feedback *type* (defined as a level by Hattie and Timperley, 2007; Gan, 2011); secondly, the feedback *mode* (whether a statement or a prompt for future action); and thirdly, the *target* or recipient(s) of the feedback.

The data was generated in four Year 10 physics classrooms, each led by a different physics teacher. The teacher dialogue was recorded *in situ* for later analysis. The researcher was sometimes present for observations. All recordings were held securely for later verification and three lessons underwent rater-reliability (see section 3.10.1), which also supported dependability (Bryman, 2016).

Due to the theoretical propositions that underlay the study, certain features of teacher feedback dialogue were designated as sampling units of a semantic kind; these related to the feedback *type* (defined as a level by Hattie and Timperley 2007), the feedback mode (whether a statement or a prompt for future action, after Ramaprasad, 1983), and the *target* or *recipient*(s) of the feedback:

First level descriptors

Task-Related (T)

- Provides information about the correctness of the learner's responses.
- Also informs the learner of the correct answer, but without suggesting how to revise the response.
- May provide indication of error/incorrect response or location of mistakes (if a prompt).

Process-related (P)

- recognises the strategy or process utilised by student
- Provides strategies/cues/hints/examples for error detection, information search or steps to revise report.
- May suggest explanation or justification for correct/incorrect response and reason for the use of a particular search strategy or revision approach.

Self-regulation (R)

- Provides reflective or probing questions that guide the learner in self-evaluation, seeking additional information, or monitoring of learning progress. Also relates to effort, persistence, resilience etc. (This can code negatively if done if reverse, e.g., you're not putting effort in, etc.)

Self/Praise (S)

- Remarks that are directed to the self mainly to give encouragement or affirmation and contains little or no task- related information.

The level codes were initially devised for the pilot study and were intended to provide a frequency count of Teacher to Student verbal feedback events, but this did not fully meet the criteria or more complex interactions. Through an inductive process, it became necessary to identify whether the teacher feedback was a statement and/or prompt, and at whom this was directed to see whether there was variation in this too, within classes, and between classes, or whether more random/undifferentiated.

The observation schedule was also devised as a timeline, enabling multiple and complex combinations to be recorded. In the case that the timeline proved to be irrelevant, the schedule would still exist as a 'tally sheet' of different possible combinations, though this might provide demands of a different complexity.

As a result of the inductive process, and in early involvement of a professional colleague as a reliability rater, it was decided to include further coding categories for first level descriptors:

Instructional (I) for dialogue that was teacher instruction [first level], for example, when doing board work

Other (O), for comments that were unrelated to the task [first level], often characterised as classroom activity or behaviour implementation comments or instructions

Second level descriptors

Statement (s): Provides information about the particular level:

- Task/Product – correctness of response, criticism,
- Process/Strategy – recognises what strategy or process has been employed
- Self-regulation – recognises effort, persistence, resilience, reflection
- Praise/self – undifferentiated, personal, non-task-related statements

Prompt (p): a cue for further development:

- Task/Product – how to improve the response, clarification
- Process/Strategy – provides strategies/suggestions/cues for error detection, next steps, information search, procedural hints
- Self-regulation – reflective probes that guide further self-evaluation
- Praise/self – non-specific prompts, unrelated to learning intentions

Third level descriptors

It became necessary to include a further category for groups of students however G already indicated the target was a girl, hence:

Team (T) when feedback was directed to students who had been arranged into groups, hence some of the feedback was group rather than class or individual related [third level]

At the second level, the inclusion of types of questions as *prompts* for next steps action is somewhat contended in the literature, with some authors (Knight, 2003; Voerman *et al*., 2012, 2015; Hargreaves, 2013; Svanes and Skagen, 2017) regarding these as instructional. However, with suitable distinctions made through the typology as to what actually constitutes an instructional prompt versus a verbal feedback prompt (which Knight did not make), it is proposed that from the student perspective, the recipients may indeed view these as feedback, which Voerman *et al*. (2012) acknowledged.

As described in the methodology, for the purposes of the study, it was important to be able to distinguish between types of oral interactions, since some of these do not relate to *learning* within the lesson. Table 3.2 outlines how sub-totals of interaction type were obtained:

Sub-total denoted by		First level codes	Second level codes	Third level codes
$\mathbf{N_T}$	*Total* oral interactions	T, P, R, S, I	p, s	B, G, C, T
		O	none	none
$\mathbf{N_L}$	Oral interactions relating to *learning*	T, P, R, S, I	p, s	B, G, C, T
$\mathbf{N_F}$	Oral *feedback* interactions only	T, P, R, S	p, s	B, G, C, T

Table 3.2 Nomenclature of sub-totals for comparison purposes

$\mathbf{N_T}$ = *Total* oral interactions (all first level codes)

$\mathbf{N_L}$ = Oral interactions relating to *learning* (code for Other (O) removed)

$\mathbf{N_F}$ = Oral *feedback* interactions only (code for Other (O) *and* Instruction (I) removed)

Examples of these oral interactions at the first two levels are given in Table 3.3:

Interaction level 1 (*Feedback* levels are asterisked*)	Interaction type exemplars	
	Statement	Prompt
Task/Product*	You have explained the meanings of the key terms. You have linked the name to the function. You have included a lot of facts in your answer.	Do your answers meet the success criteria? What other information is needed? Where did you go wrong on that?
Process/Strategy*	You have said why you have grouped those things together. You used a sequencing technique to explain how it works.	Explain to me how you … What was your starting point? How did you do it/create it? How have you used…?
Self-regulation*	You really persevered with … You worked really hard at that… Well done for carrying on when you found it difficult I particularly like the way …	What were the reasons you chose to do it that way? What was your favourite part and why? How can you check what you have done?
Praise/Self*	Good work. You are making good progress. Well done.	Can you do better than that? Keep up the good work.
Instructional	This is the set-up of the equipment. That is the circuit diagram to use. Right, going through the homework, you should all have a purple pen.	Can all of you just look here? Have you got a Paper 2? Can you get some results from another group?
Other	Yes, you can collect the cards in that's fine. Take them if you need them.	Do you need more (glue) than that?

Table 3.3 Exemplar categorisation of oral interaction types

The following is an extract given as an example of how a comment, or part comment (Hargreaves, 2013, 2014) would therefore become coded, with a short explanation of why that code was appended:

All interactions were coded separately where there was held to be a distinction to the previous utterance:

Is that the furthest forwards he goes?	PpG
prompting a girl to process a response	
Well what do you think?	PpG
prompting the girl to think further	
If that's what you think then use that number	PpG
prompting the girl to consider the number	
[girl answers in negative] well then, we'll talk about it	TsG
stating that they could discuss the answer	
Don't be frightened of making a mistake,	RpG
prompting a self-regulation response	
this is part of the learning process	RsG
acknowledging self-regulatory habits	

[Teacher 2, lesson 7, 29.50-30.05m]

If there was a prolonged period of, for example, probing thinking during the working out of a physics problem, to conflate these consecutive interactions could be used to give a time indication of how long they had occurred, but it would not convey how many interactions took place during that time. Additionally, an oral interaction need not necessarily have been a grammatically correct sentence; if a student asked if their answer was correct, and the teacher confirmed with a 'yes', then this would be coded as 'Ts' [Task-related statement]. This was deemed to be a more valid representation of the actual dialogue. Throughout the coding, if a sentence, sentence fragment or clause rendered it separable from the previous remark, then it was individually coded.

3.8.2.2 Field notes

Bryman (2016:440), remarking upon the 'frailties of human memory', asserted that observations were also useful in the form of field notes, where at the descriptive level they formed quick fragmentary jottings, *in situ* short transcriptions, short pen portrait moments of some of the student focus group individuals, and descriptions of the setting and context of the physics lesson (Cohen *et al*., 2017) adding a richness to the description as well as ecological validity. That these notes were made on the specific determined themes of the observation schedule enabled a 'chronolog' (Lincoln and Guba, 1985), an important aspect of a semi-longitudinal study seeking to investigate changes over time. Bryman (2016) advocates making clear and vivid notes, so that one does not later question what was meant by the note. Additionally, field notes proved to provide a space for recording some personal reflections as they occurred. Most often however, they were used as a contemporaneous record of what the teacher had given as verbal feedback, so that the in-the-moment coding could be subsequently checked. An example of field notes is given in Appendix 5.

3.8.3 Interviews

Interviewing is one of the most common methods used in small-scale educational research. Drever declares: *'this is not surprising…when you want to get information, canvass opinion, or exchange ideas, the natural thing to do is talk to people'* (2003:1). Mears however asserts that is more than questions and answers:

> '*in-depth interviews are purposeful interactions in which an investigator attempts to learn what another person knows about a topic, to discover and record what that person has experienced, what he or she thinks*

> *about it, and what significance or meaning it may have*.' (Mears, 2012:170]

The interview experience must be designed so that the questions asked result in data that the researcher requires to answer the research questions, and these are not identical. This transactional nature of interviews recognises that subjects are able to offer valuable insights not only into their thoughts and feelings but convey what significance these might have. Consequently, they are a direct method for providing rich, contextualised data; 'thick descriptions' (Geertz, 1973; Lincoln and Guba, 1985) of social situatedness. As a social, interpersonal encounter (Cohen *et al*, 2017) it represents a collection of data that will 'allow insightful analysis and produce defensible findings' (Mears, 2012:171), thus representing its position within a predominantly interpretive paradigm. The purpose of the interview may be to gather factual information, preferences, opinions of circumstances or explorations of motivation and experience (Drever, 2003).

Since other useful purposes of interviews are to develop or test hypotheses, to follow up unexpected or survey results and to contribute to validation of other data collection methods (Cohen *et al.*, 2017), a skilled interviewer can stimulate the flow of data in the way that they 'handle' the interviewee (Drever, 2003). Johnston and Toplis (2012) advocate a semi-structured interview format, which balances pre-determined topics or enquiries necessary for specificity with the flexibility to incorporate open and closed questions for further probing and clarification. Here, Mears suggests that while the interview guide frames the areas to be investigated, there is an inherent unpredictability in that one cannot 'be certain exactly where the answer will lead' (2012:172). In choosing to deviate from pre-determined questions, the interviewer must be sure there is merit in doing so, as this may affect comparability between groups. Cohen *et al*. (2017:510) provide a useful overview of different types of interview, and the strategy in this study is characterised as an 'interview guide' approach

A number of authors highlight the need for an interviewer persona demonstrating candour, interest, openness to responses and sensitivity (Drever, 2003; Mears, 2012; Cohen *et al.*, 2017). Challenges posed by interviewing include the demands of time in both undertaking, transcribing and analysing the data, and the possibility of inadvertent omission of topics through increasing flexibility. However, the semi-structured interview, lying on a continuum between highly structured question and response categories that have been determined in advance and an open one where questions emerge from the immediate context (Cohen *et al*, 2017), offers a compromise yielding both contextualised thick descriptions balanced with efficient collection and analysis costs (Johnston and Toplis, 2012).

3.8.3.1 Student Interviews as a Focus Group

To fully explore how feedback practices and interactions affect students, interviews were collected with selected students of the identified sub-sets in a focus group. Section 3.4 above has highlighted how many feedback studies had failed to incorporate the student view, which was required for thick descriptions of student self-belief also. Focus groups are an interactive interview method in which the 'group opinion is at least as important as the individual opinion, and the group may sometimes take on a life of its own' (Gibbs, 2012:186). They are often used to complement other data collection methods. Certain students were approached for interviews, based on both pre-intervention survey results and intervention teacher recommendation. These focus groups were audio-recorded, transcribed, and the narratives structured to show interview contents within identified themes. The semi-structured student questions are shown in Appendix 6.

Semi-structured interviews took place with purposefully sampled students selected through analysis of the survey at both pre- and post-intervention (see Figure 3.3). The structure of each phase of interviews were similar in that it began with gentle 'social' questions to develop confidence and rapport; it then comprised of some hierarchical and probing elements, with more closed questions, which could then be expanded to elucidate articulation of beliefs and responses and gain more insight into context and experience. Cohen *et al*. note that 'the interviewer is responsible for considering the dynamic of the situation, for example, how to keep the conversation going' (2017:519). Whilst focus groups offer challenges in terms of potential for confidentiality issues, conflict within the group, difficulties in managing group interaction, and possibly shallower data, Gibbs (2012) highlights the empowerment and even joy that participants can feel in talking openly about a topic. Ethically, using a focus group means that the researcher cannot assure confidentiality, as an individual may subsequently reveal the contents of the group's talk, and the interviewer must be alert for an 'imbalance of power' that students may perceive (Cohen *et al*., 2017) and respect the emotional and social dimensions of the interview.

3.8.3.2 Teacher Interviews and coaching conversations

To enable RQ4, and explore RQ5, it is necessary for the intervention teachers to undergo some continuing professional development (CPD). To ensure that the intervention of training to utilise more self-regulation feedback becomes not only introduced but embedded in their practice, the researcher met with the intervention teachers at approximately two-monthly intervals from the initial CPD to 'debrief' how the use of the feedback is progressing. This was in the form of a semi-structured interview, taking the form of a 'coaching conversation' along the 'GROW' model of coaching (Whitmore, 2002). This is a simple sequence of types of questions identifying the 'Goal', the 'Reality', the 'Options', and the 'Wrap-up'. The nature of these questions are explored through Chapter 7.

Taken together, the focus groups and teacher interviews offer 'thick descriptions' (Geertz, 1973; Lincoln and Guba, 1985), representing a detailed account of a social setting which offers a basis for others to make a judgement about its transferability to other situations; 'a powerful and user-friendly summary which can serve as a guide to professional action' (Bassey, 2001:5). Thick descriptions can include recording non-verbal communication, commentary in field notes, transcriptions, pen portraits of teachers, reconstruction of events and represent the complexity of the situation (Cohen, Manion and Morrison, 2017).

3.9 Analytical Approach

The study has been designed by engaging with existing theoretical propositions and has been characterised as a blend of experimental and evaluative types. Since the outcomes relate to the students' self-belief systems, and the embedded units are the students and their teachers, the collection of both fine-grained quantitative and qualitative data with attendant coding of constructs of potential interest support a semi-inductive strategy to the analytic technique of explanation building, especially with reference to the interviews.

Analysis of the survey data was required to answer RQ1 and 2, and this is discussed in Chapter 4. The observations are critical in answering RQ3 and 4, and the schedule was originally developed as a relatively simple tally for event occurrence. However, by introducing another variable (that of time recorded down the side), the schedule remains event-driven, but extends the information gained from teacher patterns of interaction to looking at those patterns over time. In analysing these, it will be possible to obtain the frequency, reported as proportions or percentages, enabling potential comparison across classes of the same teacher(s) and between teachers also.

To gain an answer to RQ5 requires a triangulation of evidence from multiple sources: survey, student and teacher interviews, and observations. Whilst it is valuable to gain a sense of 'voice' from the students and their teachers as to their notion of impact from the intervention, corroboration will be provided primarily by the post-intervention survey in proving student self-belief change for the student intervention group.

3.9.1 Organising and presenting the data analysis.

Cohen *et al*. (2017) discuss five ways in which the data analysis may be organised and presented:

a) by groups, whereby the data is automatically grouped, enabling patterns and themes to be seen. However, the collective responses of an individual become dispersed within the group.

b) by individuals, in which the total responses of one participant are presented before moving onto the next. The researcher must therefore subsequently perform a secondary analysis to look for patterns across the individuals.

c) by issue, sometimes decided pre-ordinately. The analysis would thus disregard data deemed irrelevant to the issue.

d) by research question, in which all data from various streams are collated to provide a collective answer [*ibid*]

e) by instrument, in which the results of each instrument are presented in turn. Further analysis is thus required to analyse the content of responses across these instruments, by issue and/or people.

Given the bimodal nature of the research questions, the analysis will proceed by *instrument*, although there are some aspects of group analysis for the survey data, and individual analysis for the classroom observations.

3.9.2 Analysing and presenting the survey data.

The initial data collection is survey-driven and seeks to collect baseline information for RQ1 on student self-belief systems in physics. Surveys were completed by both intervention and comparison groups and were collated into a grid that enabled subsequent analysis. The questions are closed in that although a Likert scale offers a range of response, each response can be coded along the range. Thus, the range of responses in the different categories can be counted and proportions of self-belief constructs calculated. These tallies were then analysed through SPSS to enable the findings detailed in Chapter 4.

The survey therefore is a key part in determining the meaning-systems that individuals have created for themselves as outlined in the self-belief theoretical framework (section 2.4), and a subsequent return to the survey at the end of the study will track whether those self-beliefs have changed for the students. Whether it is the CPD intervention that has impacted on those meaning-systems by way of the social persuasion of the verbal teacher feedback will be answered in RQ5.

3.9.3 Ensuring a systematic approach to the observation data analysis.

Becker and Geer (1960, cited in Cohen *et al*, 2017:665) outlined a systematic approach which can include comparing different groups simultaneously over time, matching the responses given in interviews to observed behaviour [surveys and field notes, *my insertion*] and calculating the frequencies of occurrences and responses. It was important to guard against subjective influences on the part of the researcher; these could include data overload, over-emphasis of confirmatory data, misplaced confidence of judgement, inconsistency of analysis and mistaking co-occurrence for association (*ibid*).

Analysing content ‘begins with a sample of texts (the units) and defines the units of analysis (e.g., words, sentences) and the categories to be used for analysis’ (Ezzy, 2002:83). The categories for coding the observations had largely been predetermined by the theoretical perspectives of previous authors, however since continuing data did not produce additional insights or codes, saturation was deemed to have been reached (Ezzy, 2002:93). Once coding and counting of all observations was completed statistical analysis and quantitative methods were applied, leading to an interpretation of the results. (Cohen *et al.,* 2017).

3.9.4 Interview data analysis

The purpose of the interview data was to provide the opportunity for in-depth exploration from a range of individuals. Using a largely inductive approach since there was no predetermined framework for the discourse analysis (Bryman, 2016), the pupil interview data was transcribed and analysed to obtain more nuanced descriptions from the participants in the development of themes, in an attempt to strengthen causal inference.

The interview data therefore provided an essential narrative of rich detail that explored both the baseline data and any potential changes from the perspective of the students, and of the teachers as well, in answering RQ4 and 5. In analysing transcripts of semi-structured interviews, the responses were reorganised, categorised and summarised to match the research questions (Drever, 2003). Rules were developed for defining and extracting categories from the data, as these were not predetermined. Finally, summaries were generated from the teacher and student perspectives, which involved the selected student groups and individuals, and these are presented as themes in Chapters 6 and 7.

3.10 Procedures to address validity and reliability

Lincoln and Guba (1985) and Guba and Lincoln (1994) assert that qualitative studies should not be subjected to the same criteria of quality as quantitative research, offering instead the notion of 'trustworthiness' as more suited to the complexities of the social world, arguing against the positivist standards of reliability and validity (Bryman, 2016). Trustworthiness comprises four aspects: *credibility* (paralleling internal validity); *transferability* (paralleling external validity); *dependability* (paralleling reliability); and *confirmability* (paralleling objectivity). However, quasi-experimentation is often described in more positivist standards, and it would be unhelpful here to use both positions, so the 'classic' notions of validity and reliability will be used, alongside a presentation of techniques employed to assure these. Validity in quantitative approaches utilises sequentially, appropriate instrumentation, careful sampling, appropriate statistical treatment of data, non-selective use of data, and presentation of the data without misrepresentation (Cohen *et al*, 2017). In qualitative research, validity is not 'a discretely identifiable element of any research project, which is capable of being located at multiple and specific stages within the research' but embraces the honesty, depth, richness and scope of data, the extent of triangulation, and the objectivity of the researcher (Winter, 2000).

3.10.1 Quasi-experimental design issues

Since a quasi-experimental approach does not utilise random assignment and/or controlled manipulation of variables, it is very important for the research design to protect against internal and external validity threats. Eight *internal* threats as originally defined by Campbell and Stanley (1966, cited in Grant and Wall, 2009) include:

History – events or influences on the test subjects, for example, in both schools, ‘mindset’ was being culturally embraced by the senior leadership;

Maturation – test subjects may themselves change over the life of the study;

Testing – the subjects may become sensitised to the aims of the study from the pre-test, or that the pre-test may influence the subjects on subsequent tests (the students were not aware of the feedback focus of the research);

Instrumentation – Changes in instrumentation (such as the survey) during the study may affect what is being measured and how it is measured. The surveys differed slightly between schools in the physics content of two questions, as each were following different examination specifications, otherwise they were not changed to assure survey stability. The Teacher 1 (intervention) cohort undertook the pre-intervention survey twice, separated by a week to investigate survey stability, and these showed ‘very good’ agreement;

Statistical Regression – regression to the mean may present a concern in studies with extreme scores (not undertaken);

Selection – If the groups are functionally inequivalent at the beginning of the study, comparison of results will be biased (schools regarded these as comparable classes);

Experimental Mortality – if test subjects (or teachers) leaving the study, the equivalence could be affected (some data had to be removed as students completed one bit not both surveys);

Selection Interaction – where the selection method interacts with one or more of the other threats to impact results.

By using multiple groups which are comparable, these threats to internal validity are minimised. Trochrim (2007) asserts that the only real multiple group threat to internal validity is if the groups are not comparable before the study and suggests examining the groups on pre-test measures to determine similarity, and 'make some judgement about the plausibility that a selection bias exists'; the schools concerned identified these as comparable groups.

Cook and Campbell (1979, cited in Bryman, 2016) identified five major threats to the *external* validity, and hence the generalizability of the study:

Interaction of selection and treatment – to which groups could a finding be generalised? Are the test subjects representative of a Year 10 science population?

Interaction of setting and treatment – would a finding be generalised to other settings?

Interaction of history and treatment – for example, would the study have the same results if conducted in a different time in the academic year?

Interaction effects of pre-testing – When test subjects become sensitised to the nature of the investigation through the pre-test

Reactive effects of experimental arrangements – test subjects' awareness that they are participating in an investigation may influence how they respond to the intervention.

The groups were all Triple Science, mixed ability within the upper range, in which males and females are represented. The actual nature of the intervention was not fully disclosed to the students so that they were less likely to become sensitised to the nature of the study. As a former secondary physics teacher, these students appeared typical of a Year 10 population, and is some justification that findings could be generalised to other settings.

In terms of quasi-experimentation, reliability refers to ‘the extent to which a measurement or an experimental procedure elicits consistent interpretations about the construct that it sets out to measure’ (Rogers and Revesz, 2019), and the double-case approach underpins efforts to secure this consistency.

3.10.2 Case study design issues

Yin, a noted case study author, applies notions of reliability and validity to case study research (2014), rather than trustworthiness. It seems that most objections to these standards relate to issues involving data collection methods, however the largest discussion centres on the external validity or *generalisability* of the case study approach; how can a single case be representative of other settings (Bryman, 2016)? The design of this study incorporates two such cases to attempt to mitigate this doubt.

Perceived weaknesses of the interview method include issues about verbatim transcriptions (it may be appropriate to use selective transcription), bias on the part of interviewee and interviewer, and the scale of interviewee reticence (Cohen *et al*., 2017). The interviewee may misremember, consciously or unconsciously modify the facts, say what s/he thinks the interviewer wants to hear, or indeed be unwilling to talk. Should the interviewer detect known bias, s/he might use (for example) classroom artefacts to provide information instead (*ibid*.) Adhering to pre-decided questions may help guard against interviewer bias in leading the interviewee, since this would be contrary to the intended purpose of the interview.

Verbatim transcripts of interviews are held to be close to a true record of an interview, though one must recognise that elements such as body language and tone of voice are lost (Drever, 2003). If selective transcripts are used, or summaries made, there is a danger of systematic distortion which may be safeguarded against by staying faithful to the language of the interviewees and by having a colleague check transcript summaries. It is necessary to develop and define valid categories that will match viewpoints across the range of interviews, and reliability is obtained when analyses across them are consistent.

Johnston and Toplis (2012) argue that detachment and objectivity are crucial in non-participant observation to avoid issues of external validity. It is possible that the presence of the observer may in some way impact the reactive behaviour of participants (see section 3.10.3). Validity concerns often relate to bias, in that ‘the observer may selectively attend’ (Cohen *et al*, 2017), which can lead to a distortion of data. Since observational responses had in effect been pre-determined by categorisation, the researcher must guard against an expectancy effect in that the researcher was observing on order to test the hypothesis, which would shape data collection (*ibid*.)

Observation validity is based on the schedule actually measuring what it purports to. Since the sole purpose of the initial observations is to collect information about the levels and proportions of feedback in use in these teachers’ classrooms, the categories selected have appropriate face-validity. However, in moving the observational focus to that of selected sub-sets of students, it will be necessary to validate the observational output with the views of the participants under observation, which helps to establish theoretical validity (Simpson and Tuson, 2003).

It should be clear what has been pre-categorised as a significant event, so that inference is kept low, as this can lead to misinterpretation. In measuring 'what is said', it is necessary to pre-determine the boundaries around the different types of feedback, so that they can be quickly recognised and recorded in situ. Exemplar phrases were given above in the analysis plan section. The data gathered were treated quantitatively and analysed in simple statistical form to enable subsequent comparison. A University colleague unconnected with the study acted as a reliability rater to check several coded samples, together and subsequently independently. The number of variables made an application of Cohen’s Kappa for inter-rater reliability ineffective, however the less statistically desirable method of percentage agreement (which does not take into account chance agreement) gave a range of between 0.88-0.97, which is classed as ‘very good’.

Triangulation is used for validation of data by cross-verifying the data obtained from more than one method and is especially useful in complex non-laboratory social situations (Bryman, 2016). It is therefore used specifically for demonstrating concurrent validity i.e., data gathered using one instrument should correlate highly with data gathered from another. This study employs ‘methodological triangulation’, since the same method is used on different occasions, as well as different methods on the same object of study. There is also an element of ‘time triangulation’ by collecting data from the same and several groups at different points in time. This enables the stability of observations and emergent themes arising in interviews to be taken into consideration as the intervention proceeds over time.

3.10.3 Role of the researcher

The researcher must establish the boundaries of the study, by firstly developing a data collection matrix in which the amount and timing of collection is specified. This thereby sets a timeframe as a boundary, in order to trigger a clean ending point (Yin, 2015).

On the whole, the researcher was not involved with the student participants on an extended basis. There were some interactions during the preliminary and post-intervention focus groups. These were semi-structured and conducted in a manner that respected potential power imbalances, and sought to put the interviewees at ease, similar to the tone of the coaching conversations

There was an extended period of 5-6 months where the researcher was present in classes for essentially non-participant observations of students and teachers. Even as a non-participant, it is possible that the students and teachers may have become more used to the researcher’s presence over the life of the study.

3.11 Ethical issues

'Ethics concerns that which is good and bad, right and wrong.' (Cohen *et al*., 2017:111), and in terms of educational research, concerns what the researcher should and should not do at all stages of the research design, both in that research and regarding other individuals (Creswell, 2013, Hammersley, 2017). Ethical consent was sought according to established guidelines and was granted by the University of Leeds' ethics committee before the research commenced; the ethics terms of reference are shown in Appendix 7.

Cohen *et al*. (2017) list a field of issues occurring in recent ethics literature and note that these are not black and white in educational settings, as do BERA (2018) who acknowledge that 'few ethical dilemmas have obvious or singular solutions' (page 7). BERA (2018:5) also summarise a list of responsibilities for researchers to adhere to a principle of respect for: participants, stakeholders, the community of educational researchers, publication and dissemination, and the researchers' wellbeing and development. In more detail, Hammersley (2017) outlines some of the philosophical diversity within research ethics in that there are a number of epistemic and practical values that can inform the researcher's work. He identified a set of three principles: minimising harm, protecting privacy and respecting autonomy, whereas Punch and Oancea (2014) outlined four: autonomy, trust, beneficence [concerning issues of harm] and acceptability, and these latter will be used to discuss the steps taken to maintain an ethical stance throughout the research.

3.11.1 Ethical autonomy

Permissions were gained from the schoolteachers, the students, and the students' guardians as well as the head teacher as 'gatekeeper'. To this end, briefing letters with the request for consent were sent to all potential participants and their guardians (Appendices 8-13), although these varied in both the readability (for the students) and the amount of information about the proposed study (see section 3.11.4 below). The researcher spoke with each class at the start of the study to explain in person and allow the students to ask questions (there were none).

The student participants were aged 14-15, and thus were classed as a vulnerable group. The University of Leeds guidelines stated under the Gillick principles that provided the students were able to show competence, they would be able to consent themselves. However, the decision was made to also inform the students’ guardians also, in case they did not wish their child to take part, and this request would have been honoured. There was however to be no pressure to sign the consent forms, and few students were ultimately involved in the interviewing process. All documentation was kept in a secured cabinet.

3.11.2 Trust

Data protection legal requirements as well as ethical concerns necessitated that all data had to be securely stored via password protection, and on an encrypted device. When the audio-recording had been downloaded and stored, it was erased from the recording device. The participants and schools had their confidentiality, anonymity and non-traceability guaranteed. All participants were reassured of their right to withdraw from the study at any time and the students received full disclosure at the end of the study. No participant withdrew during the study although there was some experimental mortality, and that data was excluded.

3.11.3 Beneficence

There was no potential for physical harm for any participant within the study, however embarrassment caused by asking questions, or risk to reputation may be construed as causing harm (Hammersley, 2017); the first of these is discussed in the Interviews section (3.8.3) concerning the sensitivity of the interviewer. It was important to respect potential power imbalances e.g. when interviewing. Cohen *et al.* (2017) describe interviews with children thus: ‘It is important to understand the world of children through their own eyes rather than the lens of the adult’ (p.528) and note several deliberate actions that the researcher can undertake to make the children feel more comfortable. The second potential harm, to participant reputation may be circumvented by observing confidentiality and non-traceability (Hammersley, 2017). However, there remains an ethical dilemma potentially influencing how openly (or not) teachers’ practices are reported.

A further tension must be recognised in undertaking experimental approaches on students’ educational experiences. Should the intervention be successful, would the comparison group have been disadvantaged in not receiving it? To circumvent this, the researcher offered to widen the professional development to the science department subsequent to

The visits were conducted in such a way as to afford maximum respect to the school sites, and to the teachers’ primary responsibility of teaching the students. There was therefore minimum disruption, and interaction with the students was kept short so as not to unduly waste their time.

3.11.4 Ethical Acceptability

Punch and Oancea (2014) notes the conflicting demands sometimes attendant on research. In this study, the student and guardian letters discussed the importance of good interaction and communication within physics, students’ attitudes towards physics and the value of physics to the economy but did not disclose the title of the study. This was because the researcher was looking for honest responses to the change in feedback, rather than contrived responses, and should the participants have been aware of the purpose of the research, they may have acted differently, and the results would have become biased.

3.12 Chapter Summary

This chapter has outlined the philosophical and methodological underpinning of the study and has justified the research design and data collection methods arising from a thorough review of both research methods literature and the investigation decisions taken by studies influencing this research.

The complexities of the layered research design of a semi-longitudinal, quasi-experimental approach to test the efficacy and impact of a teacher professional development intervention on verbal feedback has been outlined. A thorough pilot study enabled decisions to be made about both survey design and observation strategy. The review of feedback typologies in Chapter 2 supported the construction of a three-level typology for the purpose of this study that allowed teacher-student interactions to be captured and coded. A collection and analysis plan was shared, including aspects to protect the validity and reliability of instrumentation, data capture, researcher behaviour and ethical considerations. The subsequent chapter now explore the findings arising from the analysis of the data, commencing with the student self-belief surveys, in a quest to answer research questions 1 and 2; establishing what beliefs were held, and ascertaining whether these changed for the intervention students over the life of the study.

Chapter 4 Findings from the pupil surveys

This chapter presents the findings from the pupil surveys, administered to all groups at a time of pre-intervention to Groups 1 and 2, and again post-intervention. As described in Chapter 2, the survey was designed to explore the range and depth of student self-beliefs in order to answer research questions 1 and 2:

1. What are the students' self-belief systems as they enter Key Stage 4?
2. Do the students' self-belief systems in physics change over the duration of the intervention?

The survey is comprised of grouped categorical questions on self-efficacy, self-concept, anxiety and mindset, answered on a Likert scale of response. The scale is ordinal, since there is not an equal distance between each point of the scale, and these were subjective and individual student responses. Ordinal data are considered non-parametric (Cohen *et al*, 2017), since assumptions had not been made about the population; indeed, the characteristics of the student population in this sample were previously unknown. The validity of the survey items, and survey as a whole has been discussed in Chapter 3.

In sum, the validity and reliability measures taken indicate that the survey tool was both reliable and valid; it thus both identifies the self-belief constructs that it was designed for and is repeatable. Both the intervention and comparison groups yielded results that enabled the researcher to categorise the status of the self-belief constructs at the start of the survey; measures were obtained for self-concept, self-efficacy, anxiety and mindset.

There were positive changes in three of the four self-belief systems being surveyed for the intervention group; self-concept and self-efficacy in physics increased, whilst anxiety decreased. Mindset showed a very slight increase and could be regarded as stable. For the comparison group, self-concept decreased over the duration of the survey; this suggests that as a group, the students believed they were 'worse' at physics. This correlates with the change in anxiety for the comparison group, which increased (Ahmed, 2019; Lee, 2009, Morony *et al*, 2013). However, self-efficacy increased for the comparison group, which suggests, in conjunction with the self-concept finding, that they felt they could perform certain named tasks well, but not that this made them feel more disposed towards the physics domain. This increase in self-efficacy was not as large as that in the intervention group. Finally, the mindset construct showed a very slight increase; that is, it moved to a position of more individuals indicating an incremental theorist position, however this was not significant. We will now turn to a defence of the survey reliability before a more detailed presentation of the grouped constructs.

4.1 Reliability calculations

The same survey is used for both pre-intervention and post-intervention states for stability of reliability [The only item which differed between the groups was Question 9, an item test of self-efficacy, since the groups at different institutions had not encountered the same topics]. A test of the stability of a survey is that it should yield similar data from similar respondents at different times, and this was piloted before the data collection commenced. Similarly, the internal consistency of the survey was modified through the pilot study, and Cronbach's alpha reliability coefficient calculated to be 0.84 at the time. The range 0.80-0.90 is considered to be 'highly reliable' (Cohen *et al*, 2017). As a further stability test of reliability, for Teachers 1 and 3 (the same institutional setting), the test was administered before and after a half-term (one week) holiday and means and Cronbach's alpha for constructs compared.

The internal reliability for all 20 items shows very low internal reliability [Cronbach alpha = 0.266], as all four constructs have been calculated together. However, when grouped thematically as constructs in Table 4.1, the Cronbach's alpha calculation shows much higher internal reliability:

Self-construct	Number of items	Pre 1/2 Cronbach's alpha	Pre 2/2 Cronbach's alpha	Post Cronbach's alpha
Self-efficacy	6	0.850	0.866	0.838
Self-concept	5	0.782	0.881	0.853
Anxiety	5	0.857	0.878	0.849
Mindset	4	0.793	0.825	0.838

Table 4.1. Reliability statistics (Cronbach's alpha) of pre- and post-intervention surveys by self-construct.

Measures of Cronbach's alpha thus indicate high reliability from both pre and post-intervention surveys. In addition, split-half reliability calculations were also calculated, again showing high reliability of the survey instrument. In Table 4.2, the self-efficacy items show the highest reliability of the four self-belief constructs being tested:

Construct	Spearman-Brown coefficient	Guttman split-half coefficient
Self-concept	0.863	0.839
Anxiety	0.846	0.802
Self-efficacy	0.873	0.869
Mindset	0.851	0.848

Table 4.2. Split-half reliability statistics of pre- and post-intervention surveys by self-construct.

Although a total of 88 participants took part over the lifetime of the intervention, not all students completed both the pre-intervention and post-intervention surveys. Although counting all of the responses would have enabled a fuller picture of student self-belief to be obtained, for validity only those whose responses could be securely mapped throughout in SPSS were included, resulting in 33 participants in each of the intervention [IG] and comparison [CG] groups.

4.2 Summary findings

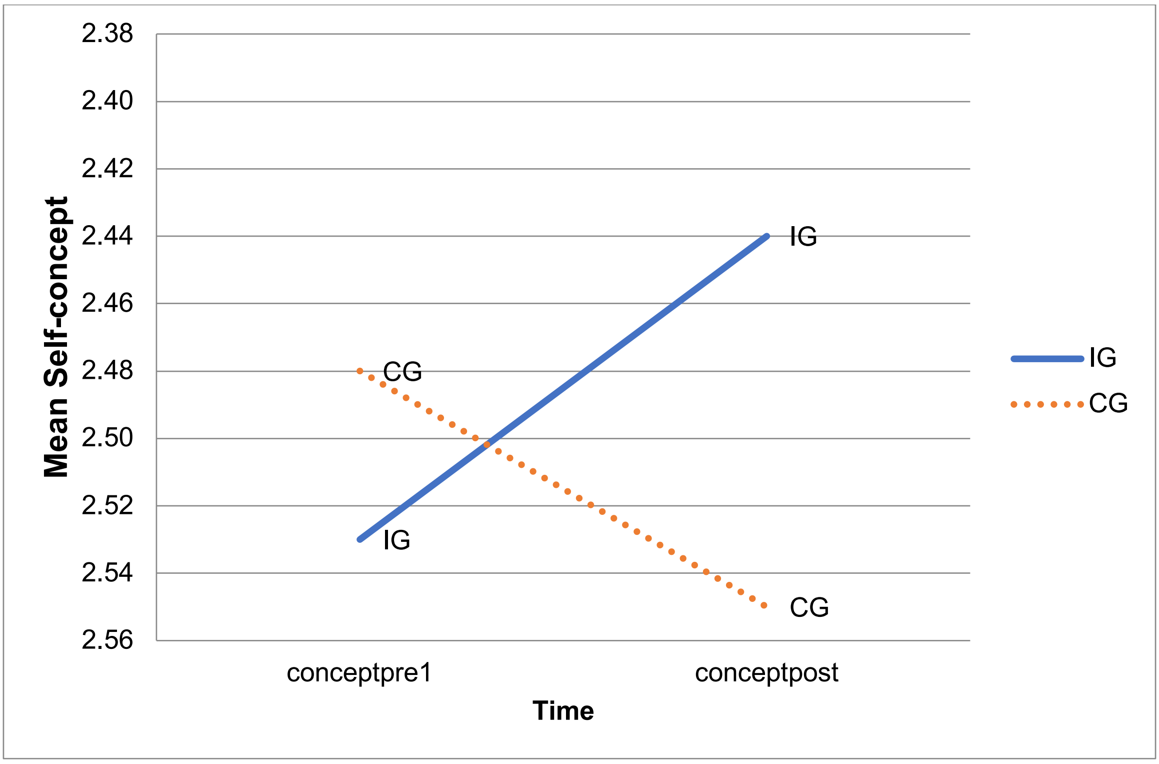
The exploratory nature of ordinal data analysis lends itself to a descriptive statistics presentation and the summarised findings of each construct, calculated as a mean as each participant’s sub-scale score, are presented below. The results are initially presented as grouped self-belief themes, and then more fine-grained analysis of items is considered. The pre- and post-intervention survey outcomes were conflated for the IG teachers (T1, T2) and CG teachers (T3, T4) and broad comparison summaries for pre- and post-states are shown in Appendix 14. The summary frequencies and cross-tabulations for individual survey questions are presented within the findings for separate items below.

For the line graphs below, which indicate the pre-test and post-test conditions, the y-axis represents the mean values of the scores of the items; Strongly Agree codes as a 1; Agree codes as a 2; Disagree codes as a 3, and Strongly Disagree codes as a 4, for all positively phrased items. Negatively phrased items are therefore reverse-coded. Thus, a line showing a downward trend indicates movement away from (for example) ‘Disagree’ and into ‘Agree’. Where necessary, scales were then inverted so that graphs would display as upwards trends, which is simpler to read when discussing positive outcomes. It is important to note that as these represent *two* data points, at the start and end of the study, the line is actually an interpolation between two points; a mid-study survey would have strengthened reliability and validity. However, as both the self-belief construct and the temporal nature of the study can be considered continuous variables, these two points are shown as trends. Occasionally, bar charts have been used to demonstrate the change in mean value of the construct sub-scale on the y-axis, however these values are mainly stated.

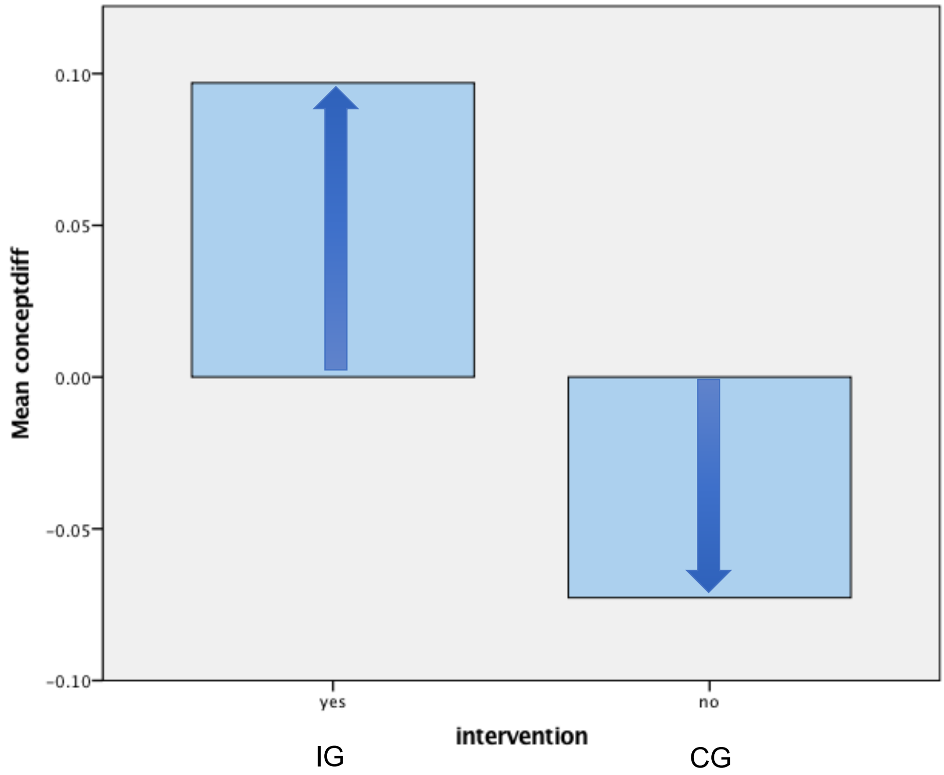
4.2.1 Grouped self-concept

Grouping the discrete variable constructs of self-beliefs together enables an analysis over each theme. Self-concept refers to a person’s perceptions and knowledge about themselves and tends to be domain- or subject-referenced in academic achievement situation; colloquially for students, self-concept is ‘how well they do’ in physics and is often expressed as a measure of themselves against others in their group (Tice and Wallace, 2003; Marsh *et al*., 2019).

Graph 4.1 shows a positive trend for the intervention group [IG], and a negative trend for the comparison group [CG]. This suggests that the IG improved their self-concept (with respect to physics) during the life of the intervention, whereas this appeared to decrease for the CG.



Graph 4.1 Line graph of mean change in self-concept over time for intervention and comparison groups.

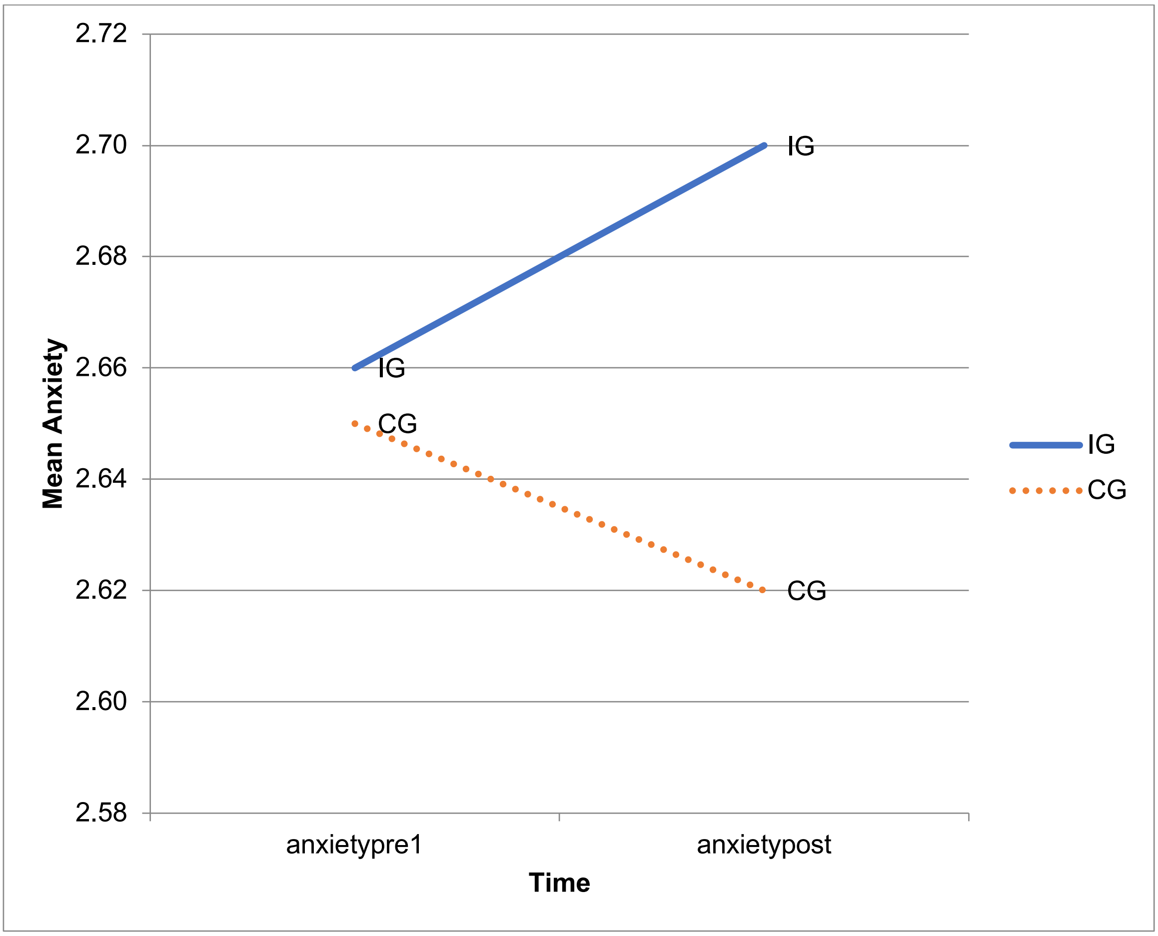


Graph 4.2 Bar chart of mean change in self-concept for intervention and comparison groups

Graph 4.2 indicates that the IG appeared to have a slightly larger positive gain in the mean of self-concept, whereas the CG have a smaller change, but negative with respect to self-concept in physics. The total mean change represented is 0.175, which is not significant. However, in a small-scale research study with qualitative insights, the apparent change will be explored with student perspectives.

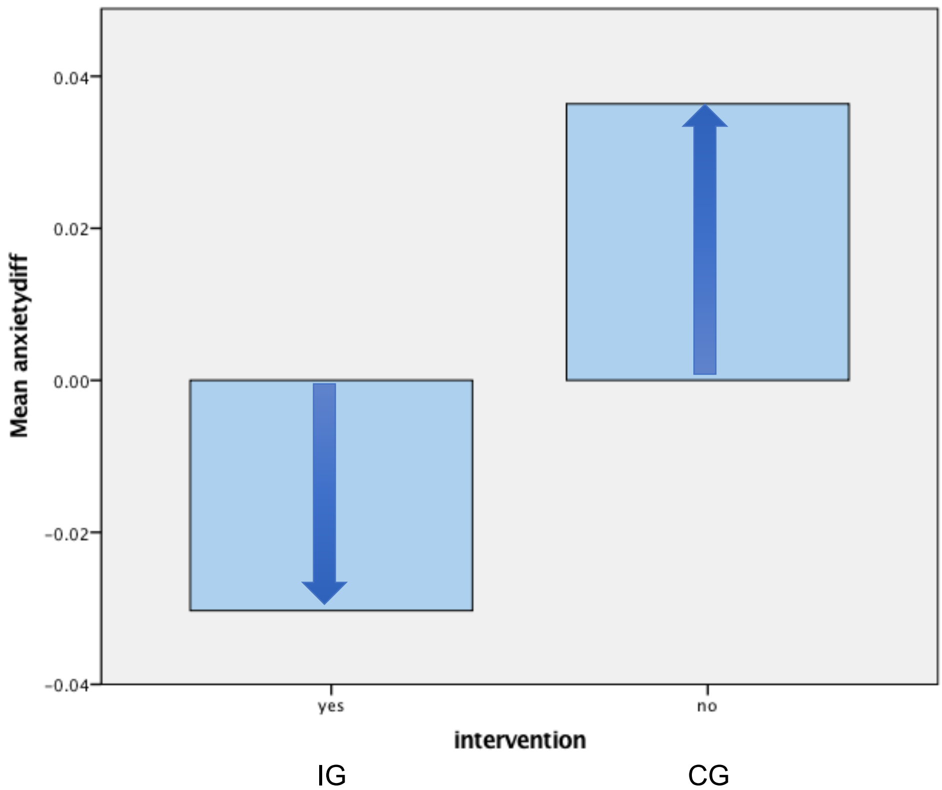
4.2.2 Grouped Anxiety

Science anxiety was first defined as a ‘diffuse or vague fear which arises in science learning situations’ (Mallow, 1978), and factors such as negative feelings arising from past difficulties, poor relationships with science teachers and self-messages perpetuating ‘I’ll never solve these problems’ can all combine to cause anxiety in science and reduce performance (Mallow and Greenburg, 1983).



Graph 4.3 Line graph of mean change in anxiety for intervention and comparison groups.

In Graph 4.3, the upward trend of the IG indicates that more disagreed with feeling anxious about their physics work as the life of the intervention progressed; conversely, the downward trend of the CG indicates that overall the physics anxiety within the classes increased. These were small mean amounts of change however, and this is represented in Graph 4.4; for this sample size, this is not considered to be statistically significant. It does support the view (Lee, 2009, Morony *et al*, 2013) that anxiety correlates *negatively* with self-concept.



Graph 4.4 Bar chart of mean change in anxiety for intervention and comparison groups.

4.2.3 Grouped Self-efficacy

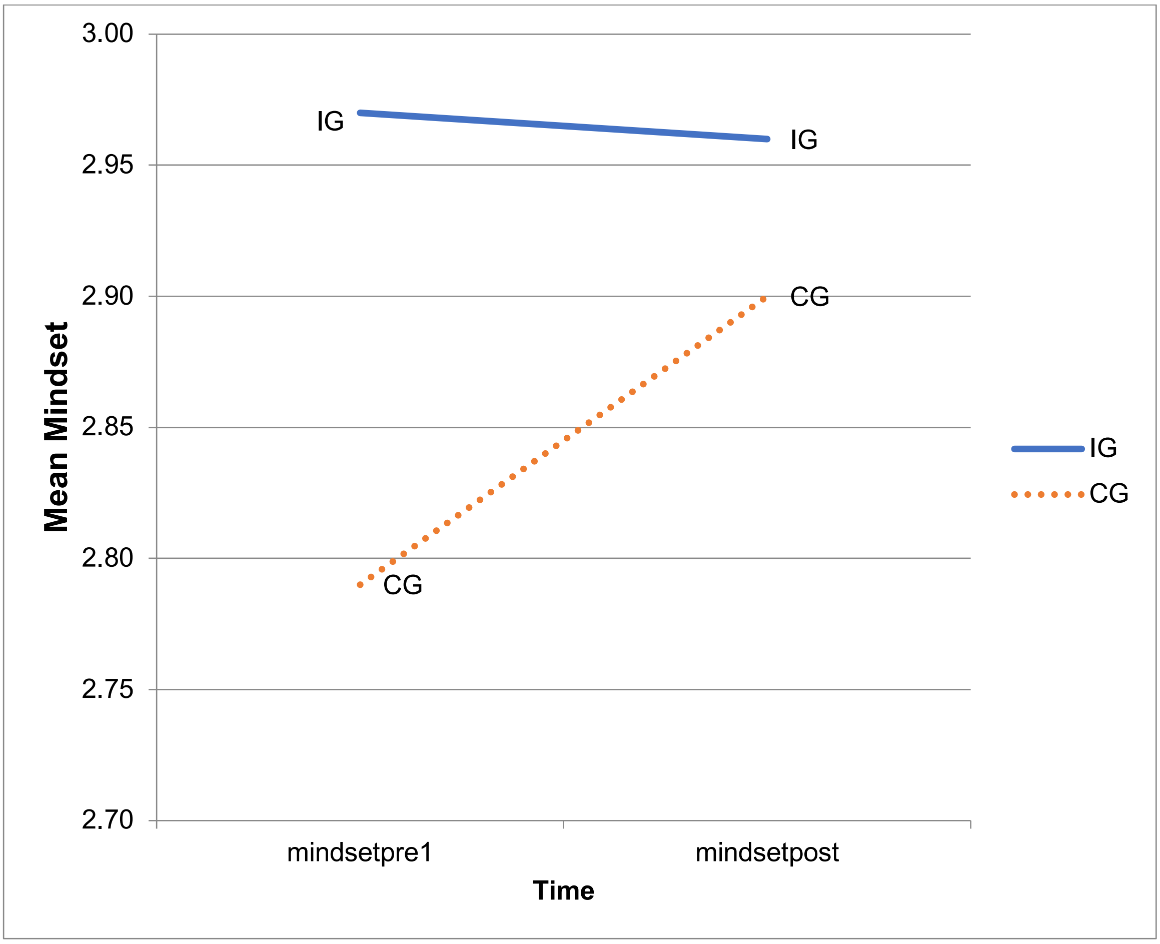
Graph 4.5 shows that the self-efficacy in physics for both groups improved over the life of the intervention, as both groups have decreased their mean numerical scores [=tending towards 'agree' conditions]; the vertical axis has been reverse-scaled to indicate this positive trend. The graph also suggests that the IG had a very slightly more positive self-efficacy starting point value, and that their gain in self-efficacy was marginally more pronounced, shown by the steeper gradient. The mean change in self-efficacy for the IG was 0.12, and 0.08 for the CG.

Graph 4.5 Line graph of mean change in self-efficacy for intervention and comparison groups.

Self-efficacy, being the judgement, that one can successfully perform a task (Bandura, 1997) is described as both domain- and context-specific, and of malleable temporal stability by Seon and Bong (2019). Unlike self-concept, whose dominant reference is past experiences (I have always done well in physics), self-efficacy, whilst derived from past mastery experiences (Bandura, 1977, 1997; Britner and Pajares, 2006; Usher and Pajares, 2008), is a forward-looking subjective conviction; the difference between ‘I am’ and ‘I can’.

4.2.4 Grouped Mindset

Both the pre- and post-intervention conditions show that there was very little change in the mean value of mindset for the IG [2.97 to 2.96], and this was at a value which indicated that as a cohort, they disagreed with the supposition that one could not change their basic intelligence to a high extent (Graph 4.6). The CG started at a slightly lower mean value [2.79] and through the life of the intervention appeared to have moved towards a similar view on intelligence, with a mean change of +0.11. It is possible that this is a maturation effect, and potentially linked to the CG's increase in self-efficacy, however the emerging limitations of mindset questionnaires being applied in school settings (Li and Bates, 2017, Lynch, 2018) will be discussed in greater depth in Chapter 8.



Graph 4.6 Line graph of mean change in mindset for intervention and comparison groups.

4.2.5 Broad comparison of intervention and non-intervention groups

Investigating the results between these two pairs of sample populations [T1+T2 (=IG) and T3+T4 (=CG)] reveals that there were some similarities and some variation between the intervention, and comparison samples. A summarised table of these is shown in Appendix 14a for easier reference. Summarising in this way indicates that the two intervention classes appear to have a higher baseline in self-concept, self-efficacy and lower anxiety, and exhibit less individuals coding for more fixed mindsets, although some of these differences are slight, given the sample size.

A similar comparison of the pairs of post-intervention surveys showed 16 items where the intervention group showed more positive outcomes than the comparison group, again included as Appendix 14b. The self-efficacy for groups 1 and 2 were higher in five out of six items, higher for self-concept in four out of five items, and lower anxiety in three out of four items. The results for mindsets were broadly similar. Overall, the two groups (from different schools) comprising the IG presented as a body of students with increasingly adaptive learning dispositions towards their learning in physics during the lifetime of the intervention.

Each group of variables is presented further below as itemised results to enable finer-grained observations to be made. Further examination will indicate whether there are differences in the way in which these students responded to the phrasing of questions, and if there was unequal contribution to the mean effects described above.

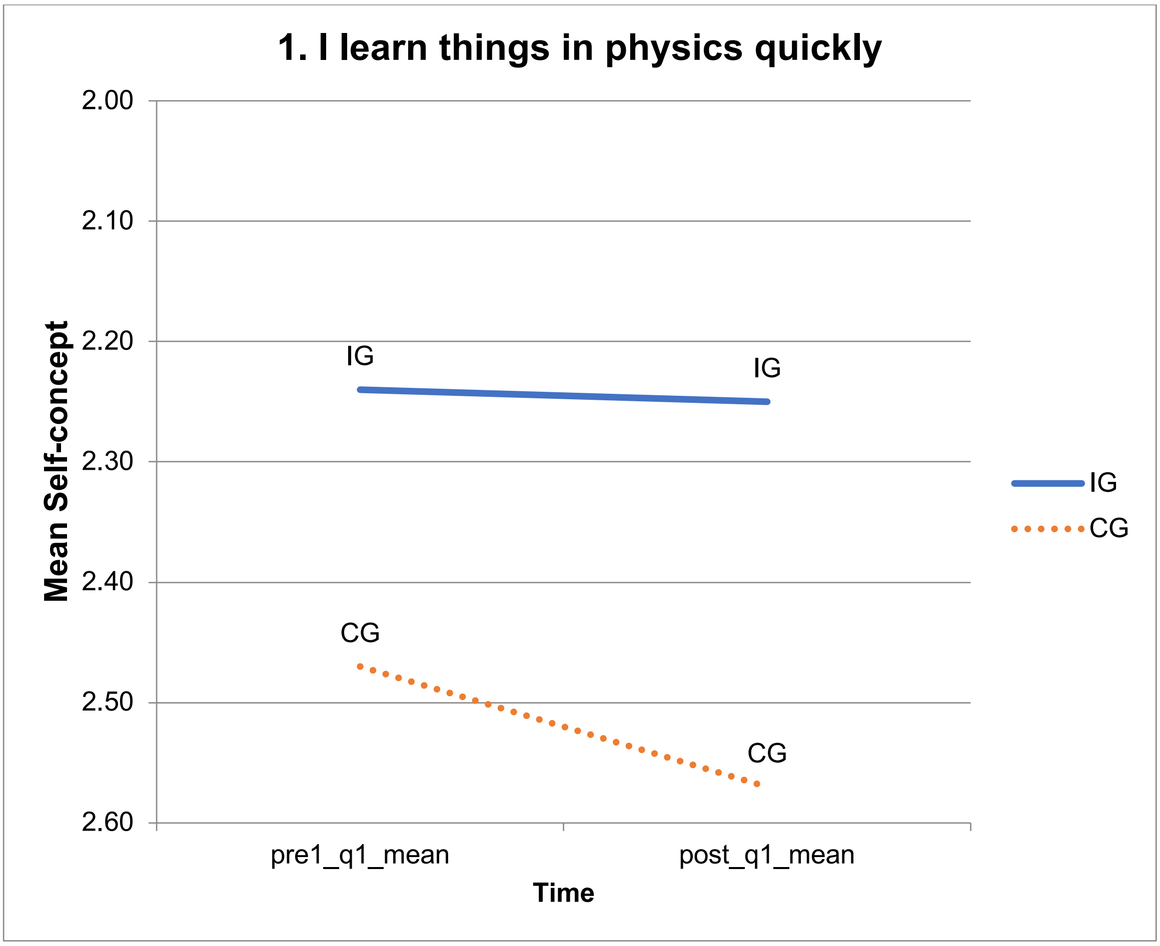
4.3 Findings from individual self-concept items

As a domain-specific construct deriving from the manner in which individual view themselves in physics (Tice and Wallace, 2003; Marsh *et al*., 2019), there is potential for questions relating to self-concept to return differing responses, as the students adjust their beliefs over time in reaction to new information being internally reorganised. Since the intervention spanned six months, there is potential for this to occur, as well as general maturation effects taking place. Although each change considered below in presenting the findings from individual items is small, they provide a consistent picture of students' perspectives within the physics classroom.

4.3.1. I learn things in physics quickly

[item 1: I learn things in physics quickly]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*5*	15.1	*2*	6.1	*2*	6.1	*2*	6.1
	Agree	*17*	51.5	*13*	39.4	*22*	66.7	*14*	42.4
	Disagree	*9*	27.3	*17*	51.5	*8*	24.2	*14*	42.4
	Strongly disagree	*2*	6.1	*1*	3.0	*1*	3.0	*3*	9.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.3 Pre- and post-survey results for item 1: I learn things in physics quickly.



Graph 4.7 Pre- and post-survey results for item 1: I learn things in physics quickly.

Although the CG post-survey score rose slightly from 45.5% to 48.5% [values given as ‘Agree’ and ‘Strongly Agree’ counts as percentages of the CG cohort] as seen in Table 4.3, the IG score rose from 66.8% to 72.6%. On this item, the values given by the y-axis on Graph 4.7 indicated that the CG decreased in self-concept; the IG had a higher self-concept initially, with a higher proportion agreeing overall, and this appeared to remain quite stable through the life of the intervention. This is a positive result for the IG compared to the CG.

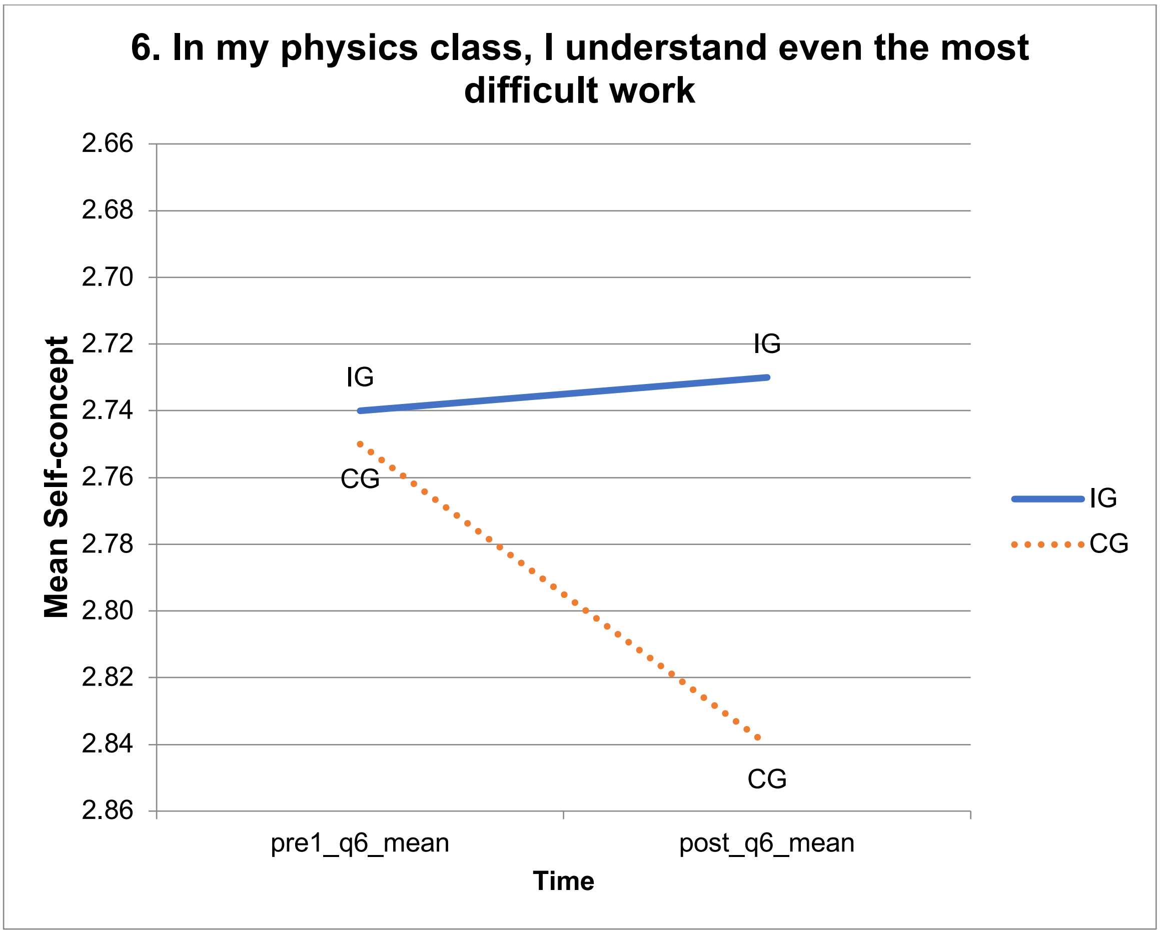
4.3.2. In my physics class, I understand even the most difficult work

The decision to use the word 'difficult' in a self-concept item question was not an easy one. 'Hard' and 'difficult' are words that have a culturally negative bias, whereas 'challenge' is culturally less negatively biased, and in certain contexts, achieves a positive bias. Difficult was a word used frequently in the 2006 *Girls in the Physics Classroom* report by the IoP, in terms of perception of students, and so it was included here.

Pre-test, 63.6% of the IG and 60.6% of the CG disagreed with this statement, although more students in the CG strongly disagreed (Table 4.4). However, the proportion of students agreeing to some extent that they could understand even the most difficult work dropped in the CG from 39.4% to 36.4%, and Graph 4.8 shows a decrease in self-concept overall for the CG. This is a positive result (for the IG, in comparison to the CG).

[item 6: In my physics class, I understand even the most difficult work]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*1*	15.1	*2*	6.1	*2*	6.1	*2*	6.1
	Agree	*11*	33.3	*11*	33.3	*10*	30.3	*10*	30.3
	Disagree	*15*	45.4	*13*	39.4	*16*	48.4	*14*	42.3
	Strongly disagree	*6*	18.2	*7*	21.2	*5*	15.2	*7*	21.3
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.4 Pre- and post-survey results for item 6: In my physics class, I understand even the most difficult work.

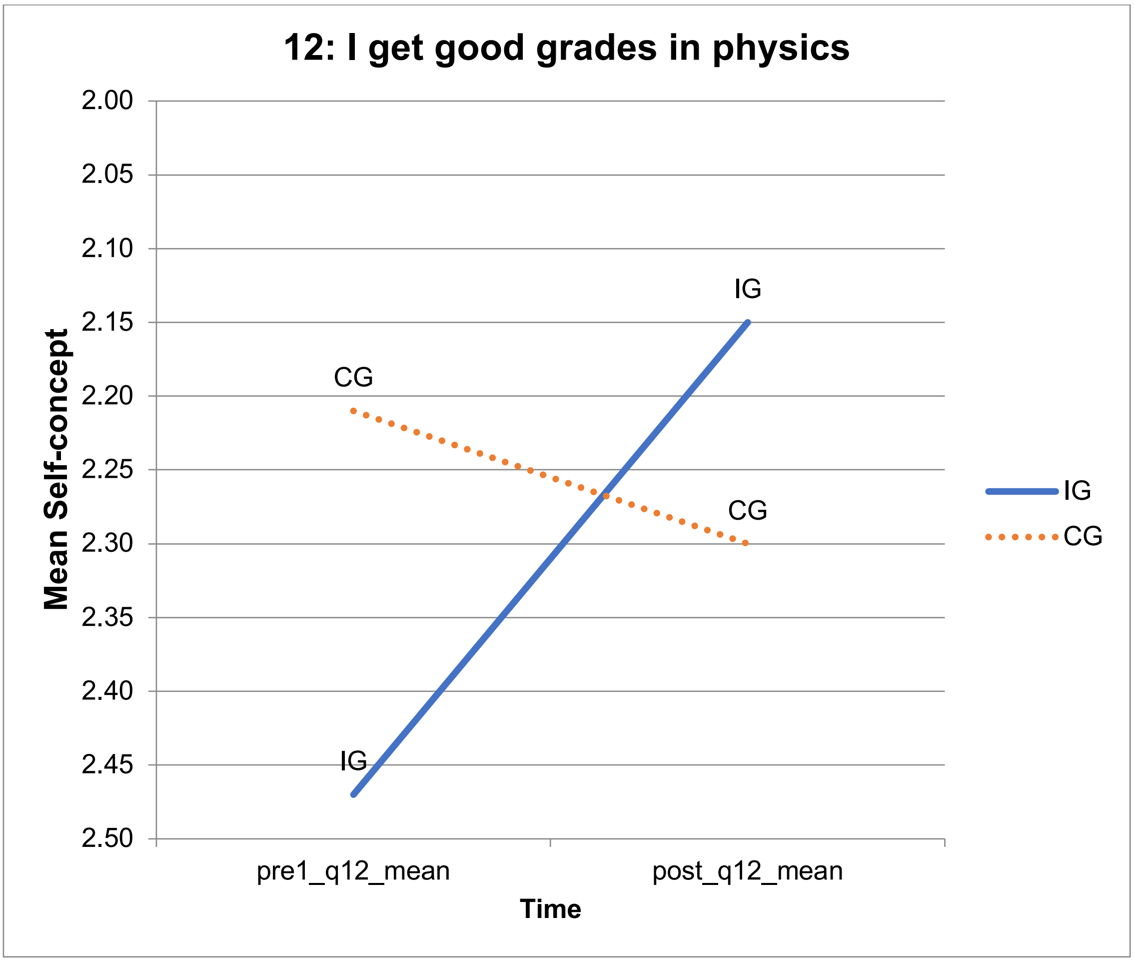


Graph 4.8 Pre- and post-survey results for item 6: In my physics class, I understand even the most difficult work.

4.3.3. I get good grades in physics

[item 12: I get good grades in physics]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*1*	3.0	*3*	9.1	*1*	3.0	*3*	9.1
	Agree	*19*	57.7	*21*	63.7	*27*	81.9	*17*	51.5
	Disagree	*10*	30.2	*8*	24.2	*4*	12.1	*13*	39.4
	Strongly disagree	*3*	9.1	*1*	3.0	*1*	3.0	*0*	0
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.5 Pre- and post-survey results for item 12: I get good grades in physics.



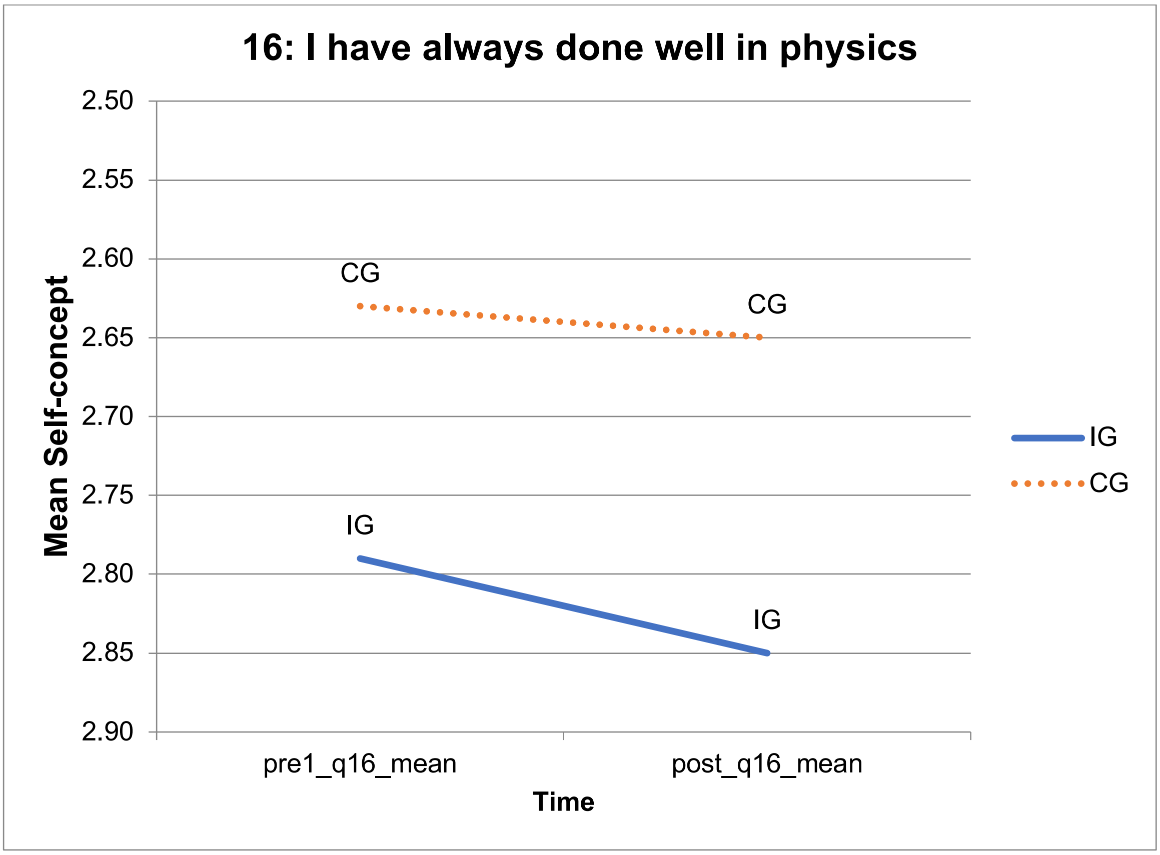
Graph 4.9 Pre- and post-survey results for item 12: I get good grades in physics.

Table 4.5 shows that for Item 12: ‘I get good grades in physics’, the IG self-concept rose from 60.6% to 84.8%, whereas the CG mean self-concept, initially higher, actually dropped from 72.6% general agreement (pre) to 60.6%, with a higher proportion disagreeing with this statement in the post-condition, as shown by Graph 4.11. This is a positive result for the IG compared to the CG.

4.3.4. I have always done well in physics

[item 16: I have always done well in physics]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*0*	0.0	*1*	3.0	*0*	0.0	*2*	6.1
	Agree	*9*	27.3	*12*	36.4	*8*	24.2	*13*	39.4
	Disagree	*21*	63.6	*17*	51.5	*22*	66.7	*14*	42.4
	Strongly disagree	*3*	9.1	*3*	9.1	*3*	9.1	*4*	12.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.6 Pre- and post-survey results for item 16: I have always done well in physics.



Graph 4.10 Pre- and post-survey results for item 16: I have always done well in physics.

For both pre and post scores, the CG indicated higher self-concept for this item. No student strongly agreed in the IG (post-test), whereas two students in the CG did, and a higher proportion overall. Indeed, the IG showed a decline in self-concept for this item, from 27.2% to 24.2%; CG were 39.4% to 45.4% agreement respectively (Table 4.6). This downward trend was mirrored to some extent by the CG as seen in Graph 4.10, though the IG show a steeper gradient. This reverses the trend of the other self-concept items, and it appears this was not a positive result for the IG. Given the fluidity and complexity of the motion of self-concept, and the dominant reference point being in the *past* (Seon and Bong, 2019), one interpretation could be that students believe they did not do well in the past, although may be doing better at this point in time; implications will be discussed in chapter 8.

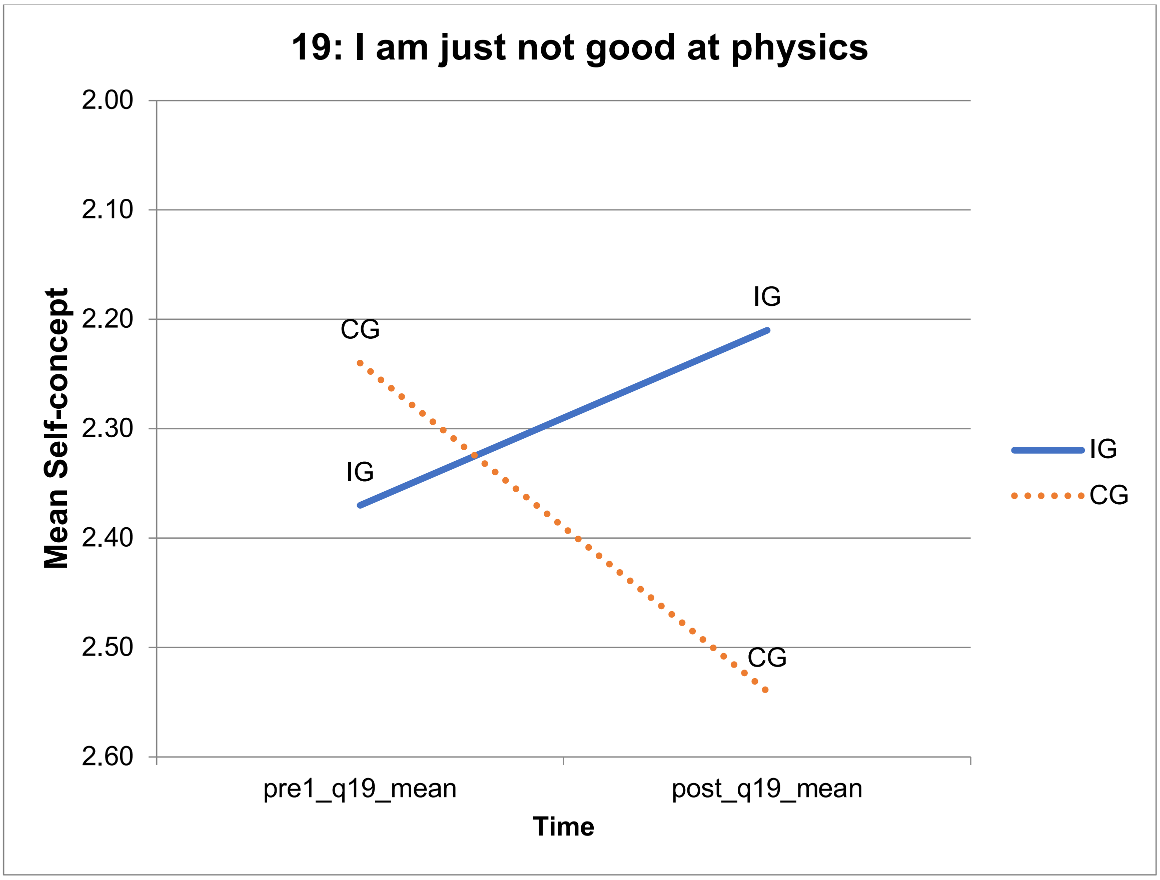
4.3.5. I am just not good at physics

[item 19: I am just not good at physics]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*3*	9.1	*4*	12.1	*3*	9.1	*5*	15.1
	Agree	*19*	57.7	*19*	57.7	*21*	63.7	*12*	36.4
	Disagree	*7*	21.1	*8*	24.1	*8*	24.2	*10*	30.3
	Strongly disagree	*4*	12.1	*2*	6.1	*1*	3.0	*6*	18.2
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.7 Pre- and post-survey results for item 19: I am just not good at physics (reverse coded).

Item 19 coded negatively for self-concept, serving as a check on other (positively-coding) self-concept items. Table 4.7 shows that 24/33 IG students disagreed post-test with this statement, representing 72.6% (previously 66.6%). For the CG, 51.6% disagreed with this statement, and this had declined from 69.8% at pre-test. In total, 17 out of 33 CG students disagreed with the statement, 5 of them strongly, which sets up an interesting tension between always doing well in physics [Item 16] despite not being good at physics [Item

19]. Graph 4.13 shows an increase in IG self-concept, and a decrease in CG self-concept.



Graph 4.11 Pre- and post-survey results for item 19: I am just not good at physics (reverse coded).

In summary, there were five self-concept items in total, with the IG cohorts showing positive trends in four of them. From a pre-test comparison, the IG showed a higher self-concept in items 1 and 6, initially lower in items 12 and 19, becoming higher than CG in 12 and 19, but lower in item 16 pre and post (I have always done well in physics). One interpretation of this could be that students in the IG group did not consider that they had a historical record of success in physics, but that more recently their perceptions of themselves in the subject had become more positive. Item 16 could be held to support this inference.

4.4 Findings from individual anxiety items

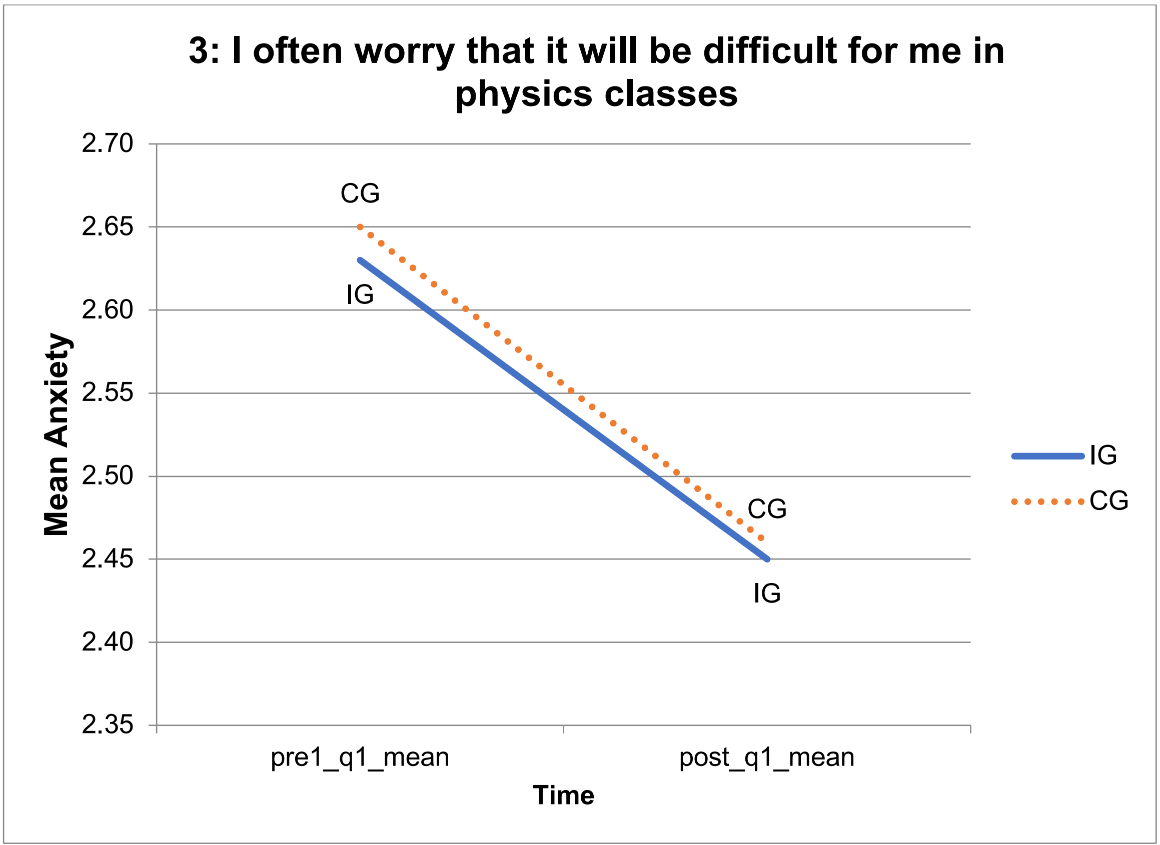
Anxiety has been described as an ‘emotional, distressing experience that is characterized by dislike, worry, and the wish to withdraw from the anxiety-provoking stimulus’ (Moeller *et al*, 2014), and ‘one’s physiological and affective responses when performing or thinking about a task (Morony *et al*, 2013). Responses to this type of question are potentially affected by the perceived stress of the situation at the time of answering the survey.

An agreement position on these items tends to lower numerical values; the y-axes have not been reversed so that the graph can show a negative trend should one exist.

4.4.1. I often worry that it will be difficult for me in physics classes

[item 3: I often worry that it will be difficult for me in physics classes]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*2*	6.1	*2*	6.1	*3*	9.1	*3*	9.1
	Agree	*11*	33.3	*10*	30.3	*15*	45.4	*16*	48.5
	Disagree	*17*	51.5	*18*	54.5	*12*	36.4	*9*	27.3
	Strongly disagree	*3*	9.1	*3*	9.1	*3*	9.1	*5*	15.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.8 Pre- and post-survey results for item 3: I often worry that it will be difficult for me in physics classes.



Graph 4.12 Pre- and post-survey results for item 3: I often worry that it will be difficult for me in physics classes.

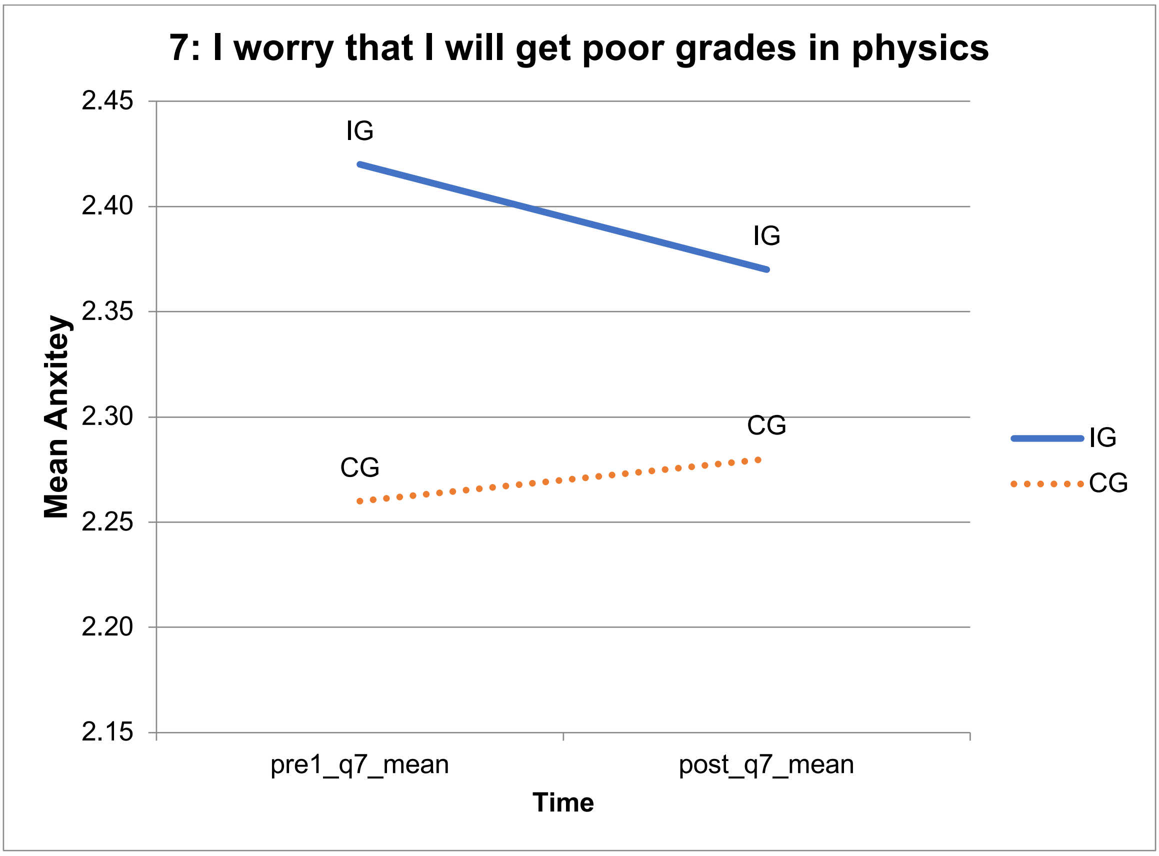
For this first survey item on anxiety, both IG and CG groups showed a like increase in anxiety (shown as a downward trend on Graph 4.12), and from very similar starting points. The IG rose from 39.4% to 54.5%; the CG rose from 36.4% to 57.6% (Table 4.8). The differences between the pre and post conditions indicate that although both increased anxiety, CG anxiety rose very slightly more than IG, though this is not significant given the sample size.

4.4.2. I worry that I will get poor grades in physics

This second anxiety item also showed an increase in anxiety in the IG group, however a slight decrease in the CG group. The IG increased from 17/33 to 19/33 in agreement and the CG from 19/33 to 21/33, as seen in Table 4.9. Nevertheless, the IG still exhibited less anxiety overall at both pre and post conditions (Graph 4.13), and the mean values indicate a position well within the ‘Disagree’ category.

[item 7: I worry that I will get poor grades in physics]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*4*	12.1	*8*	24.2	*4*	12.1	*4*	12.1
	Agree	*13*	39.4	*11*	33.3	*15*	45.4	*17*	51.5
	Disagree	*13*	39.4	*11*	33.3	*12*	36.4	*10*	30.3
	Strongly disagree	*3*	9.1	*3*	9.1	*2*	6.1	*2*	6.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.9 Pre- and post-survey results for item 7: I worry that I will get poor grades in physics.

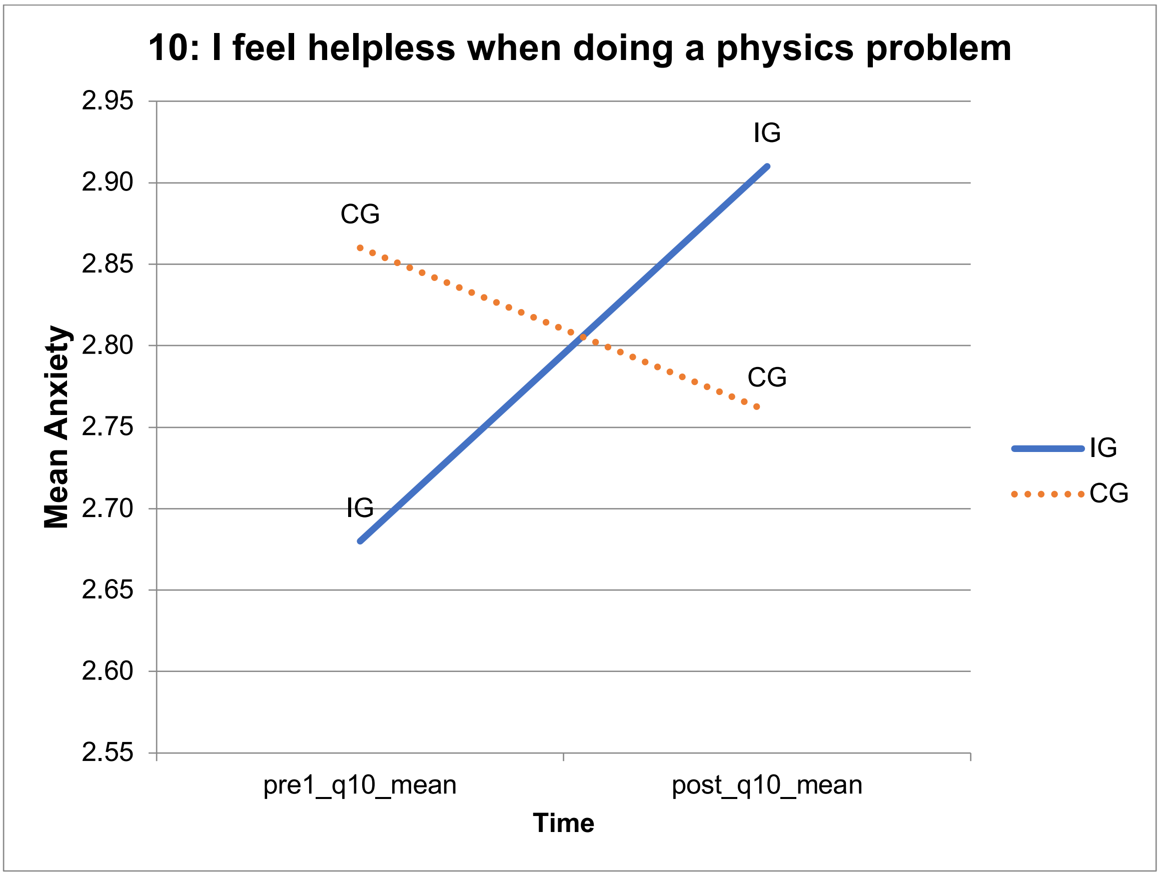


Graph 4.13 Pre- and post-survey results for item 7: I worry that I will get poor grades in physics.

4.4.3. I feel helpless when doing a physics problem

[item 10: I feel helpless when doing a physics problem]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*3*	9.1	*2*	6.1	*0*	0.0	*1*	3.0
	Agree	*7*	21.2	*9*	27.3	*6*	18.2	*8*	24.2
	Disagree	*22*	66.7	*15*	45.4	*24*	72.7	*21*	63.7
	Strongly disagree	*1*	3.0	*7*	21.2	*3*	9.1	*3*	9.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.10 Pre- and post-survey results for item 10: I feel helpless when doing a physics problem.



Graph 4.14 Pre- and post-survey results for item 10: I feel helpless when doing a physics problem.

‘Helpless’ is an emotive word, and there were gender differences on this item; females were more likely to agree across both cohorts than disagree.

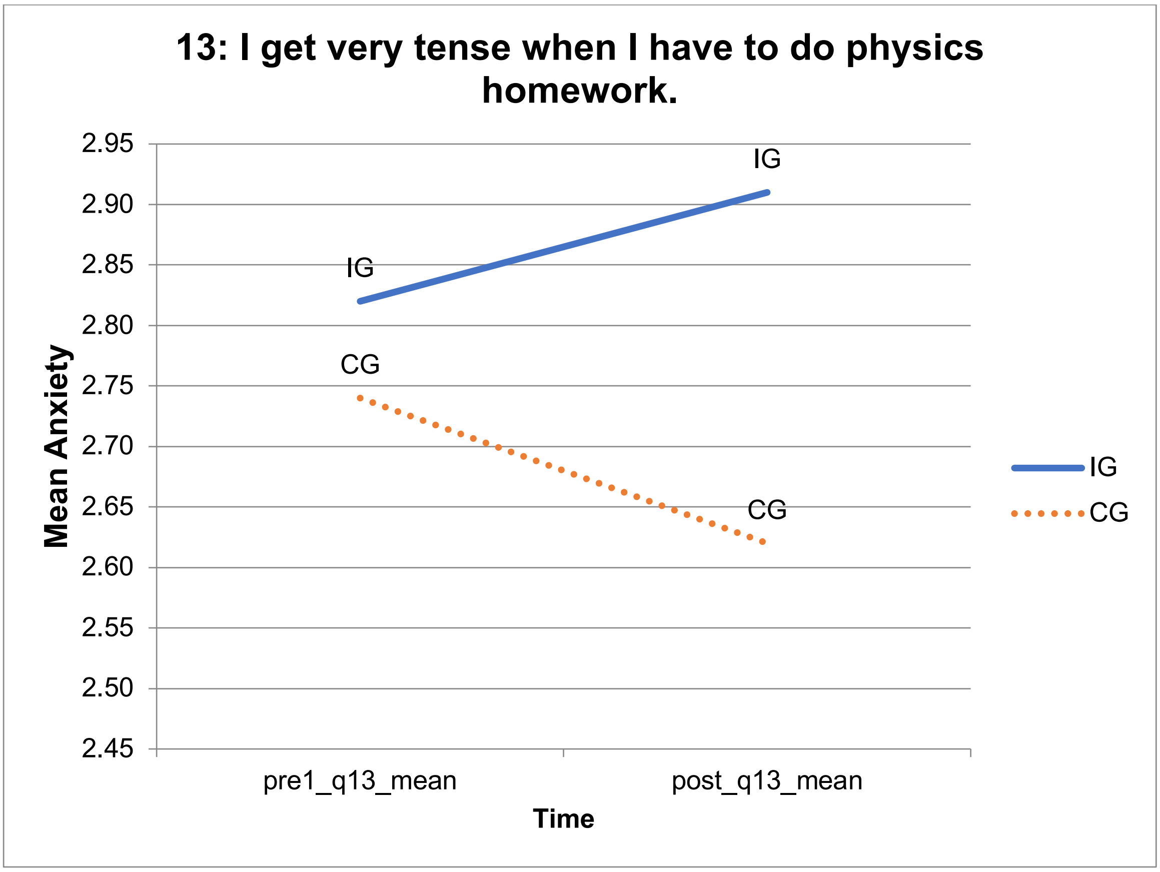
Only 18.2% of the IG agreed that they felt helpless (post-test) compared to 27.2% CG (post-test) as shown in Table 4.10. Pre-test, those figures had been 30.2% and 33.2% respectively. Graph 4.14 shows clearly that the IG decreased in anxiety, whereas the CG increased in anxiety. This is a positive result for the IG compared to the CG.

4.4.4. I get very tense when I have to do physics homework

[item 13: I get very tense when I have to do physics homework.]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*1*	3.0	*3*	9.1	*0*	0.0	*2*	6.1
	Agree	*6*	18.2	*9*	27.3	*5*	15.1	*10*	30.3
	Disagree	*24*	72.7	*18*	54.5	*26*	78.8	*18*	54.5
	Strongly disagree	*2*	6.1	*3*	9.1	*2*	6.1	*3*	9.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.11 Pre- and post-survey results for item 13: I get very tense when I have to do physics homework.

In Table 4.11, only 15.2% of the (post) IG group indicated that they felt tense when doing physics homework compared to 36.4% of the CG (post), from a starting point of 21.2% and 36.2% respectively. Graph 4.15 indicates that the anxiety level of the IG group, already lower than that of the CH group decreased further whilst the anxiety level of the CG group increased. This is a positive result for the IG compared to the CG.



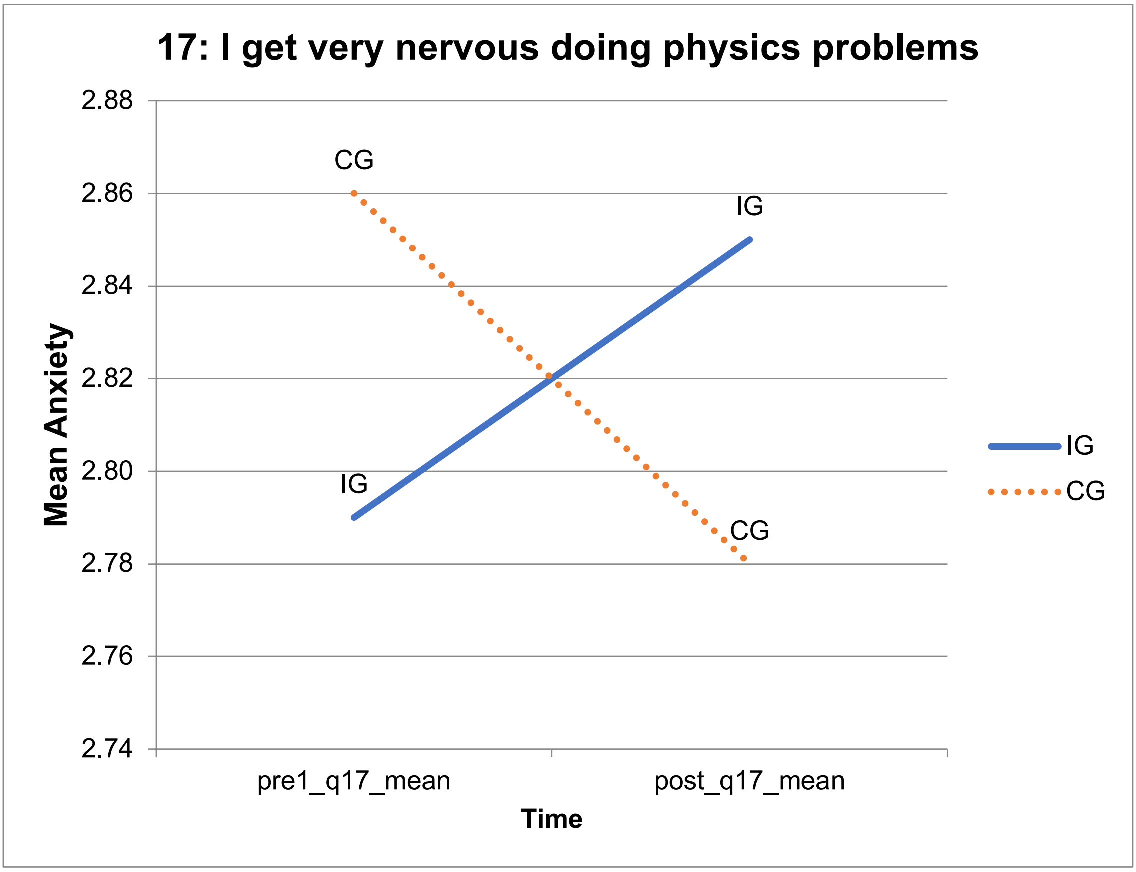
Graph 4.15 Pre- and post-survey results for item 13: I get very tense when I have to do physics homework.

4.4.5. I get very nervous doing physics problems

[item 17: I get very nervous doing physics problems]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*1*	3.0	*1*	3.0	*0*	0.0	*1*	3.0
	Agree	*8*	24.2	*9*	27.3	*7*	21.2	*10*	30.3
	Disagree	*21*	63.7	*16*	48.5	*24*	72.7	*15*	45.5
	Strongly disagree	*3*	9.1	*7*	21.2	*2*	6.1	*7*	21.2
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.12 Pre- and post-survey results for item 17: I get very nervous doing physics problems.

Although the IG had a slightly lower pre-test anxiety score for this item; pre-IG = 27.2%, pre-CG = 30.2%), the post-test scores indicated that the IG had decreased to 21.2% (and no student had indicated strongly agree), whereas the CG had increased to 33.4% (Table 4.12). Graph 4.16 indicates these trends for each group. The last three anxiety items have all straddled the 2.8 value on the y-axis for the IG group.



Graph 4.16 Pre- and post-survey results for item 17: I get very nervous doing physics problems.

There were 5 anxiety items in total, and the IG showed decreased anxiety for four of them from pre- to post-test conditions. Of these, one was comparable to the CG, however the other three indicated a decrease in anxiety against a background of an increase in the anxiety levels of the CG.

4.5 Findings from individual self-efficacy items

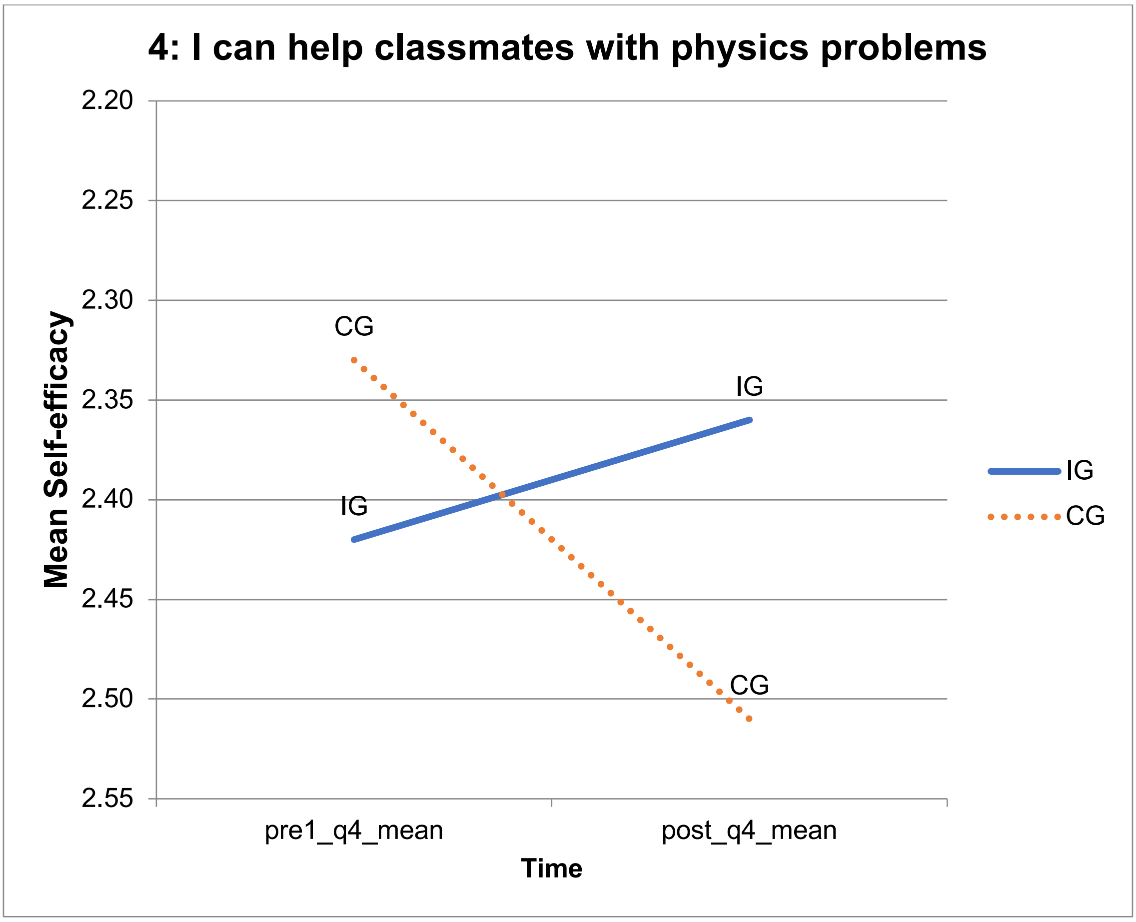
Self-efficacy items were designed to test the perception of the individual that they could competently complete a designated, specific task. The question items are presented positively, in 'can do' language.

4.5.1. I can help classmates with physics problems

[item 4: I can help classmates with physics problems]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*2*	6.1	*2*	6.1	*2*	6.1	*3*	9.1
	Agree	*17*	51.5	*20*	60.6	*20*	60.6	*16*	48.5
	Disagree	*12*	36.3	*8*	24.2	*8*	24.2	*10*	30.3
	Strongly disagree	*2*	6.1	*3*	9.1	*3*	9.1	*4*	12.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.13 Pre- and post-survey results for item 4: I can help classmates with physics problems.

Item 4 showed an increase in self-efficacy; of the IG group, 66.6% agreed overall compared to 57.4% of the comparison group [CG], and there was less strength of disagreement in the intervention classes (Table 4.13). This was an increase in self-efficacy from the pre-intervention survey in which the CG exhibited higher self-efficacy with 66.6% in the agreement categories, and the IG 57.6%. Graph 4.17 indicates these respective movements in self-efficacy, which is a positive result for the IG group.



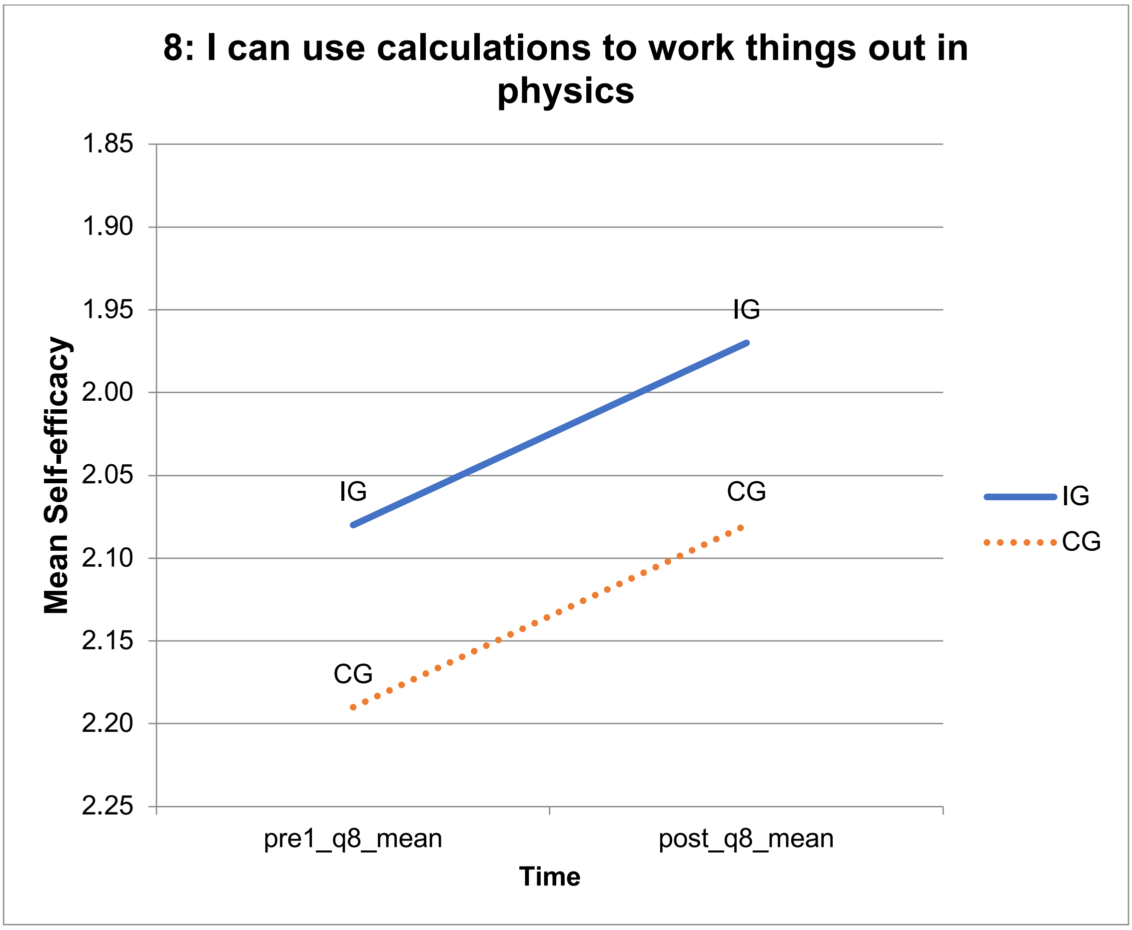
Graph 4.17 Pre- and post-survey results for item 4: I can help classmates with physics problems.

4.5.2. I can use calculations to work things out in physics

[item 8: I can use calculations to work things out in physics]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*4*	12.1	*6*	18.2	*5*	15.1	*7*	21.2
	Agree	*22*	66.7	*15*	45.4	*25*	75.8	*18*	54.5
	Disagree	*6*	18.2	*12*	36.4	*2*	6.1	*6*	18.2
	Strongly disagree	*1*	3.0	*0*	0.0	*1*	3.0	*2*	6.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.14 Pre- and post-survey results for item 8: I can use calculations to work things out in physics.

Item 8 presented as a higher degree of self-efficacy in the IG compared to the CG in the pre-intervention survey. However, the post-intervention survey also showed that the self-efficacy of the IG had actually increased further; CG rose from 63.6% agreement to 75.8% and IG rose from 78.8% agreement to 94.0%, with only three pupils disagreeing overall, as seen in Table 4.14. For CG, this was 11 pupils of the sample (=24.2%). This is graphically represented in Graph 4.18.

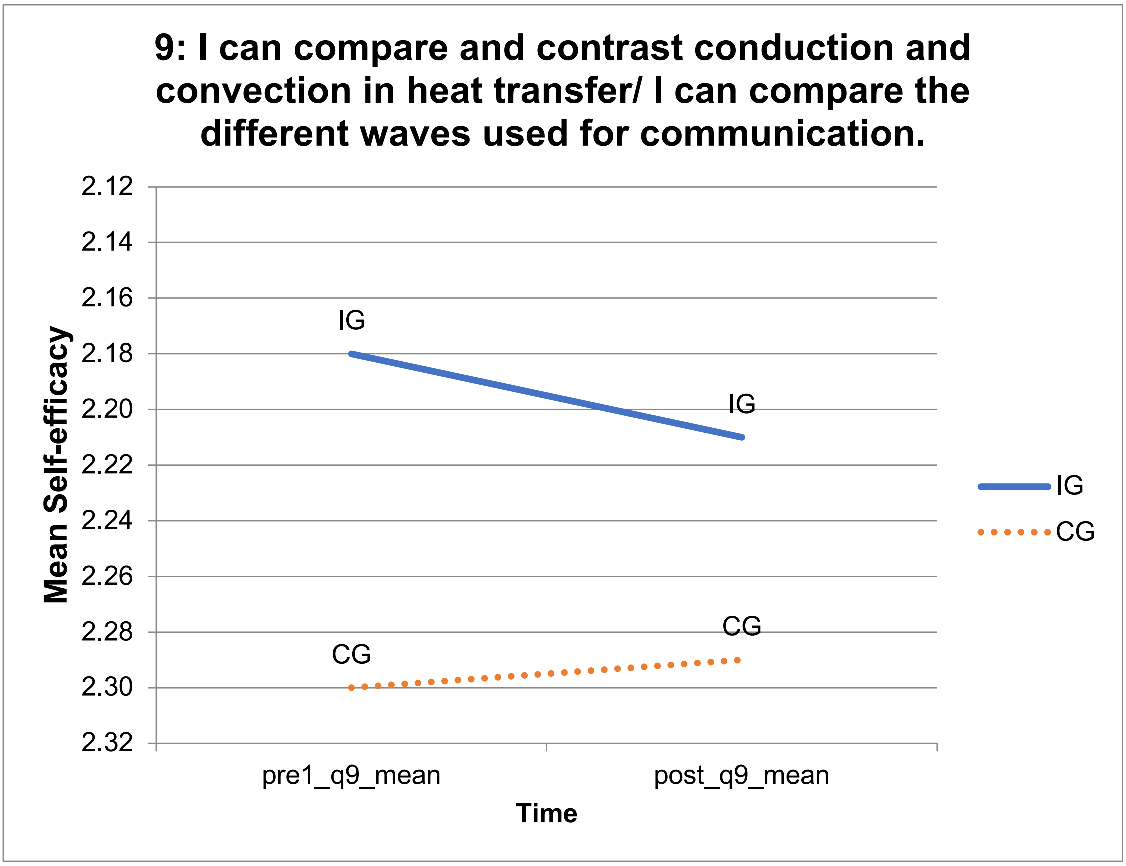


Graph 4.18 Pre- and post-survey results for item 8: I can use calculations to work things out in physics.

4.5.3. I can compare and contrast conduction and convection in heat transfer/ I can compare the different waves used for communication

[item 9: I can compare and contrast conduction and convection in heat transfer/ I can compare the different waves used for communication]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*2*	6.1	*4*	12.1	*4*	12.1	*4*	12.1
	Agree	*24*	72.7	*18*	54.5	*20*	60.6	*15*	45.5
	Disagree	*6*	18.2	*8*	24.2	*7*	21.2	*13*	39.4
	Strongly disagree	*1*	3.0	*3*	9.1	*2*	6.1	*1*	3.0
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.15 Pre- and post-survey results for item 9: I can compare and contrast conduction and convection in heat transfer/I can compare the different waves used for communication.



Graph 4.19 Pre- and post-survey results for item 9: I can compare and contrast conduction and convection in heat transfer/ I can compare the different waves used for communication.

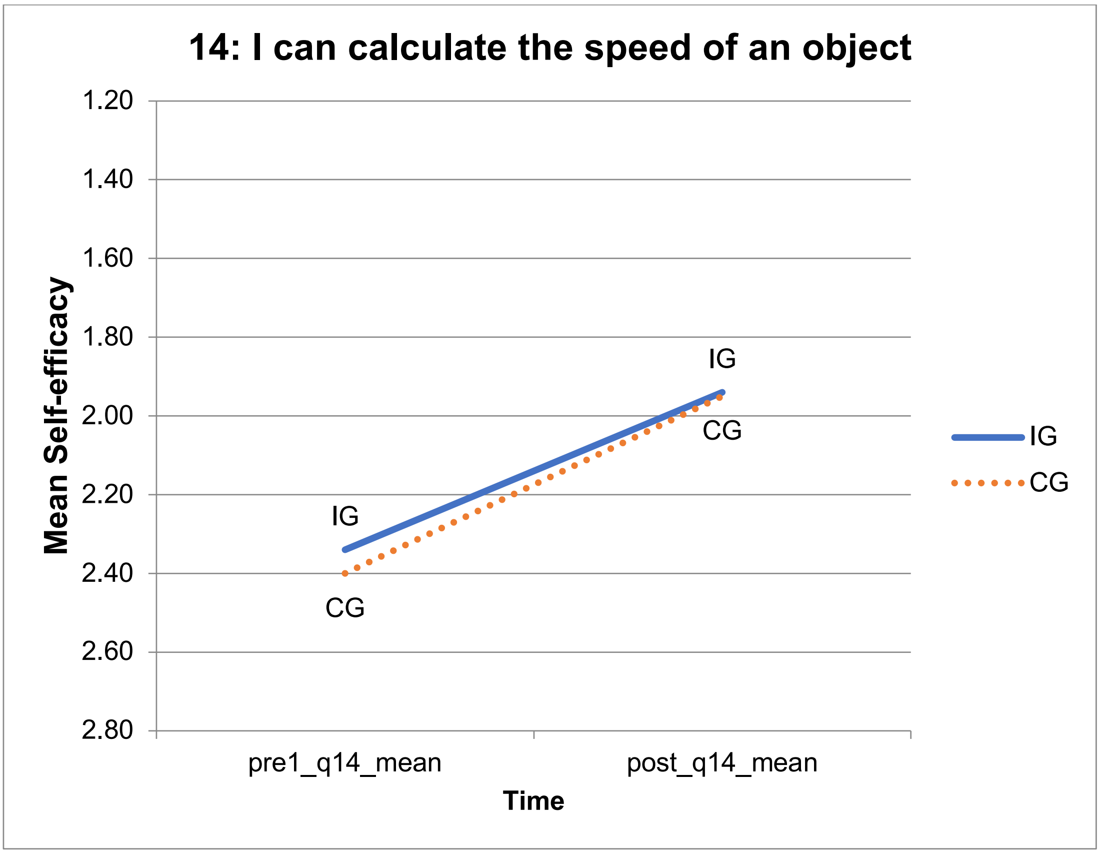
The comparison group showed broadly similar pre- and post-outcomes for self-efficacy on item 9; in both tests the IG had more students indicating agreement, although the graph indicates a very slight decrease in efficacy of the IG, as shown by Table 4.15 and Graph 4.19. Overall however, the self-efficacy levels of the IG remain higher than that of the CG. Possible inferential grounds are discussed below.

4.5.4. I can calculate the speed of an object

[item 14: I can calculate the speed of an object]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*2*	6.1	*5*	15.2	*5*	15.2	*9*	27.3
	Agree	*19*	57.6	*10*	30.3	*25*	75.7	*17*	51.5
	Disagree	*10*	30.3	*17*	51.5	*3*	9.1	*7*	21.2
	Strongly disagree	*2*	6.1	*1*	3.0	*0*	0.0	*0*	0.0
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.16 Pre- and post-survey results for item 14: I can calculate the speed of an object.

Both groups indicated an increase in self-efficacy for this item with broadly comparable scores on both pre and post-tests (Table 4.16). The difference in means for this item are shown in Graph 4.20. It is interesting that this is one of two items in which no students strongly disagreed in the post-test situation; inferentially this could be linked to the time of year (June) with examinations approaching, and recent practice having been encountered through revision.



Graph 4.20 Pre- and post-survey results for item 14: I can calculate the speed of an object.

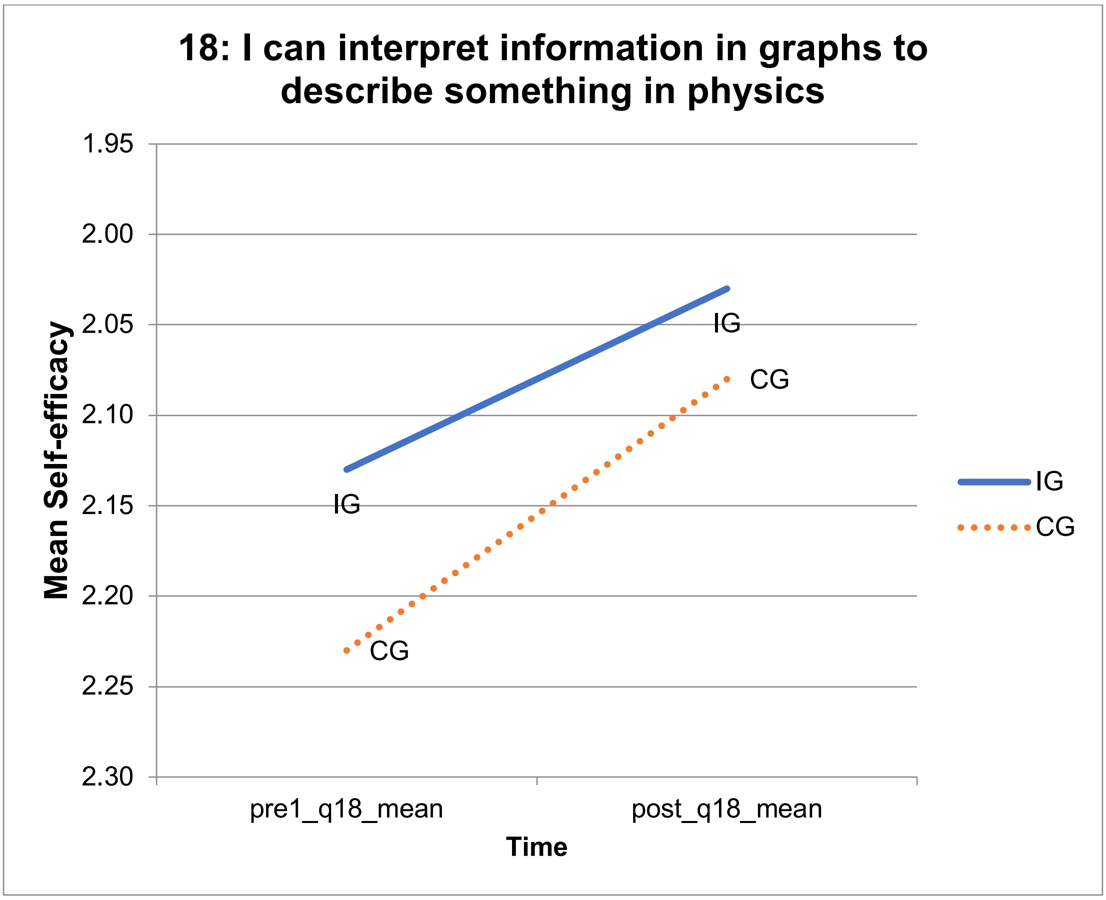
4.5.5. I can interpret information in graphs to describe something in physics

[item 18: I can interpret information in graphs to describe something in physics]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*2*	6.1	*3*	9.1	*2*	6.1	*5*	15.2
	Agree	*24*	72.7	*19*	57.6	*28*	84.8	*20*	60.6
	Disagree	*7*	21.2	*9*	27.2	*3*	9.1	*8*	24.2
	Strongly disagree	*0*	0.0	*2*	6.1	*0*	0.0	*0*	0.0
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.17 Pre- and post-survey results for item 18: I can interpret information in graphs to describe something in physics.

90.8% of the IG indicated agreement with this statement, representing a rise from 78.8%; The post-test for CG was 75.8% in total (Table 4.17). Only 3 IG students indicated that they disagreed with this, compared to 8 students in the CG. No students strongly disagreed in the post-test situation. Graph 4.21 shows the increase in self-efficacy of both groups, with the OG remaining at a higher self-efficacy level overall.

The use of ‘describing’ from a graph was chosen in preference to ‘explaining’ from a graph, since the latter is a higher cognitive demand, and students would be less likely to agree with this as a ‘can do’ statement.

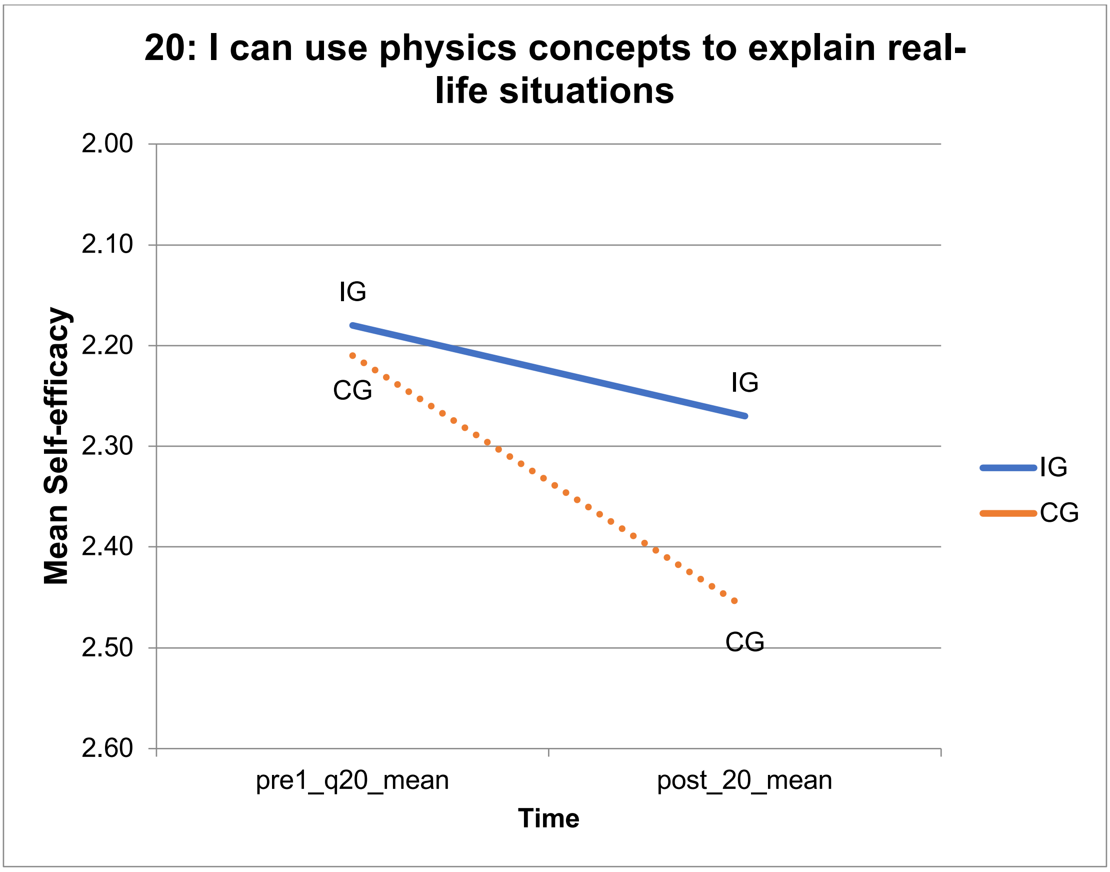


Graph 4.21 Pre- and post-survey results for item 18: I can interpret information in graphs to describe something in physics.

4.5.6. I can use physics concepts to explain real-life situations

[item 20: I can use physics concepts to explain real-life situations]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*5*	15.2	*6*	18.2	*2*	6.1	*4*	12.1
	Agree	*18*	54.5	*16*	48.5	*22*	66.6	*15*	45.5
	Disagree	*7*	21.2	*9*	27.2	*7*	21.2	*11*	33.3
	Strongly disagree	*3*	9.1	2	6.1	2	6.1	*3*	9.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.18 Pre- and post-survey results for item 20: I can use physics concepts to explain real-life situations.



Graph 4.22 Pre- and post-survey results for item 20: I can use physics concepts to explain real-life situations.

Item 20 was another positive result for the IG cohort but made so by the fact that the self-efficacy of the CG actually dropped from 66.6% (pre) to 57.6% (post); 42.4% of the CG cohort felt that they were unable to use physics to explain everyday situations (Table 4.18). The IG maintained a fairly stable, higher self-efficacy level at 72.6% (post) from 69.8% (pre). This is shown visually in Graph 4.22.

In summary, the IG indicated higher levels of self-efficacy in all six efficacy items. In five of the six, the IG mean started and ended on a score above that of the CG mean; of the sixth [item 4], the IG started at a lower self-efficacy mean score, but increased their self-efficacy, whereas the CG actually decreased theirs.

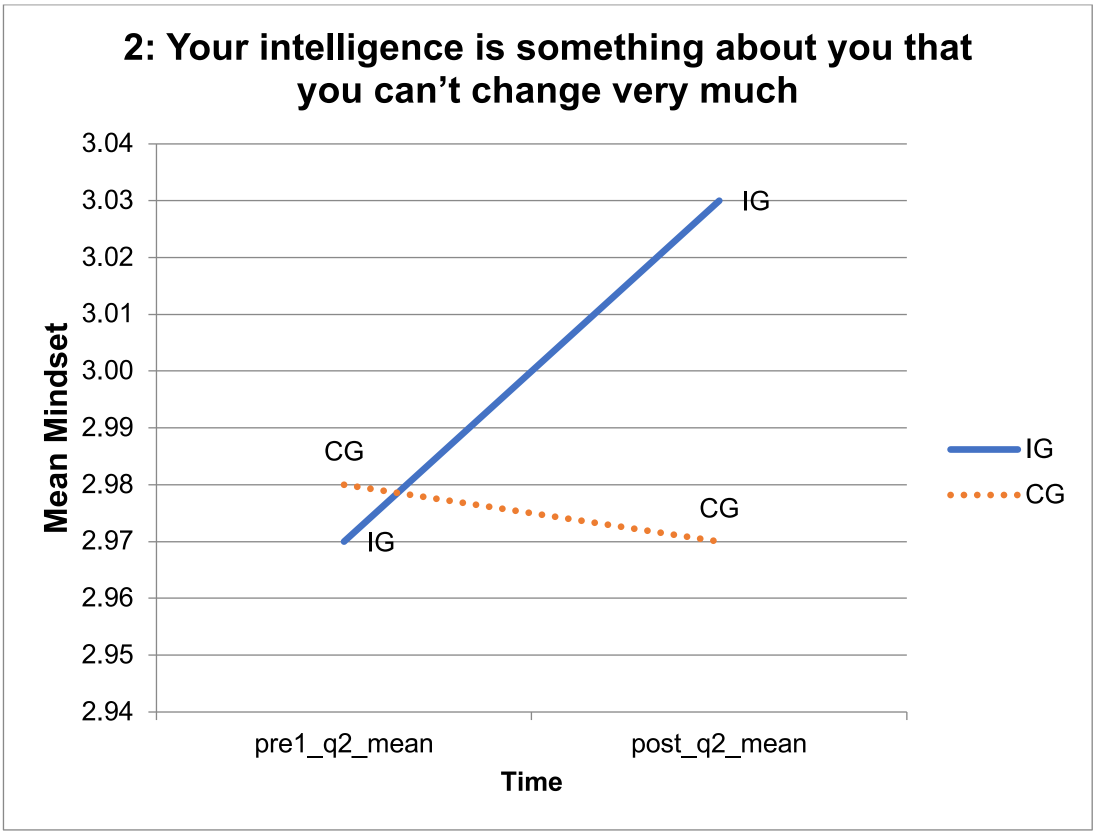
4.6 Findings from individual mindset items

The mindset items were drawn from validated surveys developed by Dweck (2000, 2006). One was reverse coded to act as a check on the other responses [item 11]. For these items, the lower the scale [i.e., agreement], the more an entity ('fixed') theory of intelligence was being espoused. Graphical trends upwards therefore show a movement towards an incremental theory of intelligence, or more colloquially 'growth mindset'.

4.6.1. Your intelligence is something about you that you can't change very much

[item 2: Your intelligence is something about you that you can't change very much]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*0*	0.0	*0*	0.0	*1*	3.0	*0*	0.0
	Agree	*6*	18.2	*8*	24.2	*4*	12.1	*7*	21.2
	Disagree	*22*	66.7	*19*	57.6	*21*	63.7	*20*	60.6
	Strongly disagree	*5*	15.1	*6*	18.2	*7*	21.2	*6*	18.2
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.19 Pre- and post-survey results for item 2: Your intelligence is something about you that you can't change very much.



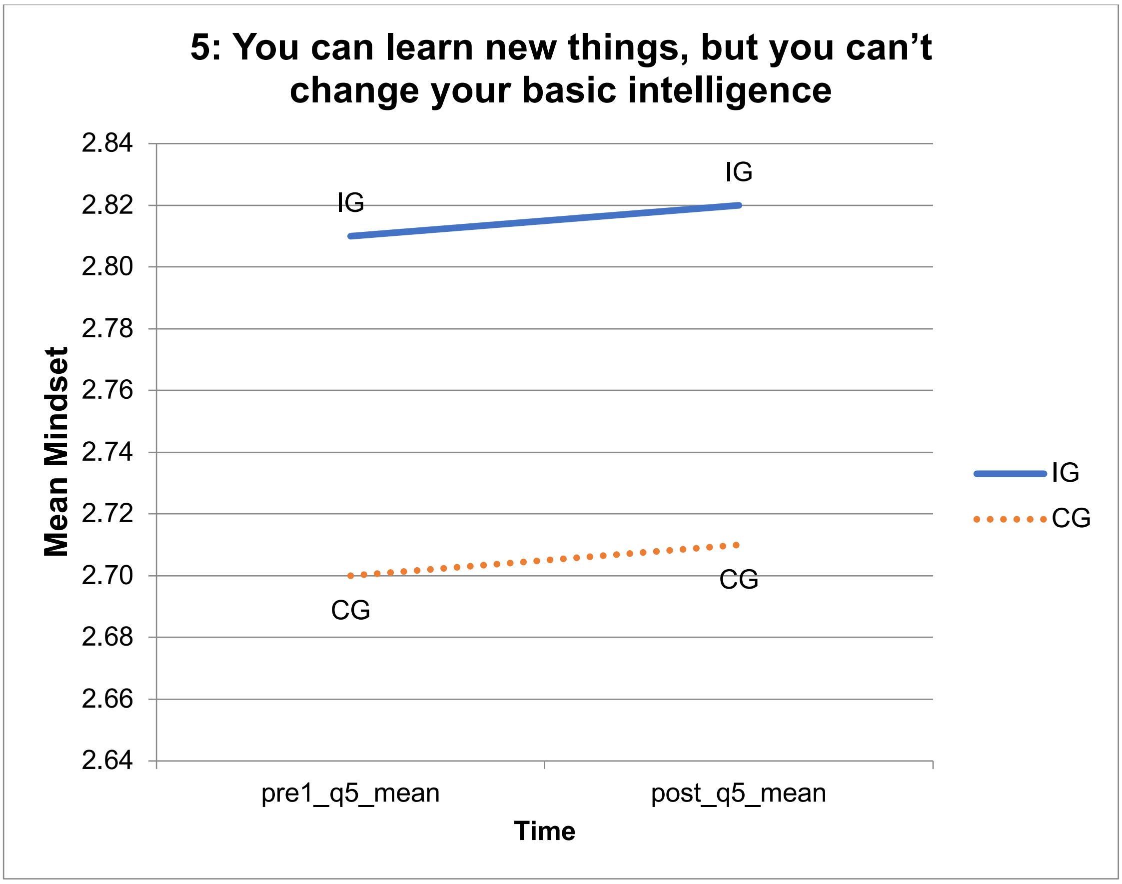
Graph 4.23 Pre- and post-survey results for item 2: Your intelligence is something about you that you can't change very much.

For this item, the CG appeared to have a relatively stable mindset score across the life of the intervention; the IG however showed a slight trend towards 'growth mindset', and this is reflected in Graph 4.23. 84.8% of the IG disagreed with this statement at the post-test stage, and less had agreed overall (Table 4.19). This is a (very slightly) positive result for the IG.

4.6.2. You can learn new things, but you can't change your basic intelligence

[item 5: You can learn new things, but you can't change your basic intelligence]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*0*	0.0	*1*	3.0	*2*	6.1	*1*	3.0
	Agree	*11*	33.3	*11*	33.3	*7*	21.2	*12*	36.4
	Disagree	*17*	51.6	*20*	60.7	*19*	57.6	*16*	48.5
	Strongly disagree	*5*	15.1	*1*	3.0	*5*	15.1	*4*	12.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.20 Pre- and post-survey results for item 5: You can learn new things, but you can't change your basic intelligence.



Graph 4.24 Pre- and post-survey results for item 5: You can learn new things, but you can't change your basic intelligence.

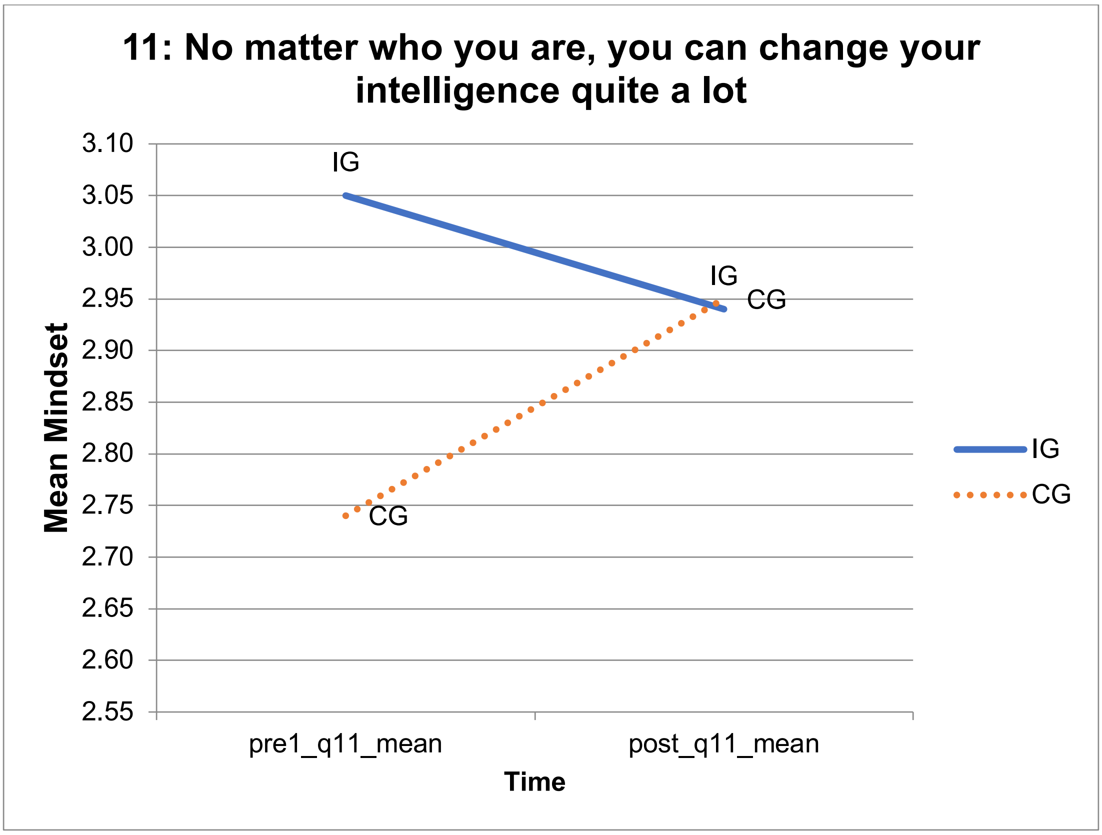
Item 5 by contrast showed that both the CG and IG had a relatively stable score for this question, and that overall, the IG had tended to a higher 'growth mindset' score (Graph 4.24). In the IG, 24 out of 33 students disagreed that intelligence couldn't be changed this proportion for the CG was 20/33 (Table 4.20).

4.6.3. No matter who you are, you can change your intelligence quite a lot

[item 11: No matter who you are, you can change your intelligence quite a lot]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*0*	0.0	*3*	9.1	*1*	3.0	*0*	0.0
	Agree	*8*	24.2	*6*	18.2	*9*	27.3	*6*	18.2
	Disagree	*17*	51.6	*22*	66.6	*14*	42.4	*23*	69.7
	Strongly disagree	*8*	24.2	*2*	6.1	*9*	27.3	*4*	12.1
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**

Table 4.21 Pre- and post-survey results for item 11: No matter who you are, you can change your intelligence quite a lot

Item 11 is reverse coded as a check for the positively coded items. Therefore, in terms of the graph, a higher number will still indicate a trend towards a 'growth mindset' position (Graph 4.25). It is possible that students found the oppositely worded statement confusing, since for the IG, nine students indicated strong disagreement, and 14 disagreement with this, not replicated in any other mindset item, seen in Table 4.21.



Graph 4.25 Pre- and post-survey results for item 11: No matter who you are, you can change your intelligence quite a lot.

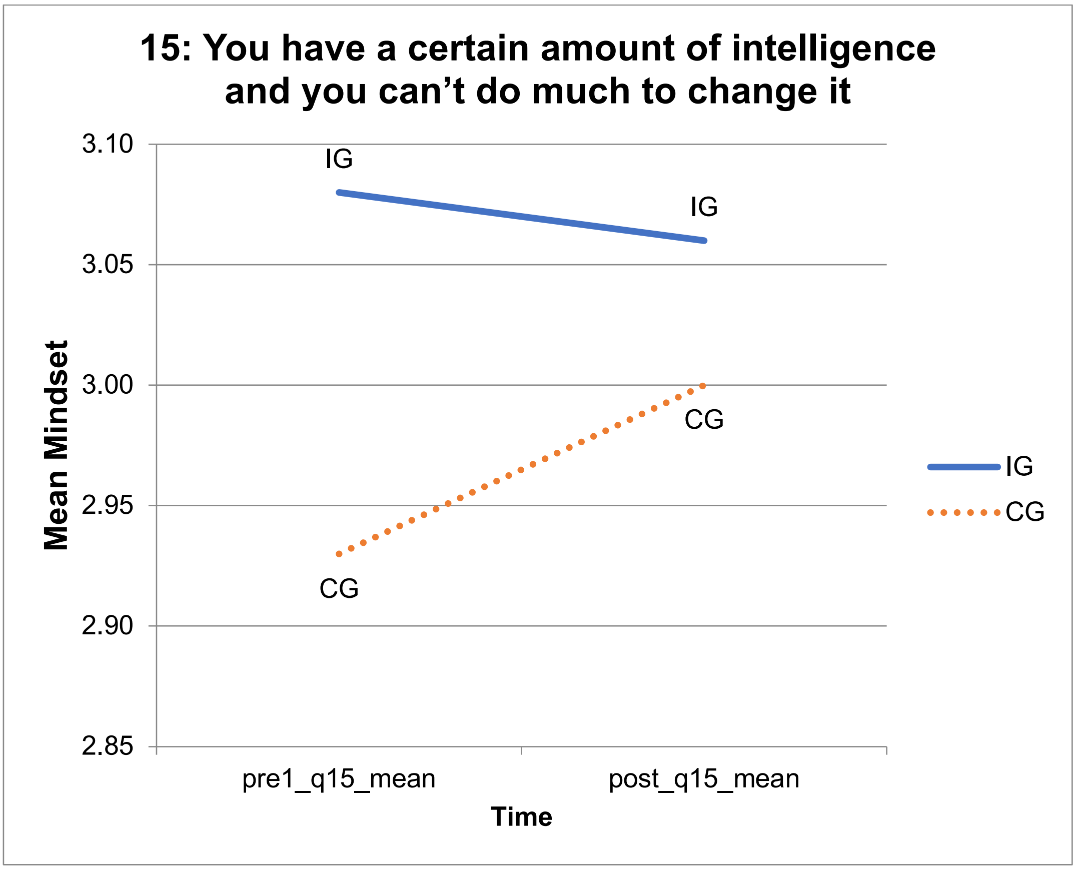
4.6.4. You have a certain amount of intelligence and you can't do much to change it

[item 15: You have a certain amount of intelligence and you can't do much to change it]									
		Pre-intervention				Post-intervention			
		Count IG	% IG	*Count CG*	% CG	*Count IG*	% IG	*Count CG*	% CG
Ordinal	Strongly agree	*0*	0.0	*0*	0.0	*1*	3.0	*0*	0.0
	Agree	*4*	12.1	*7*	21.2	*1*	3.0	*5*	15.2
	Disagree	*22*	66.7	*23*	69.7	*26*	78.8	*23*	69.7
	Strongly disagree	*7*	21.2	*3*	9.1	*5*	15.2	*5*	15.2
Total		***33***	**100.0**	***33***	**100.0**	***33***	**100.0**	***33***	**100.0**[1]

Table 4.22 Pre- and post-survey results for item 15: You have a certain amount of intelligence and you can't do much to change it.

[1] rounded

Although the IG had a very slight downward trend as shown in Graph 4.26 on this last mindset item, they retained a higher overall score in the incremental theory range, and also that the CG began to show an upward trend themselves. Table 4.22 indicates that only one student strongly agreed with this statement at t he start, and none at the end of the study.



Graph 4.26 Pre- and post-survey results for item 15: You have a certain amount of intelligence and you can't do much to change it.

In summary, mindset items showed movement in both directions for both groups, with a very slight advantage in three of the four items for the IG, although this is not a convincing claim in terms of the sample size. The position of mindsets in the self-belief system will be discussed in chapter 8 alongside the other constructs, and the implication of these findings for classroom practice.

4.7 Chapter summary

This chapter has provided an analysis of the pre- and post-intervention student surveys to present findings answering research questions one and two:

1. *What are the students' self-belief systems as they enter Key Stage 4?*
2. *Do the students' self-belief systems in physics change over the duration of the intervention?*

The findings show that self-belief constructs of self-concept, self-efficacy, anxiety and mindset were identified and measures of these obtained for all students taking the surveys, thus answering RQ1. Only those who completed both the pre- and post-intervention surveys were included in the SPSS data analysis. From this sample of 66 students (33 in each group), it could be shown that the IG students had slightly higher starting points in these constructs than the CG students, though the difference was very small.

Analysis of the post-intervention surveys answered RQ2 in showing that the students' self-beliefs were not static. However, the intervention group students showed positive gains in self-concept and self-efficacy and a decrease in anxiety, whilst mindset appeared mostly stable. In the comparison group, self-concept decreased, negatively correlating with an increase in anxiety (Lee, 2009; Morony *et al.*, 2013; Ahmed, 2019), whereas the self-efficacy measure showed a slight increase. The implications of these findings are discussed in Chapter 8 with reference to the literature, and now we turn to the analysis for the classroom observation data: findings from teacher verbal feedback.

Chapter 5. Findings from the Observations of Teacher Verbal Feedback

5.1 Introduction

The last chapter focused on the survey instruments used in the pre- and post-intervention states to ascertain what the students' self-belief systems in physics were in Year 10, and to determine whether these had changed over time. Both of these matters were established, and evidence presented to answer the first two research questions comprising Study 1. Chapter 5 moves to centre on the second area of study, considering the oral interactions used by the participant teachers in order to answer research questions three and four:

3. *What types and proportions of feedback do teachers use in their verbal interactions with students in physics?*
4. *Does the feedback pattern change during the life of the study in response to a CPD programme of flexing verbal interaction styles?*

The sources for analysis in this chapter are the coded observations, field notes and partial transcripts of a total of 45 audio-recorded lessons acquired from the intervention and comparison teachers involved in the study, over a period of six months. The lesson recordings were coded according to the constructed typology specified and discussed in Chapter 3 and analysed to establish the feedback repertoire of the teachers.

The chapter presents findings to show that analysis of teachers' verbal interactions to their students indicates that:

- using the typology constructed in Chapter 3 enables categorisation of VF at much higher frequencies that that reported in other studies (Hattie, 1999, Bond *et al*., 2000; Voerman *et al*., 2012; Campbell-Mapplebeck, 2019;
- verbal feedback [VF] comprises a large part of classroom dialogue (Hargreaves, 2013, 2014; Svanes and Skagen, 2017);

- it is possible to identify teachers' feedback style, and also indicates differences in teacher 'talk' ranging from structured, teacher-led, and less dialogic practices to responsive, individualised and authentically dialogic patterns;
- over the length of the study, the intervention teachers were observed to increase their use of regulatory, process and prompt forms of VF, and decrease their use of task type. The comparison teachers did not change;
- the intervention teachers had lower statement-to-prompt ratios which meant they 'closed the loop' more often than the comparison teachers. Consequently, intervention group students received more information on their next steps in learning than comparison group students (*instrumental* help rather than *executive* help (Hattie and Timperley, 2007)).
- consistent with both CPD and feedback intervention literature outlined in Chapter 2, the intervention teachers experienced challenge in *flexing* their VF in the learning instance, initially falling back onto their routinised repertoires.

In a subsequent chapter, these oral interaction patterns will be coordinated with the post-intervention condition survey responses, linking teacher feedback behaviours with student self-belief outcomes.

Table 5.1 presents an overview of the types and proportions of feedback used by the participant teachers in their feedback-related oral interactions as percentages over all their observed lessons; these form a *part* of the total oral interactions employed by the teachers, and subsequently coded by the researcher. These total interactions are presented in section 5.2, since they provide a useful background to establishing profiles of teacher talk that can be varied between teachers, and together with a consideration of examples of teacher dialogue, present evidence to answer RQ3, determining teacher feedback profiles.

In all tables in this chapter, the intervention teachers are represented in the shaded boxes.

	Feedback type	Teacher 1	Teacher 2	Teacher 3	Teacher 4
First level	Task	52.2	55.7	84.2*	63.4
	Process	39.3	39.9*	13.2	35.5
	Regulation	6.2*	3.7	2.6	0.2
	Self/Praise	2.3*	0.7	0.0	0.9
	% Totals	**100.0**	**100.0**	**100.0**	**100.0**
Second level	statement	55.7	56.7	60.1*	51.5
	prompt	44.3	43.3	39.9	48.5*
	% Totals	**100.0**	**100.0**	**100.0**	**100.0**

Table 5.1. An overview of the types and proportions of *feedback* used by the participant teachers in their learning-related oral interactions as percentages. *(* denotes the highest value for that type)*

The findings show that:

- At the first level, the intervention teachers [T1 and T2] had similar amounts of both Task and Process feedback as each other, and that these formed the majority of their feedback in the range 91.5 to 95.6%;
- The comparison teachers [T3 and T4] used higher amounts of Task feedback, and in the case of T3, this was appreciably higher at 84.2%, and Process type was very low at 13.2% compared to the other teachers. The Task plus Process totals for T3 and T4 were 97.7% and 98.9% respectively.
- T4 had a slightly higher rate of Task feedback than the intervention teachers, but the use of Process feedback was broadly similar;
- Both of the IG teachers had higher proportions of self-regulation feedback; this type formed a very small part of T4’s profile;

- T1 used more ‘undifferentiated’ (self/praise) feedback than the other teachers;
- At the second level, the intervention teachers had lower statement-to-prompt ratios than T3, but T4 had the lowest statement-to-prompt ratio of all. However, as discussed below, an analysis of the second level related to the first level indicated that for T4, the students were receiving nearly twice as many statement and prompts related to the Task level rather than the Process; that is, relating to the product rather than the thinking around the learning.

These findings can be compared with existing literature on types of feedback and their frequencies, such as shown in Table 2.2, and also compared over time; the fourth research question related to how the types and proportions of feedback may have *changed* over the duration of the intervention, and this is presented in light of the total coded data. In sum, the intervention teachers were seen to change their feedback types and therefore proportions in response to the CPD intervention and coaching conversations. At the first level, the amount of Task feedback decreased, whereas Process feedback increased; the amount of Regulation feedback fluctuated but showed ‘peaks’ following coaching conversations. At the second level, the teachers were able to increase the prompt (discrepancy) feedback used.

An example of the frequencies and percentages of all data are presented in a spreadsheet in Appendix 15. The analysed classroom observation data will be presented in an increasingly granular manner, and additional commentary by the intervention teachers in Chapter 7 will add rich detail to the process in which they challenged their existing feedback practices.

5.2 Summary of all oral interaction types across all teachers

Table 5.2 gives an overview of the total of observed lessons over the duration of the intervention.

Teacher	Number of lessons	Total interaction observation time in hours and minutes
1	14	13h 6m
2	22	18h 28m
3	4*	2h 26m
4	5	4h 34m
Total	**45**	**38h 34m**

Table 5.2 Summary of lessons and total interaction observation time by teacher

* T3 did not record full lessons

There were additional lessons which were observed by the researcher, for which no recording was obtained; these have not been included in the analysis since verification of field coding could not be corroborated. These were however helpful in providing events and dialogue that could be discussed and evaluated in a coaching conversation. The teachers of the comparison groups (T3 and T4) were observed only at the start and end of the study, to provide a baseline and end-point comparison to the two teachers of the intervention groups. The intervention group teachers (T1 and T2) were observed throughout the study, although not to the same extent due to both teacher and researcher availability.

As described in the analysis plan in Chapter 3, each type of code, and some combinations of code were tallied and summed for each lesson, and each teacher. From this operation, it is possible to show the percentage frequencies of all *types* of oral interaction (N_T), as summarised for all teachers in Table 5.3 below. The coding from the constructed typology shown in Table 3.1 employed three semantic sampling levels; the first related to the feedback *category* or *type* (defined as *level* by Hattie and Timperley, 2007; Gan, 2011); secondly, the feedback *mode* (whether a statement or a prompt for future action, equating to Voerman *et al.*'s (2012, 2015) progress and discrepancy feedback); and thirdly, the *target* or recipient(s) of the feedback.

Interaction type		Frequency of occurrence for **all** teachers as % of *total* interactions
First level – Interaction *category*		
T	Task/product*	39.2
P	Process/strategy	19.7
R	Self-regulation	2.1
S	Self/praise	0.6
I	Instruction	24.8
O	Other (not related to learning goals)	13.6
	Total	**100.0**
Second level – feedback/interaction *mode*		
s	Statement	55.7
p	Prompt	30.7
	Total	**86.4****
Third level – target/recipient		
B	Male recipient	31.1
G	Female recipient	22.6
C	Class as recipient	30.8
T	Team, or small group as recipient	1.9
	Total	**86.4****

Table 5.3 A summary of *total* oral interaction types (N_T for all teachers)

(*shaded boxes are oral *feedback* types; **figure cannot be 100% since O is not coded with second/third levels)

The results have been initially presented in this way to enable ‘headline’ comparison with the findings from similar studies (Brooks *et al*., 2019; Hattie and Timperley, 2011; Gan, 2011; and van der Bergh *et al*., 2012), although subsequent discussion will focus on the proportions of the intervention teachers rather than the comparison teachers. The table shows that:

- the amounts of Task and Process feedback are similar to those reported in Brooks *et al*., 2019; Hattie and Timperley, 2011; Gan, 2011; and van der Bergh *et al*, 2012;
- the self-Regulation level is slightly higher than in most of these studies; subsequent analysis indicated that this largely resulted from the intervention group teachers;
- the Self or praise level is much lower than in the above studies;
- there are almost twice as many feedback *statements* as there are feedback *prompts;* subsequent analysis indicated that these amounts differed between the IG and CG teachers;
- at the third level, teachers appear to have more interactions with male students than female students (approximately 40% more), and slightly more with male students than the whole class. However, when considering the ratio of males to females (1.5:1) in the composition of each of the four groups, this is proportionate.
- interactions with the whole class appeared to account for nearly one-third of all interactions, however analysis by teacher showed a large variation between the IG and CG teachers, and between the two comparison group teachers (Table 5.5);
- the teachers in this study did not often interact with groups of students; it was predominantly either with individuals or whole class.

As described in the methodology, for the purposes of the study, it was important to be able to distinguish between the different types of oral interactions, since some of these do not relate to *learning* within the lesson. Table 3.3 outlined how sub-totals of interaction type were obtained, and for ease of reference with respect to the following tables and graphs, these are summed here also:

N_T = *Total* oral feedback interactions: these included all the first level codes (T, P, R, S, I and O)

N_L = Oral interactions relating to *learning*: since Instruction was considered to relate to learning, unless it was administrational dialogue, all tallies coding for Other (O) were removed)

N_F = Oral *feedback* interactions only: finally, the frequencies coding for Instruction (I) were removed, leaving verbal feedback expressions only.

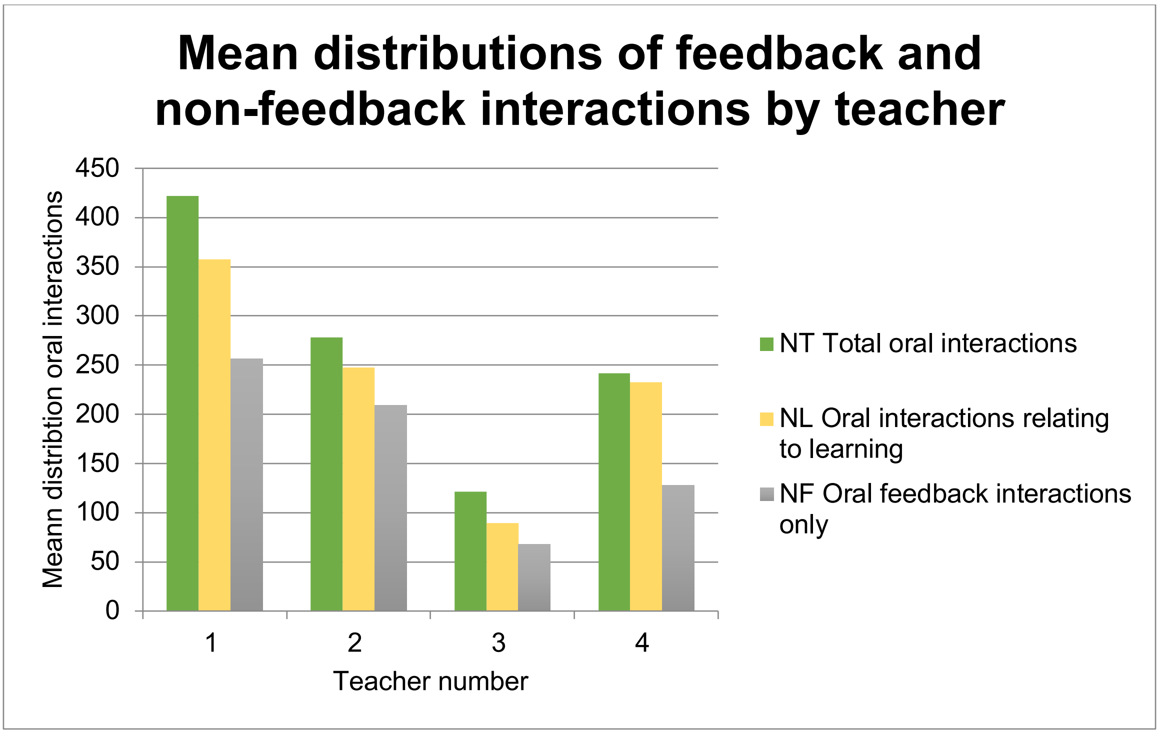
From this, it is possible to see the ratio of feedback-related and non-feedback related verbal interaction per teacher, sorted according to the first level, wherein only T, P, R and S are considered *feedback* interactions, whereas I (Instruction) and O (Other) are not. Additionally, types coded as O (Other) have been subtracted from the total observed oral interactions so that a distinction can be made between communication which is relevant to the *learning* overall, and that which is not, and this is presented below in Table 5.4. The totals for N_T, N_L and N_F have been divided by the number of lessons observed to obtain a mean amount per lesson.

Teacher	N_T *Total* oral interactions	N_L Oral interactions relating to *learning*	N_F Oral *feedback* interactions only
1	422.0	357.6	256.5
2	278.0	247.7	209.2
3	121.0	89.5	68.3
4	241.8	232.4	127.8

Table 5.4 Mean distributions of feedback and non-feedback interactions by teacher per lesson.

Table 5.4 and Graph 5.1 indicate that all teachers used interactions characterised in this study as ‘feedback’ and non-feedback (the *efficacy* of such feedback is not discussed here), and these vary in proportions between teachers:

- Both IG teachers had higher *total* oral interactions [N_T] than the CG teachers;
- T1 has a notably higher amount of total oral interactions [N_T], and this is 51% higher than T2, the next nearest;
- T4 appeared to have a broadly similar amount of both total and learning-related interactions as T2 in the same school, however once Instruction-category frequencies were removed to provide feedback-related interactions, the N_F was much lower than that of T2;
- Both IG teachers had higher interactions related to *learning* [N_L] than the CG teachers;



Graph 5.1 Mean distributions of feedback and non-feedback interactions by teacher

- Teacher 3 had much lower oral interactions in all cases, although it is important to note that fewer observations were made of T3, and that whole lessons were not recorded;
- Teacher 2 appears to have a similar N_T to T4, although when subtractions have been made for O and I codes for each of these, T2's *feedback* interactions with their class is much higher than T4 (209.2 compared to 127.8);
- Both IG teachers use more than twice as much feedback with their classes than T3 and T4; they provided between 1.6 and 2.0 times as much feedback to their students;
- All teachers used varying amounts of O interaction (range = 9-65), but the *nature* of these communications also varied. The 'Other' of T1 could often be characterised as 'banter', whereas the 'Other' of T3 was nearly all warnings for behaviour. This may have respectively impacted on relationships between teacher and class, and this will be explored in Chapter 7.

Table 5.5 presents a breakdown of these total oral interaction findings by teacher, across all observations. The figures are presented as both an aggregate count, and also a percentage of their individual interaction *totals* (N_T) for comparative analysis, since not all teachers were observed for the same amount of time. These figures therefore do not correspond to those in Table 5.1, which is of *feedback* types only (N_F). Individual breakdowns per lesson are shown in the sample spreadsheet in Appendix 15. These frequencies have also been displayed as graphs for comparative purposes and summaries have been provided.

Table 5.5 shows that:

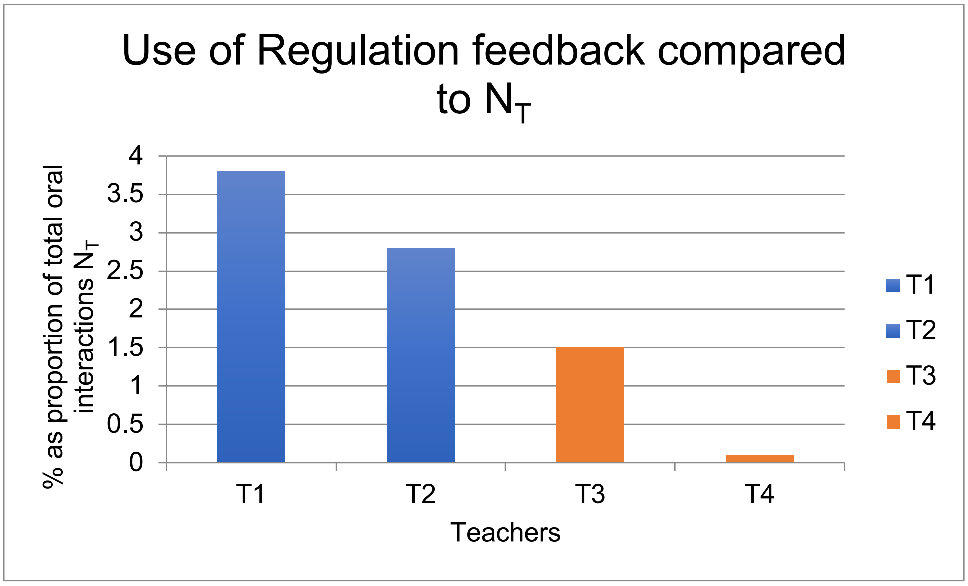
- Both T1 and T2 [IG] used higher amounts of Regulation feedback than the CG teachers, and this is shown in Graphs 5.2a and 5.2b;
- Both T1 and T2 [IG] used higher amounts of Process feedback than the CG teachers, shown in Graph 5.3. T3 had the lowest Process feedback at just 7.4% of all oral interactions compared to T1 (in the same school) at 23.9% and T2 at 30.0% (T4 was 18.8%);

Interaction type		Teacher 1	Teacher 2	Teacher 3	Teacher 4
First level – feedback *category*					
T	Task/product	31.6	41.9	47.5*	33.5
P	Process/strategy	23.9	30.0*	7.4	18.8
R	Self-regulation	3.8*	2.8	1.5	0.1
S	Self/praise	1.4*	0.5	0.0	0.5
I	Instruction	24.0	13.8	17.6	43.2*
O	Other (not related to learning goals)	15.3	11.0	26.0	3.9
Totals		100.0%	100.0%	100.0%	100.0%
Second level – feedback *mode*					
s	Statement	49.0	54.4	47.7	63.9*
p	Prompt	35.7*	34.6	26.3	32.2
Totals		84.7%	89.0%	74.0%	96.1%
Third level – target/recipient					
B	Male recipient	27.2	35.6	40.1*	22.8
G	Female recipient	26.4*	25.6	17.2	17.7
C	Class as recipient	27.5	24.5	16.7	53.8*
T	Team, or small group as recipient	3.6*	1.6	0.0	1.8
Totals		84.7%	87.3%	74.0%	96.1%

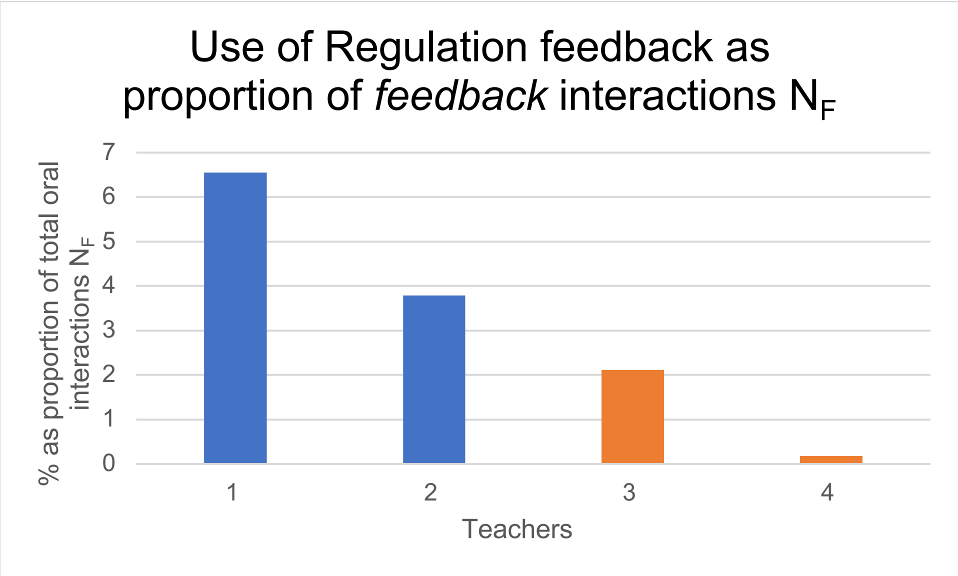
Table 5.5 Distribution of coded sematic units across all teachers (mean distribution of N_T as a percentage)

Shaded columns indicate intervention teachers. Highlighted values marked with an asterisk [*] indicate the highest values at each level for that teacher.

- T1 employed the lowest amount of Task feedback (31.6%); T3 used the most at 47.5% (Graph 5.4);
- The most frequent first level interaction was Task-related for all teachers in the study except for T4, whose most frequent category was Instruction-related (Graph 5.7).

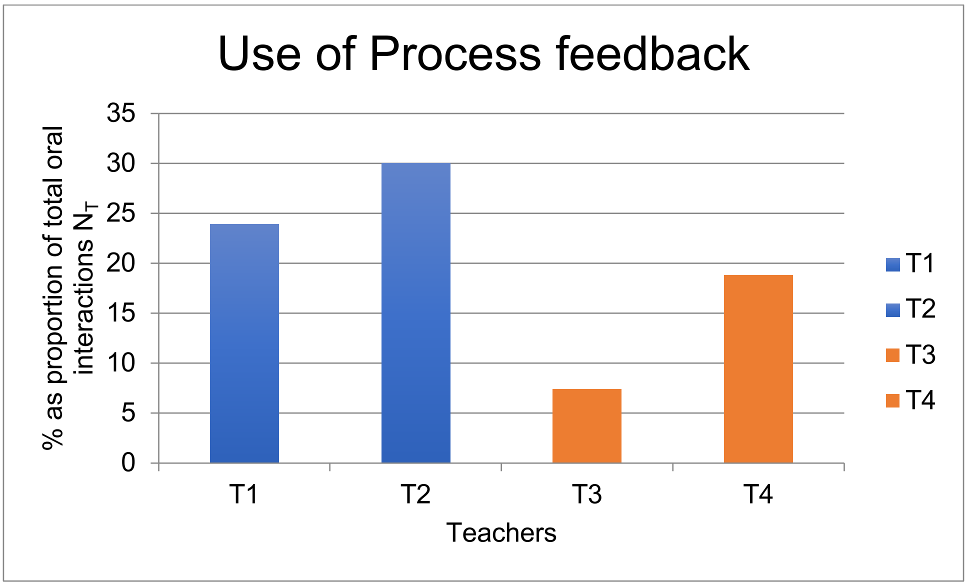


Graph 5.2a. The frequency of Regulation feedback used by the teachers in the study as a proportion of total oral interactions.

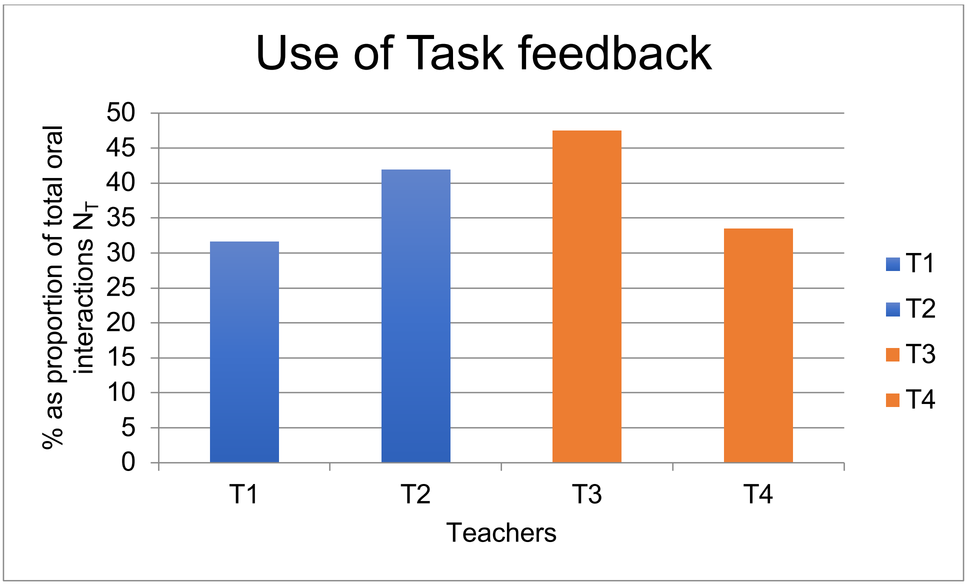


Graph 5.2b. The frequency of Regulation feedback used by the teachers in the study as a proportion of verbal feedback interactions.

In Graph 5.2a, the frequency of self-Regulation feedback is shown as a proportion of the *total* oral interactions given by the teacher (N_T), whereas Graph 5.2b presents it as a proportion of the verbal feedback only (N_F). For teachers in the intervention group, this is clearly higher than in similar typology studies (Gan, 2011; Hattie and Timperley, 2011; Bergh *et al*., 2012; Brooks *et al*., 2019).

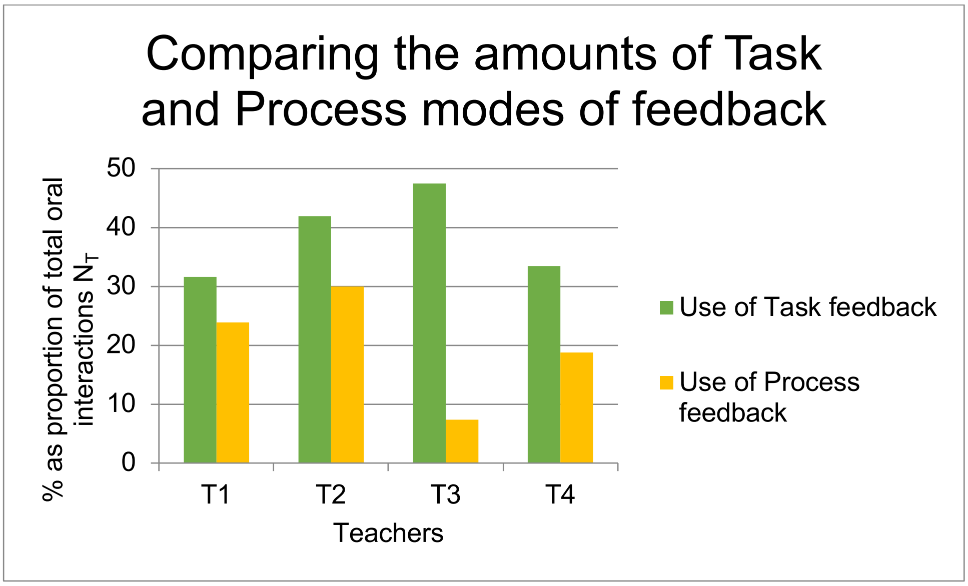


Graph 5.3. The proportions of Process feedback used by the teachers in the study.



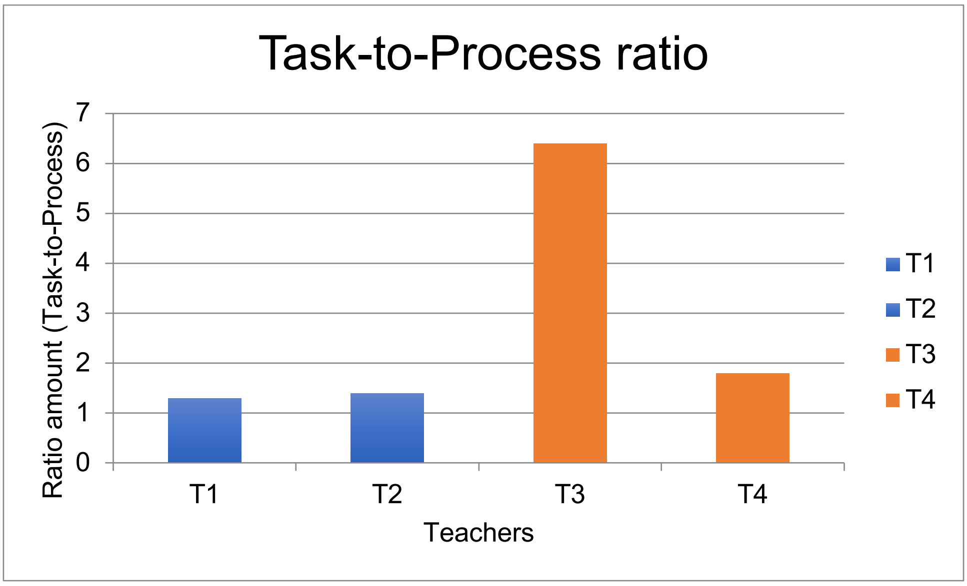
Graph 5.4. The proportions of Process feedback used by the teachers in the study.

Graph 5.3 indicates that T2 used the highest frequency of Process feedback to his students, and that both intervention teachers used more Process feedback overall than their partner teachers in the comparison groups. Of the T3 lessons observed, there was the smallest proportion of Process feedback, however this may have been due to the context of those lessons. As seen in Graph 5.4, this corresponded with the highest frequency of Task feedback observed in all teachers. A comparison of Task and Process teacher VF is therefore shown in Graph 5.5:

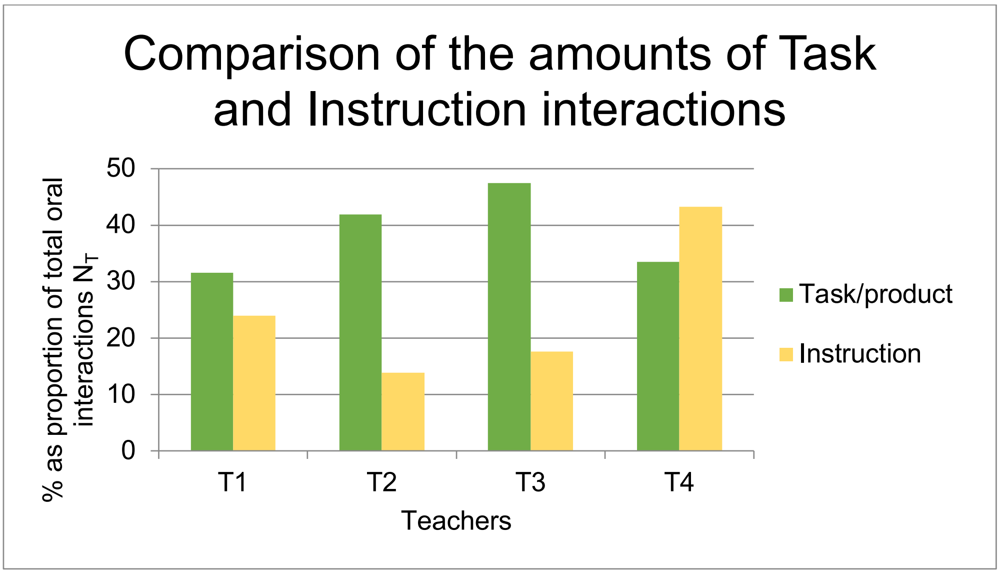


Graph 5.5. A comparison of the Task and Process feedback categories of the teachers in the study.

All teachers used more Task feedback than Process feedback, as shown in Graph 5.5, however these varied between teachers, and between the amounts of each category. By comparing the amount of Task feedback to Process feedback as a ratios (Graph 5.6), it can be seen that both IG teachers had the *lowest* Task-to-Process ratios [T1 = 1.3:1, T2 = 1.4:1], whereas the CG teachers were higher at T3 = 6.4:1 and T4 = 1.8:1. This means that the IG cohort received more Process-type feedback, designed to focus the students on *how* they were thinking and working (rather than directing the feedback at the task or product) than the CG cohort; *instrumental* rather than *executive* help (Hattie and Timperley, 2007).



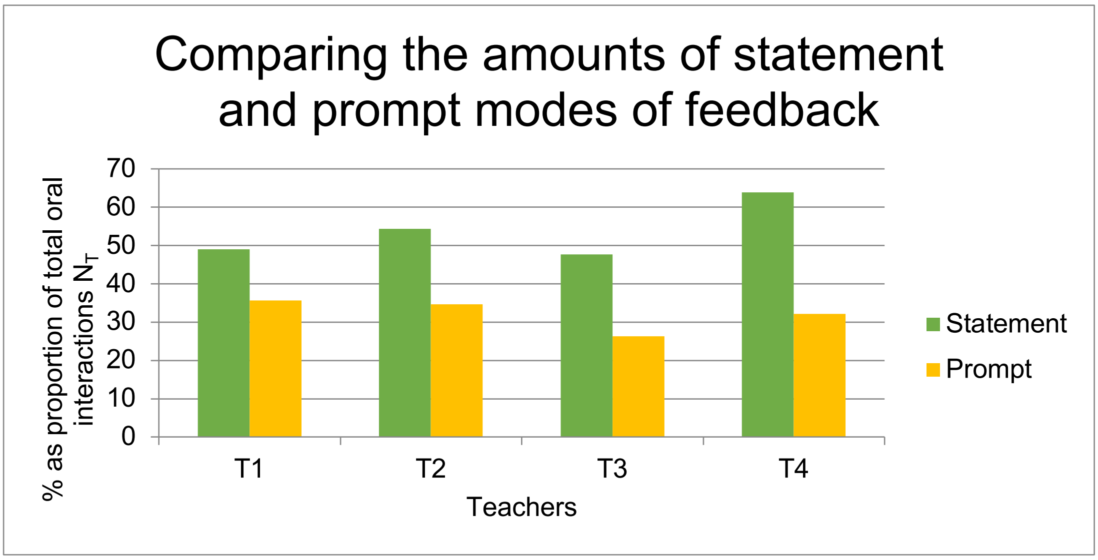
Graph 5.6. The Task-to-Process ratios of the teachers in the study.



Graph 5.7. A comparison of the Task/product and Instruction categories used by the teachers in the study.

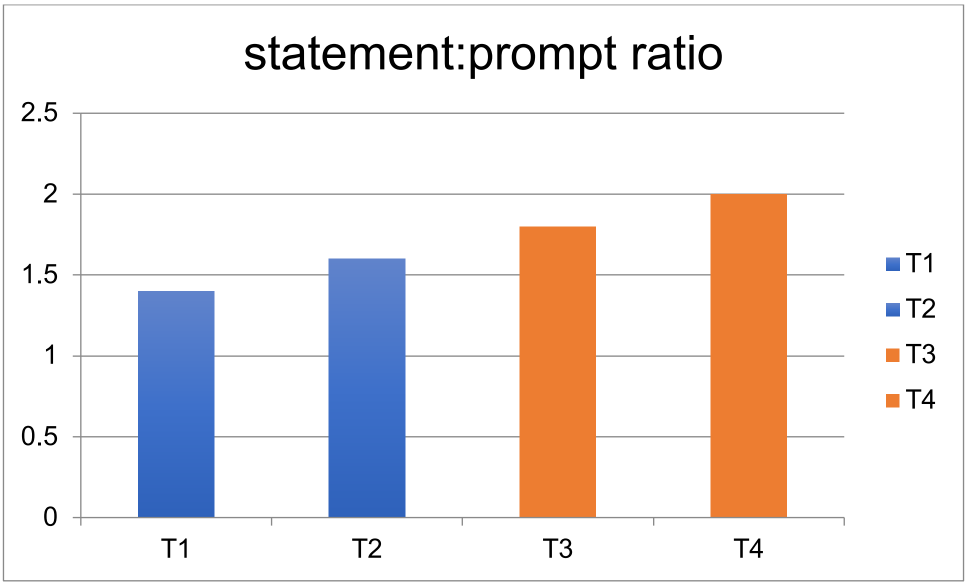
There was a difference between the intervention and the comparison group teachers in the amount of (board-located) whole-class instruction that took place; T3 and T4 used more whole-class instruction and interaction centred on board-work, whereas T1 and T2 circulated the class more, observed in lessons and described in field notes. Graph 5.7 shows the comparison of the amounts of Task and Instruction for each teacher, and here it can be seen that T4, unlike the other teachers, used more Instruction interactions (which were mainly whole-class) than Task-type.

Of second level interaction/feedback *mode*, statements (s) were more common than prompts (p). The statement-to-prompt ratio increased across Teachers 1-4, indicating that the IG students received *more* prompt 'next steps' feedback than those students in the CG. This is shown in Graphs 5.8 and 5.9.



Graph 5.8. A comparison of the statement and prompt feedback modes of the teachers in the study.

Applying the notion of a feedback loop re-imagined from Voerman *et al*., (2012, 2015) and Ramaprasad, (1983), wherein feedback statements and prompts form two sides of the loop (Figure 2.3), it can be seen that the intervention teachers, with *lower* statement-to-prompt ratios, 'closed the loop' more often than the comparison teachers. Consequently, intervention group students received more information on their next steps in learning than comparison group students; instrumental rather than executive help (Hattie and Timperley, 2007).



Graph 5.9. The statement-to-prompt ratios of the teachers in the study.

5.3 Teacher Profiles

Analysing the different proportions of the teachers' oral interactions with their students and enriching it with field note observations enabled teacher profiles to be characterised. The intervention teachers were observed throughout the study, whereas the comparison teachers were observed only at the start and end, so a less detailed, and less reliable picture can be built of T3 and T4. However, with a large number of observations of the intervention teachers, together with the detail provided by those observations, a representation of their classroom interaction style is offered. Features of how the intervention teachers attempted to change their feedback *in situ* is then offered in a later section, to provide evidence to answer research question 4 regarding the impact of the CPD intervention. For the following characteristics, the frequencies of feedback types are given as a proportion of the total oral interaction, N_T, unless otherwise stated.

Looking at the average number of feedback segments per minute demonstrates how language is not utilised equally by these teachers:

Teacher	Total length interactions in minutes	Sum of oral *feedback* interactions	Average number of oral feedback interactions per minute
1 [IG]	786	3591	4.6
2 [IG]	1108	4394	4.0
3 [CG]	146	273	1.9
4 [CG]	274	639	2.3

Table 5.6 Quantitative overview of feedback interactions per minute by teacher

Table 5.6 indicates that the intervention group teachers give on average twice as much feedback per minute than the comparison group teachers. Taken with Table 5.1 which shows that the total oral interactions, and the interactions related to learning are higher for IG teachers, it would appear that there are differences in the sociocultural learning environments between these conditions (Alexander, 2014). These numbers might imply that both IG teachers simply talked a lot more than their counterparts in the sense of ‘talking at’ their class, however this was not observed, and it is possible to demonstrate richer aspects of these teachers’ classroom practice by interrogating the data.

5.3.1. Overview of Teacher 1

T1 had the lowest Task category amount of the teachers in the study, whereas the Process and Instruction categories had similar values at 23.9% and 24.0% respectively.

Both intervention teachers used higher amounts of Process feedback; T1 Process feedback at 24.0% was noticeably higher than that of T3 at the same school at 7.4%. The value of 24.0% for T1 is consistent with both the Hattie and Masters (2011) and the Gan (2011) studies however, indicating therefore that from the lessons observed, T3 used sizably lower amounts of Process feedback.

The opportunity to use Process feedback is enhanced when the context of the physics lesson is focused on ‘working out’. When the teacher demonstrated this on the board to the whole class, this was categorised as Instructional, however it is not prejudicial to suggest that an aspect of this was also modelling the thought process, and as such would qualify as Process. This was not coded as such however, since it was not considered to be feedback *per se*. This again demonstrates the interwoven complexity of feedback with other classroom pedagogy, however as it remains a form of both comment/clarification/correction (*progress*), and content or process development advice (*discrepancy*). When T1 circulated to monitor how the students were progressing, learning conversations were common, and in coding, arrows were used to indicate staying with the individual to form a chain of interaction, usually feedback:

So how many months each year will there not be enough solar energy?	PpB
So, if you draw a line across 16.8, OK?	→PpB
So, if that’s 17.5…that must be around 16…17...so it’s going to be around about here, yeah?	→PpB
So, its 1, 2, 3, 4, 5, 6, 7 months that it doesn't reach that level that it is the level it needs to be able to do that job.	→PpB
You can almost do it, here and here, but not quite.	→PsB
So, in order to make that amount of energy, it has to be at that height, and at only those months out of the year would you be able to do it.	→PsB
Does that make sense?	→PpB
So, all you did…its dead easy…	→RsB
You’re calculating the energy	→TsB
So, each day, the average European family uses this…	→PsB
OK, calculate the energy, you’re finding the energy,	→PpB
You don’t need to rearrange	→PsB
You wrote this down	→TsB
Then all you needed to do was pull the amounts out of here,	→PsB
So, m = 100, c = 4200 and theta, the change in temperature was 40.	→TpB
So, all you have to do is put the numbers in and press = on your calculator.	→RpB

The difficulty is when you have to rearrange it, or pulling the information out of this here, and that's what you have to practise.	→RpB

Teacher 1, lesson 5, 08.49-10.15

Here there is a focused conversation on process in a combination of Process-prompt, e.g., advice to use a strategy; ‘calculate the energy, you’re finding the energy’ and Process-statement, where the strategy is modelled; ‘So, m = 100, c = 4200 and theta, the change in temperature was 40’. This was punctuated by product-related statements and prompts such as, ‘You wrote this down’ and ‘all you have to do is put the numbers in and press = on your calculator’, culminating in an encouragement to practise, i.e., a self-regulatory recommendation.

Both intervention group teachers used higher amounts of Regulation feedback than the comparison group teachers. T1 used nearly three times (2.8) as much Regulation feedback as T3 in the same school. Following the CPD intervention, T1 used a differentiated approach to homework so that the students would experience a ‘can-do’ approach at a level to suit themselves. In moving around the class and distributing the sheets, T1 would sometimes preferentially give one type, perhaps with a suggestion of moving onto the other one afterwards, or free choice. There was a sense of expectation of effort and completion, and the tone of voice was positive and occasionally beguiling:

You can manage it; it’s not going to take that long	RsG
Do you want the easier one, or the one with the calculations?	RpG
I would suggest you do this more challenging one on refractive index	RpB
If you want all three sheets, you can, or you can choose between these two	RpB
That one’s explanations, and that one is calculations	TsG
What are you doing? You’re doing that one, that’s fine [*more challenging*].	TpB RsB
Do you want this one, and then you can come back for the other one? [*girl asks for both*]	RpG

Good yes, of course you can [*another girl asks why would you want both?*]	RsG
To challenge yourself	→RsG

(Teacher 1, lesson 4)

According to Hattie (2012), regulation feedback may include assisting learners to monitor, and direct their own actions, engage with tasks, and expend effort. T1 used some whole class instruction and modelling often at the start of the lesson, then would move the students to independent work, including practical activity. There would sometimes be an opportunity for the students to make decisions about their own physics learning, and these were useful occasions to direct the students into more self-regulatory habits:

[*all*] Year 10! get yourself a book. Work out what it is that you need to look at next lesson.	RpC
So that when I come round and ask you, you can tell me what I'm doing my quiz about.	RpC
[*boy says heat particles!*] No, no heat particles!	O
[*girl*] Right, so what do we think?	RpG
Everything. Oh excellent.	O
So do we want to do a quiz on a bit of everything?	RpG
Or do you want to do each lesson on something different?	→RpG
Gosh, your enthusiasm girls is unbounding. (*sic*)	→RsG
[*girl*] that's not very good	→SsG
You can, you can do it!	→RsG
Stop being so negative.	→RsG
You can do it.	→RsG
You're both actually very good at it.	→SsG

(Teacher 1, lesson 9)

T1 had the highest amount of Self/Praise (undifferentiated) feedback of the teachers at 1.4%. This is much lower than the frequencies in the studies of Brook *et al*., (2019); Hattie and Masters (2011), Gan (2011) and Bergh *et al.* (2013), but higher than the other teachers in the study. This was often characterised as a swift 'well done', 'very good'; terms that T1 later herself described as ‘back-slapping’. Somewhat ironically, and counter to the received wisdom that generic praise is ineffective and sometimes detrimental to both learners and learning (Brooks *et al*., 2019; Hattie and Timperley, 2007; Kluger and DeNisi, 1996; Mueller and Dweck, 1998; Shute, 2008), this praise could be seen to lubricate interactions, and the students felt valued by their teacher. Voerman *et al*. (2014) suggest that such feedback may influence the emotional space within the classroom, and that a more nuanced view on Self-category feedback could be adopted.

T1 had 15.3% of oral interactions coded as Other; much of this could be characterised as ‘banter’ with the individuals in the class, with whom there was a good camaraderie observed. In this short vignette, T1 is chatting with two boys, one of whom was Billy in the focus group, during an investigation. The student response was not transcribed, but T1 is teasing them over a timer, and all T1 responses are coded as ‘Other’:

> I’ve put it in my drawer so that it’s out of the way [*to Billy*]
>
> So that you don’t feel the need to set my alarm over and over again to irritate me over and over again Billy.
>
> Yes, that’s why you were looking for it.
>
> Yes, it’s away, its hidden from you. *My* stop clock.
>
> Yes [*to another boy*]. I have a stop clock, a magnetic one, that Billy likes to set the time on so that it goes off later.
>
> What was the other thing I noticed the other day…and I knew it was you lot?
>
> Oh! [*to boy, remembering incident*]
>
> I didn’t have a go at you, not really. [*to boy*]
>
> Yes, I know [*back to Billy*]
>
> (Teacher 1, lesson 1, 30.02-30.50)

The interaction took less than a minute and then the teacher quickly checked that they had taken a time measurement on the activity and rapidly passed on to another individual. As described for Self-category feedback above, interactions not related to the learning may enhance the emotional space of the learning climate.

T1 had a higher Instruction frequency at 24.0% compared to T3 at 17.6% in the same school. This category was mostly associated with whole-class recipients at the third level since T1 often started the lesson with modelling or explanation at the front, moved to more individual activities and then drew the lesson to a close with a whole-class instructional plenary.

At the second level, feedback *mode*, T1 had 49.0% statements compared to 35.7% prompts of total oral interactions. This is a statement-to-prompt ratio of 1.4:1. Graph 5.9 shows the statement-to-prompt ratios of all the teachers in the study. This means that the T1 students received more 'next steps' or discrepancy advice than the students in the other groups. T2 was the next lowest ratio at 1.6:1, indicating that both the intervention groups received more prompt feedback than the comparison groups;

In terms of the third level, T1 gave approximately equal feedback to males, females and the class at 27.2%, 26.4% and 27.5% respectively. There were 12 males and 9 females, which could account for the slightly higher male value for target B. T1 also had the largest amount of Team or group feedback at 3.6%, which was more than twice that of the other IG teacher, T2, and twice that of CG teacher T4. [T3 did not give team or group feedback in any of the observed lessons];

T1 gave the highest amount of Team/group feedback of the teachers in the study. On observation, this seemed to be a consequence of the way in which the lessons were conducted: initial teacher input (e.g., modelling), then students working individually or in groups to complete tasks. There were high monitoring and discussion rates as the teacher circulated. This is supported by the considerably higher interaction rate of T1 at a mean of 422.0 per lesson, as seen in Table 5.5.

The classroom style of T1 was characterised as energetic, 'everywhere', bantering and supportive. T1 acknowledged themselves as sometimes the hardest working person in the room and that they could be quite controlling of the learning approaches. One of the most notable features was the affection for the class as a whole and for individuals within it although there was no preferencing. Equally, the students held T1 in high regard and there was a sense of trust, which authors have highlighted as a key factor in receiving feedback (Carless, 2009; Eva *et al*., 2012; Gamlem and Smith, 2013: Dann, 2019; Elbra-Ramsay, 2019) The teacher-student relationship is explored further in Chapters 6 and 7 from each of their perspectives.

5.3.2. Overview of Teacher 2

T2 had the highest amount of Process feedback, and the second highest amount of Task feedback of the teachers in the study. However, when the Task-to-Process feedback ratios are considered, T2 had the second lowest T:P ratio (both IG teachers had the lowest Task-to-Process ratios). T2 had identified Process feedback as a key area to work on as he was motivated by developing the students into more independent learners. Early in the study, T2 was able to use Process feedback as he circulated the room, but this tended to be statement rather than prompt at the second level as he explained the thinking, rather than prompting the thinking:

OK, that was just showing where the unit for acceleration came from	PsB
It's a similar thing up there [on the board]	PsB
You've got to think of metres per second as just being a speed	PpB
It's in its own little box with brackets around it	PsB
It is [*confirming a boy's question*]	TsB
It's velocity, or rather I should say it's change in velocity	PsB
So, I'll write it out. It's change in speed per unit of time.	PsB
And speed is distance divided by time	PsB
Because the other thing they'll [exam questions] make you do is work out the speed, and then the acceleration afterwards	PsB
To change it up and make it more complicated	RsB
But the plus side of those is that if you make a mistake in the first half and get the speed wrong, you'll still get the marks for using the acceleration equation right	RsB
So, practise doing those for yourself	RpB

(Teacher 2, lesson 3)

T2 had a frequency of 2.8% Regulation feedback, whereas T4 in the same school was just 0.1%; this was also nearly twice as much as CG teacher T3 at 1.5%. After the CPD intervention, T2 saw the opportunity to encourage students in self-regulatory thinking when circulating and monitoring learning:

The next step for you is a bit more challenging one	RsG
Why is it not possible, or very difficult to run on ice?	PpG
Use the forces to explain why it's difficult to run on smooth ice.	PpG
Just a challenge for you for the rest of the lesson	RsG
I like how you're thinking; let's focus on the forces	RsB PpB

Similarly, to T1, T2 used differentiated worksheets to challenge students at different levels, and was also supportive in their choices:

These ones are really, really good practice	TsG
You can write an example down yes, but I want you to be doing really good, hard, effective practising	RpG
So, writing down an example does not mean that you're learning it, it means you're mechanically copying something down	PsG
…both [worksheets]. This is optional, so that's the trickier stuff; but these things usually start simple and work their way up.	RsG

Occasionally, Regulation category feedback arose as a negatively phrased response or question, such as 'Can the four of you work as a four or is that not going to work?' (RpT), but generally it was phrased positively, and occasionally in a way that would help reduce physics anxiety:

Is that the furthest forwards he goes?	TpG
Well what do you think?	PpG
If that's what you think then use that number	PpG
[*girl answers in negative*] well then, we'll talk about it	TsG

Don't be frightened of making a mistake,	RpG
this is part of the learning process	RsG

As a teacher of physics, T2 was able to bring the nature and processes of working scientifically into the learning rather than a focus on the physics content only:

People used to make this argument; I don't need to wear a seatbelt because I will just hold onto the steering wheel, and I'll just brace, and I'll be safe in a crash	PsC
So that's what the practical is about; it's about seeing whether that argument has any scientific validity	PpC
Are you happy with the consistency of that?	RpB
What's the speed?	TpB
OK, do one more and check it.	TpB RpB

(Lesson 8)

A good example of building resilience in physics learning is using the troublesome topic of electricity, for here students find circuits both conceptually and practically challenging, especially when circuit faults are not obvious. This may generate feelings of physics anxiety and low self-concept as students struggle to find answers. More than half-way through the study in a continuation of a circuits lesson (Lesson 12) on resistance, there were issues connecting ammeters and voltmeters correctly. A large proportion of the interaction was concerned with T2 talking the students through why ammeters go in series and voltmeters in parallel as the circuits were 'trouble-shooted' further:

Well that's why I'm making you do all these practicals so that you get more and more of a feel for what works and doesn't work in the circuits	RsG
So, the effort will pay off, if you see what I mean?	RsG
If you don't understand why you're doing something, then you must feel free to ask.	RpG

(27.15-27.36)

The practical concludes twenty minutes before the end with T2 asking for individual results and collating on the board. The teacher then uses them as ‘field scientists’ to derive the relationship between the values, demonstrating student ownership of the learning to them:

> Neil: The amps plus the voltage gives the input voltage.
>
> T2: Hmmm, amps plus voltage..?
>
> Neil: Yes, gives you the input voltage.
>
> Ava: that works for our data too.
>
> T2: it works for your data? OK, let’s check that. Do the amps plus this voltage here give the total voltage that the power supply is supposed to be putting out? Check that. Everyone who’s got data, check whether that holds.
>
> T2: Edward, does that hold for you?
>
> T2: It doesn’t hold for Edward’s set-up. There is a problem with that.
>
> T2: Add I and the V_1 together to see if it equals V_T.
>
> T2: OK, so we’ve got a suggestion that V_T equals I plus V_1 [*writing on board*] for these groups at the back. Doesn’t work for Karl and Andy, doesn’t work for Edward…
>
> [*discusses value differences between groups*]
>
> T2: OK, excellent first attempt, doesn’t seem to be borne out by the rest of the experimental evidence. Hey this is proper science, right? A result from one team that seems to say one thing, we’re verifying it independently by other researchers working in the field – yes, we are; we’re doing insta-peer-review [sic]. Good stuff. No, this is really good, I am really happy.
>
> T2: We need to find a different relationship here; this one is not working.
>
> T2: What other one could we use? [*students make suggestions*]
>
> T2: When you increase the voltage from the battery, the voltage and the current reading increases. So, hang on a second; if I increase the voltage here, not surprisingly the voltage goes up. Does the current go up as well? [*class agrees*]
>
> T2: So, we can start by saying one thing. We can say the voltage appears to be proportional to the current. It looks like voltage is proportional to current.

[*boy*] there might be something reducing the voltage, see, given that I have so many components, the voltage might be distributed over more components?

T2: so how can we put that? If we have more resistance, we have less current, do you agree?

[*boy talking indistinctly*]

T2: Look, Edward's come up with an idea. He says look my circuit has a lot of components in it, that means it must have more resistance than Yvonne's circuit, so the idea is that current and resistance are opposing each other. If I have more resistance, I have less current. (36.05-40.50)

The Ohm's Law relationship is then introduced and discussed, and subsequently consolidated by question practice, including a challenge question of calculating the resistance using power information, previously covered in the academic year. This was a confident, engaging section in which the students were challenged to create a relationship from their empirical results, modelling working scientifically as they did so. Although there was teacher talk here, it was student learning that was socially constructed in the dialogue (Alexander, 2014).

Overall, T2 used very little undifferentiated praise overall (S=0.5%), which is much lower than the feedback studies cited above; there was specificity in the way he offered responses.

At the second level, T2 had the second highest amount of statement interactions at 54.4%, however when the ratio of statement-to-prompt is considered for T2, this was found to be 1.6:1, and this is lower than those of the CG teachers [see Graph 5.9]. As before, having a low statement-to-prompt ratio indicates that T2 students received more 'next steps' guidance or discrepancy feedback than comparison group students

T2 interacted more with the males (B=35.6%) than the females (G=25.6%). Although there were only two more males than females, more males were interacted with in an effort to keep them on task and engaged. Almost one-quarter of all oral interactions were to the class (C=24.5%), and 1.6% to groups. It was noted during observations that T2 circulated the class less frequently than T1.

The classroom style of T2 was characterised as appearing extremely knowledgeable, and that this subject knowledge was frequently both connected and contextualised. His demeanour was slightly teasing and ‘dry’, showing a good sense of humour although the students did not always understand the references. T2 was more reserved with the class than was T1 and connected less well with them personally. It was clear however that he knew them, and their individual learning needs well.

5.3.3. Overview of Teacher 3

T3 had the highest value of Task-related interactions (4.5% of N_T) combined with the lowest amount of Process, yielding a T:P ratio of 6.4:1, which is considerably higher than the other teachers in the study.

T3 used 1.5% Regulation feedback but was not observed to use undifferentiated praise at all (S=0.0%) in the lessons. However, the majority of the responses coded as category R were affect or motivational requests such as ‘Are you okay with this?’ and ‘do you need any help with anything?’ rather than specific assistance to help them direct their own efforts to regulate their learning.

The teacher had the highest amount of oral interaction coding as Other at 26.0%, and a large proportion of this related to behaviour management. From observations in class, there were higher noise levels than in the other teachers’ lessons, a higher proportion of being off-task and associated behavioural issues;

Initially, T3 appeared to have the lowest number of statement interactions; however, this was impacted by the removal of the high number of Other and Instruction category codes. When statement-to-prompt ratios are considered, T3 had a ratio of 1.8:1, thus higher than the IG teachers, as in the N_F tally, 60.1% were statements and 39.9% were prompts.

40.1% of total interactions were directed at males and 17.2% to females; however, there were 17 males and 6 females in the group, which would account for the imbalance. 17.5% of interactions were directed at the class overall, and none to small groups in the observed lessons. The females were described by the teacher as ‘very quiet’, and somewhat intimidated by the larger number of boys;

T3 did not circulate and monitor in the lessons observed, and mostly remained at the front desk. As only 4 lessons were observed, it is not possible to say whether this was typical practice.

The classroom style of T3 thus appeared to be ‘from the front’. There was continued reaction to the behaviour issues in the class, the dynamic of which was affected by the much larger number of males; there were observed difficulties with certain individuals or groups working together or near each other. T3 remained calm, direct and focused on both the physics content and the direction of travel of the learning, but without exhibiting much connection with the class. It is very possible that T3 exhibits a different teacher persona with another class; this was not observed. Whilst there seemed to be limited affection for the class; T3 was concerned about the small number of females who were described as ‘being very hard to get anything out of’ and ‘they tend to stick together’ (field notes, T3, lesson 1).

5.3.4. Overview of Teacher 4

T4 had the second lowest amount of Task feedback but combining with a Process value of 18.8% gives a T:P ratio of 1.8:1, the highest of these ratios. The students in this class thus received almost twice as much Task/product feedback as they did Process feedback (Graph 5.6), showing more comment on *what* had been done, rather than *how* it had been done. Similarly, at the second level, the statement-to-prompt ratio was 2.0:1 for T4, meaning that the students received twice as much comment/ clarification/correction type feedback as ‘next steps’ advice. This was also the highest s:p ratio of all the teachers in the study (Graph 5.9);

T4 was not observed to use Regulation feedback overall (0.1%, comprising one comment which was ‘do you think you can work that out?’ in lesson 3). The amount of undifferentiated praise feedback was low at S=0.5%, identical to T2 in the same school;

The amount of interactions coding as Other was the lowest of all the teachers in the study at just 3.9%, and much of this arose from calling for quiet and dealing with behavioural issues. There was not much ‘banter’ with the students. One main difference between T4 and the other teachers in the study was the amount of interaction coded as Instructional (I=43.2%). This was nearly twice as much as that utilised by T1, and over three times as much as T2 in the same school. All the other teachers had the Task category as their largest first level value; for T4, this was the Instructional category (Graph 5.7).

The other distinction in T4’s practice compared to the other teachers was in the amount of class-directed interaction (C=53.8%). Much of this also correlated with the first level Instructional category. This was more than twice the class-directed interaction than T2 t the same school. In lesson 1, T4 taught by the board for 36 minutes before releasing the students into individual activities. In modelling answers on the board, T4 would give the equation, substitute numbers and provide answers rather than ask students to contribute, which seemed a disabling approach rather than an enabling one. All observed lessons mirrored this front-loaded process.

Despite having 15 males and 4 females in the class, there was a broadly similar amount of male and female recipient feedback (B=22.8%, G=17.7%). The females were spaced around the room and worked well with the males; they contributed well in each lesson, exhibiting confidence. There was a small amount of Team/group directed interactions as the teacher circulated (T=1.8%), but from observation, T4 usually interacted with individuals rather than groups once circulating and monitoring began.

T4 was a relatively new teacher of physics, and the style was characterised as highly instructional, and directed to the whole class. T4 was observed to teach ‘from the front’, but in providing the answers during the explanation, did not use the opportunity for the students *themselves* to experience the thinking process. Instead, the students’ only undertaking at this point was to copy material into their books, as doing rather than learning, before then working on their own questions. Consequently, many first level Task and second level statements (s=63.9%) were confirmatory acknowledgements of successfully answered questions. T4 used names well and appeared to have a good connection with the class. A dynamic and animated presence as well as the use of a variety of contexts to illustrate the physics content appeared to be indicative of practice within those observed lessons.

In summarising from the evidence, the study has found that it was possible to answer the question ‘*What types and proportions of feedback do teachers use in their verbal interactions with students in physics?’* in providing a rich overview of a range of oral interaction frequencies, and combined these in analysis to show how classroom teachers can develop a dialogue profile that is very much their own.

The teachers in the intervention group overall demonstrated that they used more level 1 Process and Regulation feedback than the comparison teachers, as well as more second level prompt. To enable RQ4 to be answered, we must investigate whether these intervention teachers naturally had these differences from the start, or whether differences arose as a result of undertaking CPD to flex their feedback responses:

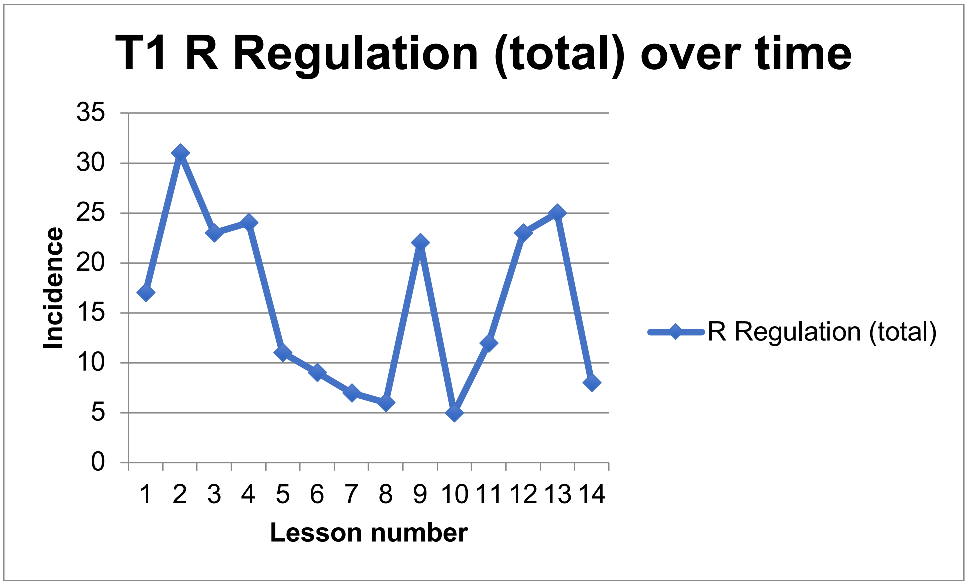
4. *Does the feedback pattern change during the life of the study in response to a CPD programme of flexing verbal interaction styles?*

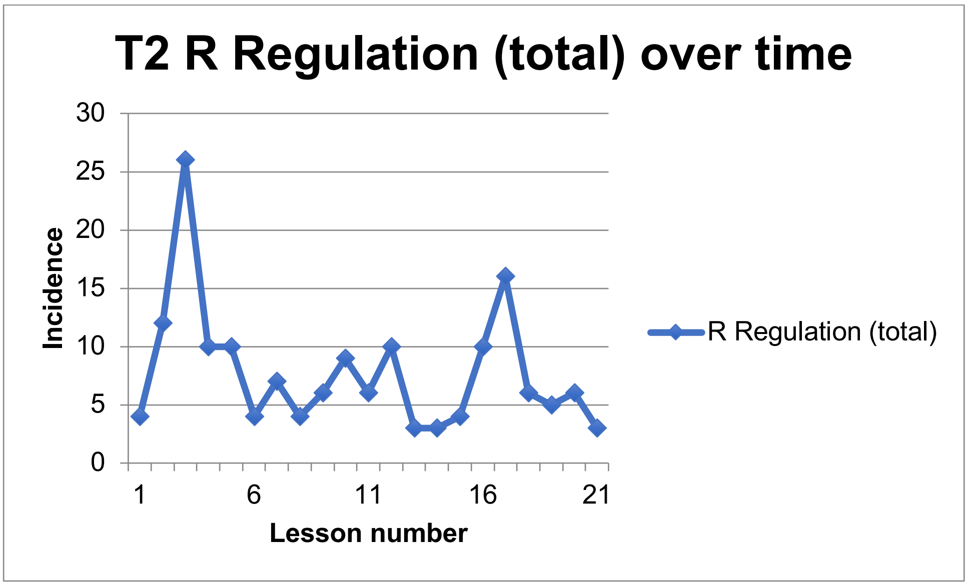
5.4 Types and proportions of feedback over time

5.4.1 Towards a self-regulatory ideal

The main focus of the study, the use of self-regulation feedback, displays a varied pattern, as shown in Graph 5.10. The sample size of Regulation feedback is too small to support a reliable trendline, however several large spikes occur in the use of Regulation feedback, which follow the coaching conversations held with the teachers in their schools (see Table 7.1). Since the coaching took the form of feedback, reflection and practise, the teachers had the opportunity to try to phrase some of their observed classroom responses differently, and both intervention teachers reported that this was helpful in their efforts to flex their feedback (see Chapter 7). They also acknowledged that they found this extremely challenging to do ‘in the middle of teaching’ (Eraut, 2004; Hammerness *et al*., 2005; Kothagen, 2010; Voerman *et al*., 2015), and this study has found that Regulation feedback can be quite contextualised, so the teachers felt less able to insert it unless it made sense in the learning instance, and that this was more likely to happen on an individual basis.

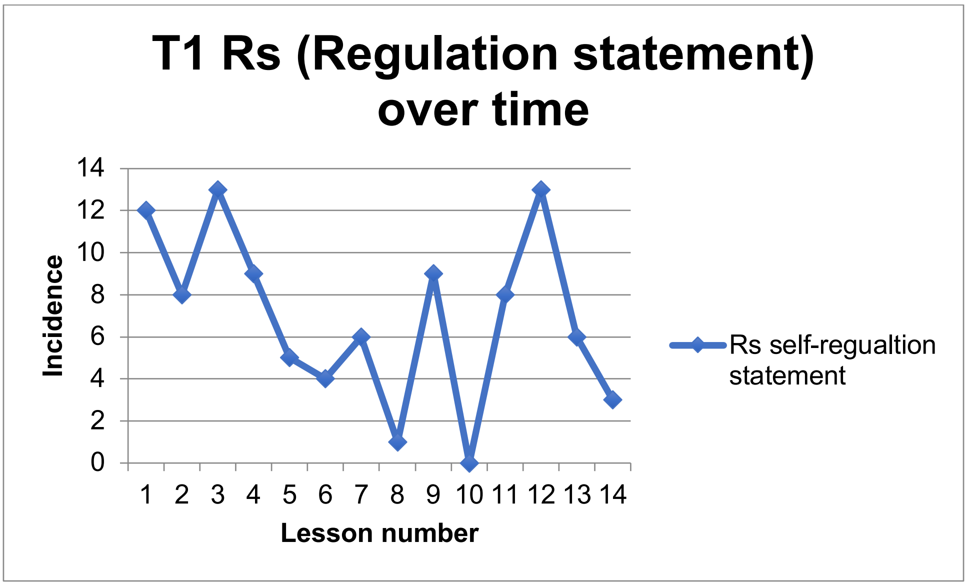
Both teachers reported that they felt most able to give Regulatory feedback when circulating and supporting individuals during independent work. When one considers Hattie’s assertion that when one gives feedback to the whole class, it is received by no-one (2012), it would seem most beneficial when regulatory feedback is given in a calibrated manner to a student to enable them to direct efforts towards their own learning goal. Feedback at this level ‘fosters the willingness and capability to seek and effectively deal with feedback, to self-assess and self-correct, to attribute success to effort more than to ability, and to develop effective help-seeking skills’ (ibid:20). Graph 5.10 shows how the deployment of Regulation feedback varied over time for both IG teachers.

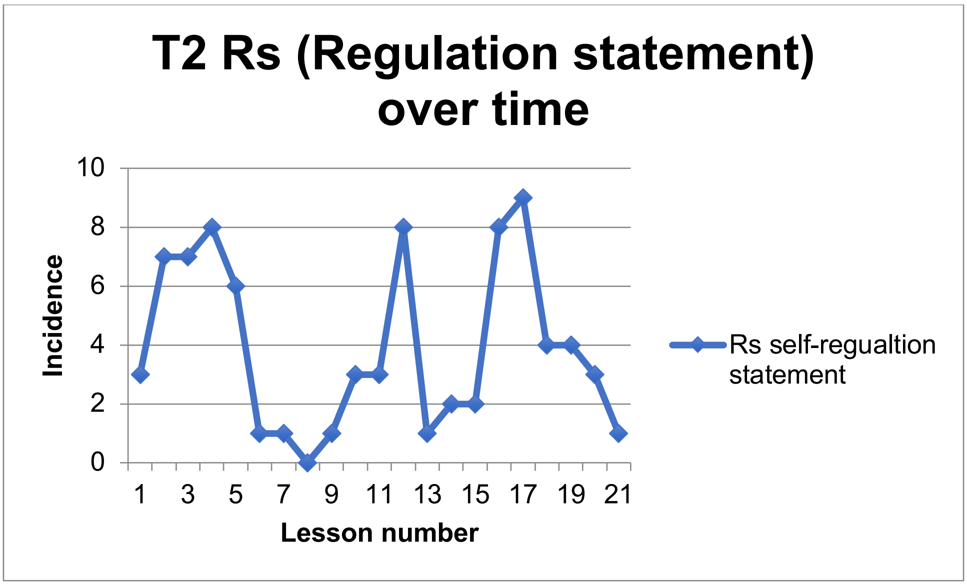




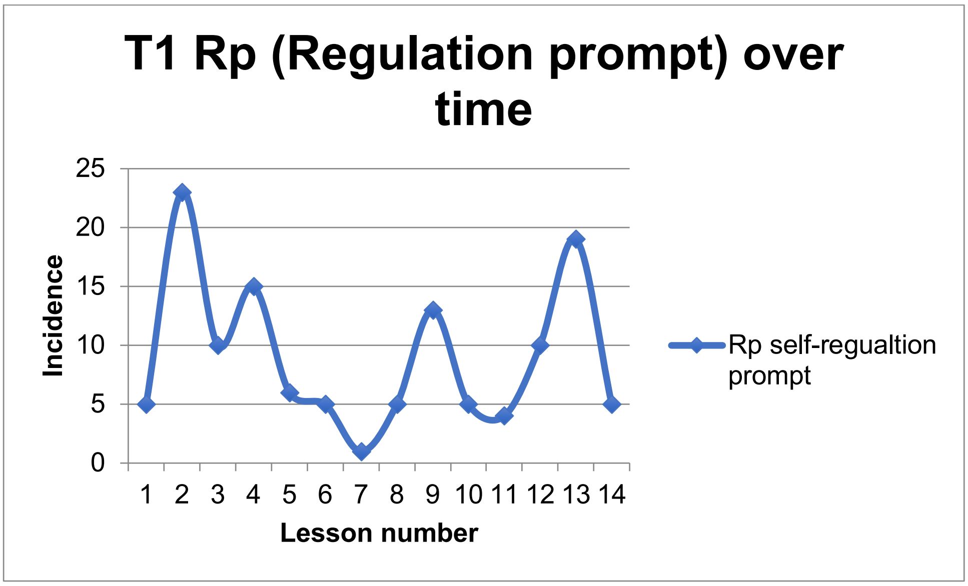
Graph 5.10 The change in total use of Regulation (total) feedback by the IG teachers over time.

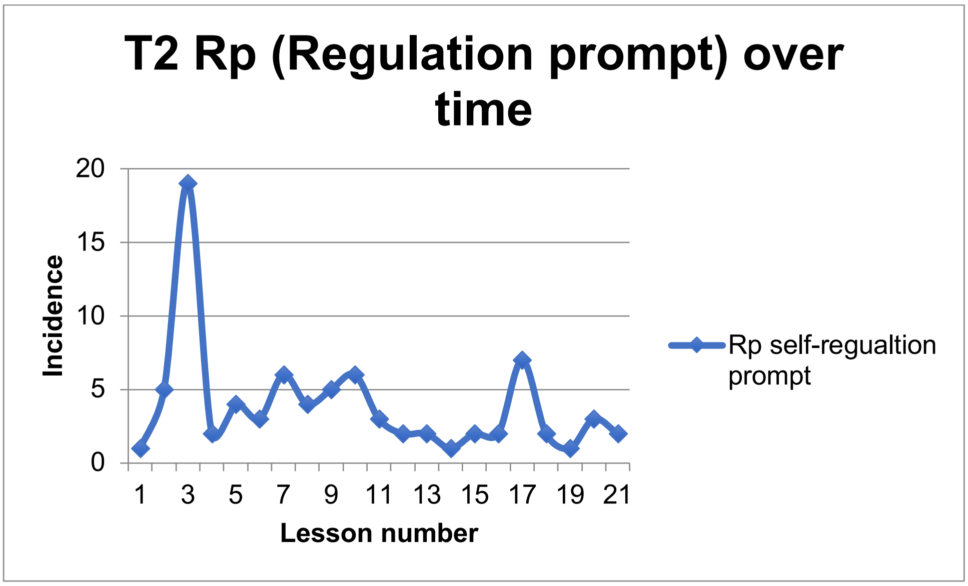
Examining Regulation feedback in terms of its second level shows slight differences in statements and prompts between the teachers. Again, the peaks appear to follow a meeting in which the teacher rehearsed flexed feedback responses (Graphs 5.11 and 5.12). For T1, there is continued variation in incidence of both second level types, however for T2, there is more variation in statement than there is in prompt, and what appears to have happened here is that T2 fluctuated in ability to provide statements indicating commentary on (e.g.) recognising a student's effort, engagement, self-evaluation and even willingness to accept feedback, but was less able to provide 'feedforward' in terms of these characteristics. This would involve not just recognising the above, but immediately re-phrasing a statement as perhaps a question, and asking: 'Do you see why you got it wrong?'; 'Can you find the place where the error was made?'; 'What strategy did you use?'; 'Can you think of a different strategy?'; 'Which do you think is the better strategy?'; 'How will you know if your answer is correct?'





Graph 5.11 The change in total use of Regulation-statement feedback by the IG teachers over time.





Graph 5.12 The change in total use of Regulation process feedback by the IG teachers over time.

In the latter half of the study, lesson 12 for T2 on building and testing circuits offered a good opportunity to practise self-regulatory habits, as described above. In trying to help students 'trouble-shoot' their problematic circuits, T2 provided some motivational statements and prompts:

[*girl experiencing difficulty*] It's certainly tricky [*to build it correctly*] in several ways	RsG
First, you can't see what's happening, secondly, sometimes the equipment doesn't work on you	TsG
It does require a certain amount of patience and a certain amount of – [*girl says she doesn't have that*]	RsG
Well I think you do, I disagree, I would argue that you do	RsG
It's just that when something doesn't work you have to sit there and think, 'how can I fix this?'	RpG
So don't think that you've permanently broken it, let's just see where all this spaghetti is going…	TsG TpG

(30.02- 30.41)

T2 then guided the girl through how to build the circuit, explaining the strategy, and when it still did not work, discussed how to check each component, looking for (e.g.) a damaged lead until the fault is found.

Did you see what I did there? I experimented, and I tried a few combinations of things to try to isolate where the problem is.	PpG RsG
You don't have to be an engineer to do it, it's just trying different combinations until something works, OK?	RsG

The total supporting interaction takes just under three and a half minutes from 29.31-33.00. The noise level from the rest of the class remained manageable.

Near the end of the study, T1 had gained more confidence in flexing feedback and had practised some regulatory phrases (See Chapter 7). Lesson 13 (of 14) contained several passages where T2 had provided both R-s and R-p feedback upon circulating and supporting individuals' work. The context of the lesson was momentum, and the difficulties encountered by some of the students in processing the equations afforded T1 an opportunity to not only model and prompt the strategies, but also discuss and stimulate self-regulation and motivation in succeeding at it:

You stuck already? [*girl says 'I don't get it'*]	RsG
You do, I bet you get something.	TpG
Have a go, that's good!	RpG
Let's have a look, yes that's right	TsG
So, when the gun shoots, the bullet comes out of the gun at high velocity and the gun recoils. explain why the gun recoils.	PpG
Can you remember what I said?	PpG
So you've then worked out that the momentum of the bullet is 35kgm/s,	TsG
I thought you said you couldn't do any of this [*supportive tone*]	RsG
You've done the two difficult questions, it was just the wording bit	RsG
So, you've got 35 going that way, what must you have this way…?	PpG
Minus 35 [*confirms*	TsG
so then it will get you to think about the mass of the gun in comparison to the mass of the bullet and therefore how fast the gun will be going – it won't be going fast, anywhere near the 700m/s that the bullet's going.	PsG
Does that make sense?	PpG
See if you can write something for that, then have a go at that one.	RpG
(29.45-31.22)	

[to girl] tell me, how are you doing? [on the work]	RpG
That's good, well done.	TsG
Hey, look at you, that's excellent, very good.	SsG
Why do I like that?	RpG
You're one of the only people who have put a unit. Most people have just put zero.	RsG
So well done.	SsG
Bullet explodes outwards [*reading answer*] OK, yes	TsG
[*girl: is that right?*] yes, it is,	TsG
you could improve it slightly by adding in that the gun must have an equal momentum in the opposite direction,	RpG
but you'll want to talk about that later anyway so that's fine.	TpG
Erm, 700m/s, brilliant, yes.	TsG
[girl] good start, well done, keep going.	SsG RsG
Good effort, [*said altogether*]	RsG
(37.08-37.51)	

So the initial velocity of the bullet is 700 and the velocity of the gun is 2.5% of that, so you need to divide by 100 and multiply by 2.5.	PsB
It works out as about 17.5 m/s.	TsB
So if that's the speed of the gun, what must its mass be?	PpB
Because what's its momentum?	PpB
Right, and you know its velocity is 17.5…?	TsB
So therefore you can work out its mass, can't you?	PpB
By rearranging the formula	TpB
Can you do that?	RpB
[*boy: should be able to*] Yeah [*confirming*	RsB

(40.02-40.28)

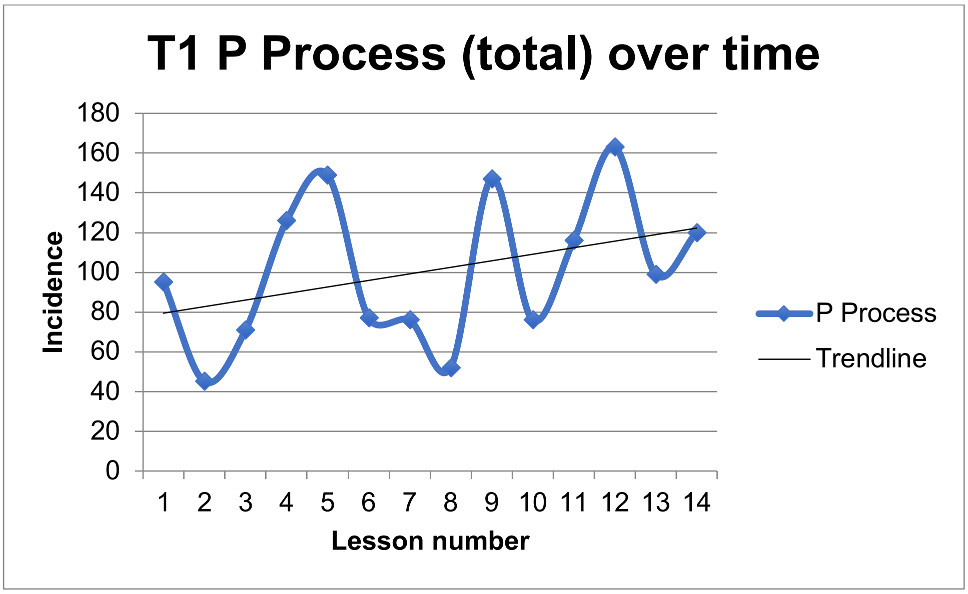
[*Charlotte*] yes it is, well done. yes, it is because what?	TsG
Yes, that's right	TsG
Yes, it's brilliant [*answer*]	TsG
You need to have some faith in yourself [*Charlotte*] because you can do this. I know you don't like to do this but you can!	RpG RsG
Why do you think you can't do it? [*Charlotte: I just don't think I can*]	RpG
Well you can, as you just proved by doing that.	RsG
Well done. Excellent	SsG

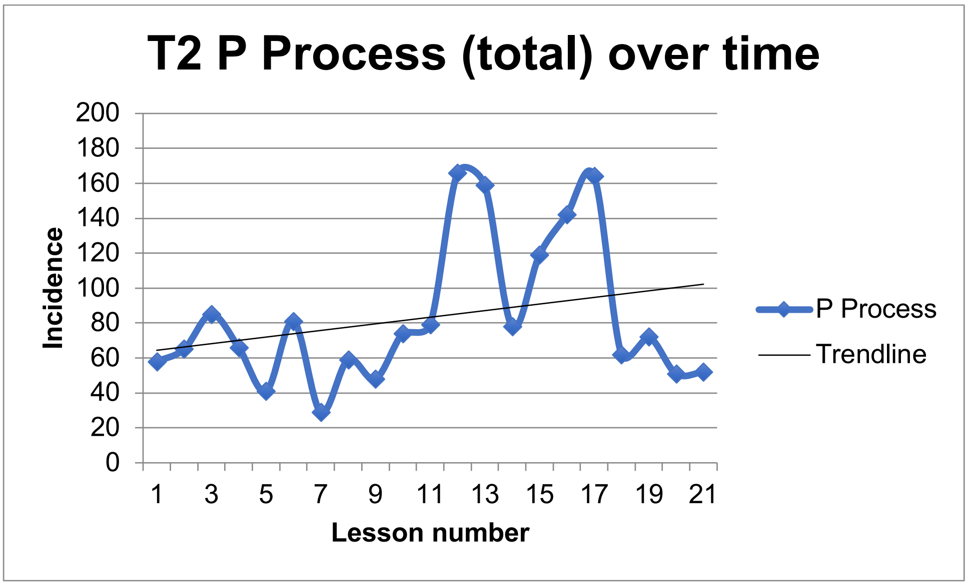
(43.55-44.17)

Both the evidence from the graph collation and the text examples provided from the numerical coded data indicate that the intervention teachers were able to increase their regulation category feedback to their students, particularly in response to on-going coaching conversations.

5.4.2 Increasing the Process (strategy) feedback

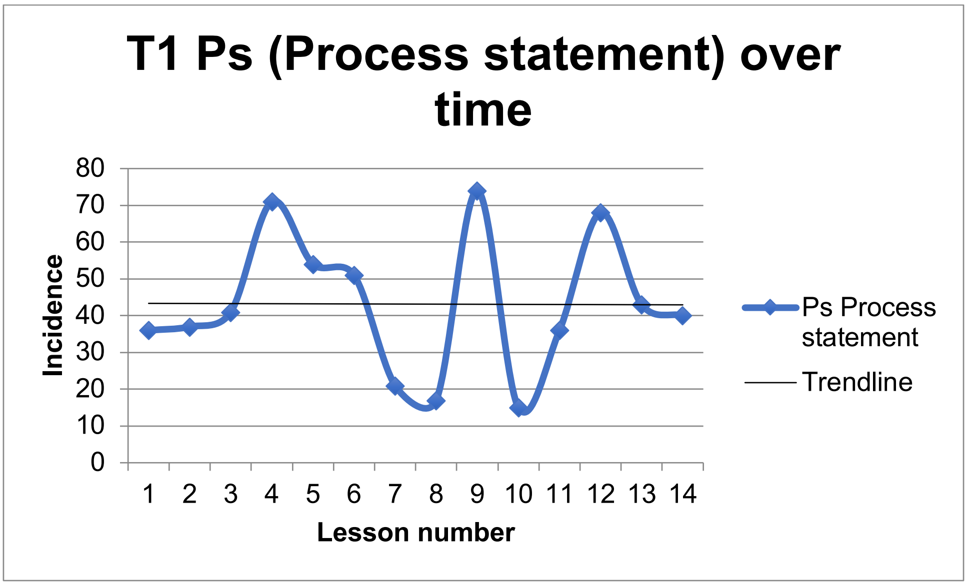
With a larger sample size, it is possible to use a trendline on the graph with greater confidence. Despite peaks and troughs in the coded observations, both intervention teachers indicated an overall increase in Process-category feedback over the duration of the study (Graph 5.13). The peaks here indicate lesson incidences of prompt feedback alone which contradicts Hattie's (1999) inaugural lecture assertion that feedback can be measured in only seconds per day.

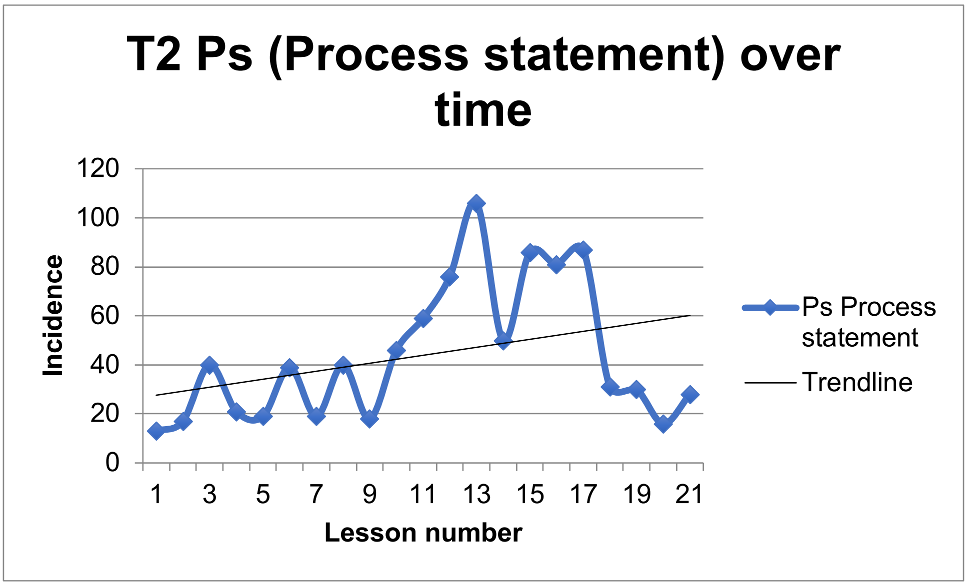




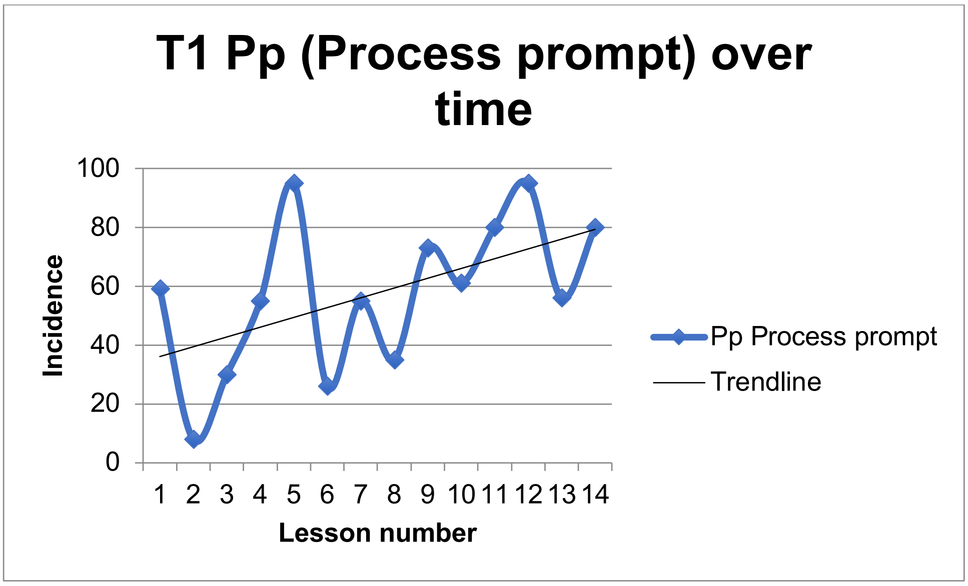
Graph 5.13 The change in total use of Process (total) feedback by the IG teachers over time.

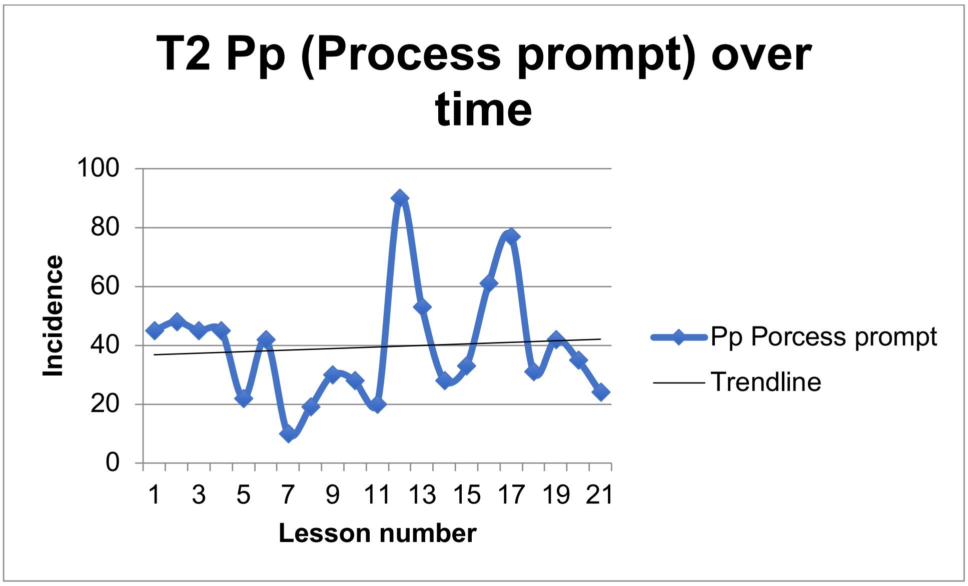
Graphs 5.14 and 5.15 map the changes in Process-statement and Process-prompt for both IG teachers. Whilst T1 shows peak-trough variation throughout, T2 appears to have struggled with Process-category feedback until at least the tenth lesson, whereupon there is an increase in both second level Process types. However, a closer analysis of T1 indicates that the incidence of Process-statement feedback was relatively stable (Graph 5.14), whereas the incidence of Process-*prompt* showed a marked increase over the duration of the study (Graph 5.15).





Graph 5.14 The change in total use of Process statement feedback by the IG teachers over time.





Graph 5.15 The change in total use of Process prompt feedback by the IG teachers over time.

Similarly, to the Regulation feedback, T1 benefited from the coaching conversations to practise flexing the feedback into other forms, and some of this is captured in Chapter 7. Increasingly, T1 was able to use Process feedback well in individual interactions with students as seen in Graphs 5.14 and 5.15; some of the exchanges above highlight how Prompt feedback appear in the dialogue sequences of lesson 13. In this lesson, there were 240 oral feedback interactions (NF), of which Task = 107 and Process = 99 (ratio =

1.1:1), and statements = 139, to prompts = 101 (ratio = 1.4:1), so there was a strong focus throughout the lesson of both Process and prompt relating to learning goals rather than a performance goal.

By lesson 10, T2 was beginning to draw upon more Process feedback, predominantly in use with individuals through monitoring progress around the class. A slightly extended interval with a girl experiencing a problem during one of the last mechanics lessons provided an example of how now Process-prompt was coming more easily in ‘interrogating’ the student:

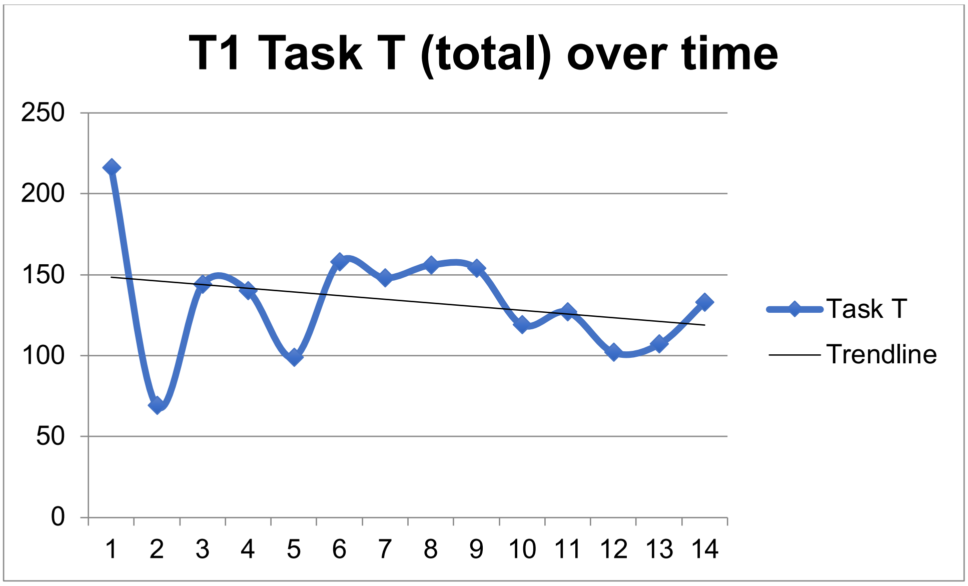
How do you get m/s2 from kilos multiplied by seconds?	PpG
[*girl explaining difficulties*] Force equals time multiplied by change in momentum? [*sic*]	PsG
Why not?	PpG
Ok, let’s think of what equations we’ve got	PpG
We’ve got momentum is equal to?	TpG
Yes, and what do I want to get?	PpG
I want to get the force.	TpG
If we’re trying to find the force exerted by the seatbelt, we’re trying to figure out what that change in momentum is	PsG
So what is that? What is the change in momentum?	PpG
How much momentum does this person have?	PpG
Don’t they? [have a momentum] so he’s got a velocity?	PpG
Ah, well we’ve got a velocity, and we’ve got a mass, don’t we, 75 kg?	PspG
So what’s the momentum?	PpG
Hmm, that’s the momentum [*confirms*]	TsG
What’s the change in momentum?	PpG
OK, work that out and it will tell you what momentum they started at.	PpG
What momentum do they end up with?	PpG
OK, so the change in momentum is just going to be from this down to zero.	PsG
Evie, you work fast; you need to think things through as well.	TsG RpG
(Lesson 10, 32.19-35.04)	

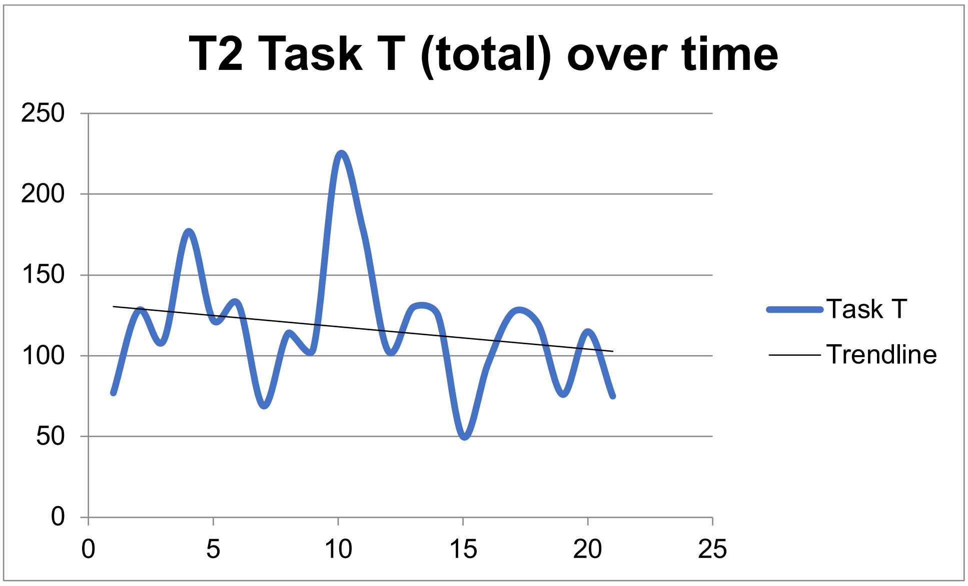
In reflection, it would appear that teachers may be more able to use this category more fluently on an individual basis, perhaps before moving to a whole-class interrogation model, which would require confidence in moving away from I-R-E questioning formats; here there is a much more fluid I-R-P-R-

P sequence developing (Sinclair and Coulthard, 1975, cited in Howe and Abedin, 2013).

5.4.3 Shifting from Task feedback

As previously described, the proportion of Task feedback decreased over the duration of the study, as the amount of Process feedback increased. For comparison, this is shown in Graph 5.16. The second-level graphs for Task





Graph 5.16 The change in total use of Task feedback by the IG teachers over time.

have not been included in this chapter, but also all show a downwards trend. The large spike in the graph of T2 at lesson 11 was the making circuits lesson in which there was equipment failure, necessitating a lot of commentary about ‘getting it right’, hence a larger proportion of Task-statement and Task-prompt, where the goal was to get a working circuit. Generally, the increase of Process-type VF corresponded with a decrease in Task-type, as the intervention teachers changed their feedback styles.

5.5 Chapter summary

This data chapter has been concerned with the findings from the classroom observations of the teachers’ deployment of oral interactions with their students in order to answer research questions 3 and 4. Evidence has been presented that demonstrates that it is possible to create a teacher profile detailing the types and proportions of feedback that they use as a signature of their teaching and dialogic style. It has been shown that these four different teachers use language in varying ways to their students, and that this has implications not only in the academic sense of leading students forwards towards both performance and learning goals, but that the way in which the teachers use their language may inform the students of what the teacher values more of these. Additionally, as suggested by Dann (2019) the language of feedback as a relational concept in the classroom has implications for the emotional space of the learning climate, and ultimately, the perception of the teacher as supportive.

The findings show that there are a large number of oral interactions made by the teacher during a lesson, contrary to suggestions by existing literature (Hattie, 1999; Bond *et al*., 2000; Voerman *et al*., 2012), and that VF forms a large part of classroom dialogue (Hargreaves, 2013, 2014; Svanes and Skagen, 2017). Teachers have dialogic characteristics that can be mapped across lessons to form a ‘teacher profile’ or repertoire (Svanes and Skagen, 2017), and that these vary between teachers and are context dependent.

Teachers vary in the amount of ‘instructional’ dialogue they give to students and oral interactions *including* teacher VF are a critical component of building relationships between teacher and students.

The intervention teachers were observed to increase their use of regulatory, process and prompt forms of VF, and decrease their use of task type, whereas the comparison teachers did not change. Additionally, the intervention teachers had lower statement-to-prompt ratios which meant they ‘closed the loop’ more often than the comparison teachers. Consequently, intervention group students received more *instrumental* information on their next steps in learning than comparison group students. However, and consistent with existing CPD literature outlined in Chapter 2, the intervention teachers experienced challenge in *flexing* their VF in the learning instance, initially falling back onto their routinised repertoires.

In summary, evidence has been presented to answer RQ3 that types and proportions of feedback used by teachers can be identified using a constructed typology, and that through a sequence of professional development, the two teachers involved in the intervention group were seen to be able to change their feedback types and proportions. This provides some evidence to answer RQ4, however, to gain a richer, and more granular picture, Chapter 7 will use the teachers’ voices to present their own learning journey as a small situated case study.

Chapter 6 Findings from the pupil interviews

6.1 Introduction and preamble

This section attempts to provide findings for RQ3 *What are the students' self-belief systems in physics as they enter Key Stage 4, and how do they change over the duration of the intervention?*

Part of this discourse analysis draws upon the use of pre-and post-intervention surveys (chapter 4) together with a deeper exploration of selected pupils' thoughts and feelings to enable a richer consideration of their individual self-beliefs over the period of the study; their 'voice' is therefore valuable. Chapter 5 identified changes that occurred in the IG teachers' feedback styles, and it was crucial to examine whether the students were able to acknowledge and recognise the feedback they had received, and the impact, if any, that they perceived it had had on their learning. This will assist forming part of the response to RQ1 *What types and proportions of feedback do teachers use in their verbal interactions with students*, and RQ2a *Has the feedback pattern changed*, and ultimately provide evidence to answer RQ5: *To what extent can data evidence an impact of RQ4 on RQ2?* [see schematic image]

Several themes were developed from an examination of the pupil focus groups, and these are presented below. Most relate to the self-belief constructs under examination and described in chapter 2, however there are additional emergent aspects, such as the pupils' notions of the personal competencies of and regard for their respective teachers. In line with RQ3, these themes will be individually indicated, and any change over the life of the intervention disclosed.

At the start of the study, one term had elapsed in their GCSE physics course; the students could be considered to have started to develop individual awareness of how they were performing in the subject. A range of questions were utilised to ascertain both student self-beliefs, and their opinion of the feedback they received in physics.

		T1 cohort						T2 cohort				
		Anna	Bridget	Charlotte	Alfie	Billy	Connor	Delia	Eloise	Dan	Edward	Felix
Pre-int	Self-concept*	1.8	0.7	2.0	2.8	1.9	2.4	1.6	2.4	2.8	2.0	1.0
	Self-efficacy*	2.1	2.0	2.3	3.1	2.6	2.7	2.2	3.0	2.8	2.8	1.8
	Anxiety	2.4	2.0	2.7	2.9	2.4	3.0	2.0	2.6	3.0	3.0	2.0
	Mindset*	3.0	2.0	2.3	2.7	3.0	2.2	3.0	3.0	4.0	2.3	2.3
Post-int	Self-concept	2.0	2.0	2.2	2.8	2.8	2.6	2.0	2.8	2.8	2.4	1.4
	Self-efficacy	2.0	2.7	2.8	3.2	2.7	2.8	2.3	3.0	3.0	2.8	2.0
	Anxiety	2.0	2.0	2.8	2.6	2.8	3.0	2.0	2.6	3.0	3.3	2.0
	Mindset	3.3	3.0	2.8	3.0	3.0	2.2	3.0	2.0	4.0	3.0	3.0
Change (qualitative)		increase	increase	increase	stable, quite high	increase	increase	increase	stable	stable, high	increase	increase

Table 6.1 The pre- and post-intervention survey scores for the student focus groups

*Items reverse-coded so that on all scores, 1=very low, 2=low, 3= high, 4=very high

Six students under T1 took part in the focus groups, comprising three boys and three girls. Five students under T2 also took part, three boys and two girls. All were aged 14-15 during the study. The teachers in the study had chosen them for the group for several reasons; through discussion with each of them and as a result of an early analysis of the pre-intervention survey, it was felt that they not only represented differing self-belief profiles, but that they were all likely to have the social confidence to contribute to interview questions. There was no experimental mortality in either focus group, and a summary of both the 'names', genders and changes in self-belief constructs is summarised in Table 6.1.

The findings will be further combined with the survey results in the discussion chapter.

6.2 Self-concept

Measures of self-concept [SC] are often based on an individual's notion of performance in a specific domain, often in comparison with peers (Marsh *et al.*, 2009). At the start of the study, it was important to gain a sense of finding out how the students felt 'they were doing' in physics, to ask in a colloquial sense. Five of the T1 students gave negative responses, exhibiting low self-concept in physics:

> Interviewer: How do you do at Physics?
> Anna: Awful.
> Bridget: Not very well.
> Anna: it's like a difficult subject because its triple, it's more intense than what I've heard core is.
> (*Connor agrees, nodding*)
> Interviewer: So, Anna, do you think it's awful because it is intense?
> Anna: it's not just because it's intense, I just find it really difficult to understand and I don't know why
> Bridget: Some of it's alright to get your head around but then there'll be a bit that's really complicated, and you can't get your head around it.

In one case, Charlotte, a girl who attains highly both in physics tests and classwork, denies that she is 'good' at physics. As seen in the detailed breakdown of the focus group students above, Charlotte has high anxiety, low self-concept, and despite her high attainment, claims extremely low self-efficacy at tasks. The group then argue about her abilities:

> Interviewer: Charlotte how do you think you are at Physics?
> Charlotte: I dunno, I just can't do it
> Anna: you can do it Charlotte, you got an A* in your last test [*the highest grade*]
> Interviewer: So, you got an A* but you think you can't do physics?
> Anna: you always say you can't do anything, but you totally always can
> Billy: how does that work??
> Interviewer: but this is really interesting…so can you, can you say, do the calculations really well?
> Charlotte: I dunno, I just can't like write, I dunno.
> Anna: what can't you do? You can, you can do most of it.
> Charlotte: I just can't remember anything.
> Bridget: she has no self-confidence whatsoever.
> Anna: she doesn't believe in herself.
> Alfie: how can you get an A*…? [*group continue to argue about Charlotte's abilities and confidence until interrupted*]

One male, Connor, unfavourably compares physics to his other science subjects, and when pressed, admits to self-efficacy in one area [calculations], but a lack of this in another area [making explanations], and this brings his self-concept lower. A second boy, Billy, expresses similar concerns:

> Interviewer: Connor, what about you, how do you think you do at Physics?
> Connor: Not very well.
> Interviewer: Why?
> Connor: I dunno, it's just difficult in some ways
> Interviewer: Which bits do you find difficult?
> Connor: Everything practically
> Interviewer: Everything? Every topic?
> Connor: well not all of them but most of them
> Interviewer: why is that? is it that you think you don't think in a physics way or you're not interested in it...?
> Connor: I don't know really, it's a bit of both I think
> Interviewer: how do you compare physics with the other sciences then?
> Connor: it's one of my worst ones I'd say anyway.
> Interviewer: but when you do things like calculations, do you think you can do those calculations?
> Connor: I'm OK at the calculations, but like explaining stuff – I can't [do]
> Interviewer: OK, thank you, Billy, what about you?
> Billy: Erm, the sort of stuff that involves maths, I think I find easier. But where you've got to explain stuff, I think is a lot harder.

Only one male expresses belief about physics in a positive way, and in enough contrast to the others to make Anna exclaim at his confidence:

> Alfie: Well generally, I can get my head round them
> Interviewer: and by *them* do you mean the maths, or the ideas…?
> Alfie: well, most of physics, yes, I can get my head around. Like I struggled a bit on convection when it was…I thought I was a bit stupid not knowing what that was, and then judging from my last scores, if you cross out the U then I think I'm pretty good at it. Not in a whatever way I explained that [*sic*]
> Anna: You're confident! [*to Alfie*]

In the pre-test condition for T2 students, all expressed feelings that physics was hard, although one male, Dan, was not as quick to evaluate it as such immediately, eventually saying it was 'more complicated' than the other sciences. Unlike the T1 group, they assign the reason for the difficulty to mathematics and use of equations in physics. One female, Delia, also expressed her belief that physics is her 'worst' subject [in attainment], exhibited low self-concept in saying that 'she just can't do it, and helplessness in how to address her own learning need, despite later acknowledging that as a Year 10 she [*they*] were expected to display more independence in their learning approach:

> Interviewer: How do you feel that you do at physics?
> Delia: It's my worst subject out of all of them
> Interviewer: Really?
> Delia: Yes, in every subject that I've chosen and every subject like maths and English, my physics is the one that sticks out like a sore thumb; on my report card, when you look at it, it's As Bs, A-stars, but physics drops to a D/C. So that just makes you feel bad
> Interviewer: OK, do you know why that is?
> Delia: probably because I just can't do it. I wouldn't know where to start to fix it. Cause I do everything that everyone else does, I just can't seem to do it.
> Interviewer: Do you find physics hard?
> Delia: Yes
> Eloise: Yes
> Felix: Yes
> Dan: [*pause*] yes
> Interviewer: OK so you guys answered yes straight away, you [*Dan*] thought a little bit and went 'yeah', what is it that's hard? Why do you think it's hard?
> Dan: It's a lot more complicated than the other sciences

> Eloise: Yeah.
> Interviewer: Can you give me an example?
> Delia: Maths, the equations.
> [*Eloise agreeing, nodding*]
> Delia: I mean, Year 10 is more independent isn't it, so if I felt that I hadn't learned what I was supposed to, we're supposed to go away and look through our revision guides.

Delia demonstrates an example-in-action of salience hierarchy (Merolla *et al.*, 2012) in which her positioning of her physics identity is lower than her other subjects. Her preference for the other subjects is offered against a backdrop of low attainment, self-concept and efficacy and her low perceived enjoyment characterised as 'that just makes you feel bad'.

In the post-intervention condition, five of the six T1 students now showed more positive physics self-concept [PSC] and acknowledges that they had improved. Charlotte, who had shown low SC and low confidence, yet high attainment, continued to claim that she was 'about the same', meaning that 'she just couldn't do it':

> Interviewer: How do you do at physics?
> Anna: Better than before in the year
> Billy: About the same
> Alfie: Really improved a lot
> Bridget: Improved
> Interviewer: So, Alfie said better, Anna said better,
> Bridget said better, Billy said about the same – Charlotte?
> Charlotte: about the same

The same question posed to T2 students yielded an observation of PSC that still seemed low, but now some of the students had shifted their reasoning away from the perceived difficulty of applied mathematics, and instead had started to verbalise what could make an adjustment in their learning, such as Edward requiring a different form of feedback to enable him to 'do better', but admitting that he too was 'doing okay'. Delia still displayed a lack of confidence at being able to attain in the subject. Eloise expressed the idea that she 'found it all quite silly, but okay'. The most positive student previously, Dan, conveyed feelings of 'doing okay', whilst maintaining that 'physics was hard':

Interviewer: So, in general, how do you feel about your physics?
Delia: I don't think I'm going to pass
Felix: Not very well
Eloise: I dunno, I find it quite difficult because the concept doesn't make sense to me [*sic*], I find it all quite silly but, it's OK.
Delia: I hate maths, so I hate physics
Edward: It doesn't really feel like I'm moving forward when he's giving the feedback like we said. But I reckon if he could change the way he puts feedback out to all of us I reckon I could do better.
Dan: I reckon I'm doing OK. It's obviously difficult, just because its physics, but I don't know, I'm kind of progressing but the feedback could be better
Interviewer: So, you said something really interesting there. You said that 'it's difficult because its physics', do you think that physics is a hard subject?
Dan: Yes
Felix: Yes
Interviewer: Felix, how do you do at physics?
Felix: Rubbish, that's all I can say
Delia: What did you get in your test?
Felix: One out of twenty. No, no, no, no three. I didn't understand the topic at all.
Edward: I feel like I'm doing ok, I'm doing quite well, but I don't know really.

Over the span of the study, T1 students showed greater gains in PSC than T2 students, since they changed from five of the six displaying low PSC to four of the six now asserting that they had improved. However, the T2 students were not quite as fixed in their original self-concept beliefs as at the start of the study, moving from one of the five stating it was 'okay' to three of the five students displaying higher PSC than previously.

6.3 Anxiety and confidence

Chapter 5 sought to show that anxiety had decreased across the IG cohort. Delia's [T2] response in the section above indicated that she worried she would not pass the examination, and analysis of both her pre- and post-intervention survey showed higher than average anxiety and low self-concept (Table 6.1). In the pre-intervention T1 student group, Anna initially spoke about how tests 'stressed her'. Yet in the post-intervention interview, spoke with equanimity in explaining her decision to prioritise another subject to physics in order to consciously avoid stress, announcing this to the apparent incredulity of the group:

> [*Anna making worried expression*]
> Interviewer: So, Anna's worrying already about this mock exam…tell me about the last test that you did with [*T1*] if you're so worried about this one.
> Anna: well, it wasn't too bad actually. I got a C, but I don't think…I probably should revise a bit more than I have done…if I think, oh it's not really worth anything, then I might not stress as much…
> Interviewer: do you get stressed?
> Anna: yes
> Interviewer: about physics or about anything?
> Anna: mainly sciences, especially chemistry…
> Interviewer: why is that do you think?
> Anna: because we haven't got the best teacher for chemistry, but for physics, it's just that its really difficult to understand all of it.
>
> [*Anna, pre-intervention*]

> Interviewer: Anna? How did you do?
> Anna: I did really well, considering I didn't revise, so I was expecting really bad, because I did more for my biology so... I thought that physics was my hardest subject, but I've got the best that I did out of all three.
> Interviewer: So, when you say that you didn't revise...?
> Anna: I really didn't revise at all.
> Interviewer: you did nothing?
> Anna: Absolutely nothing
> Interviewer: You didn't look at your book
> Anna: No, nothing
> Interviewer: was this a conscious decision?
> Anna: yeah because I thought, they don't matter
> [*group scoff and laugh*]
> Anna: OK, give me a chance! I would revise for my real mocks and my real GCSE obviously, but they [*these tests*] didn't mean anything, and it was stressing me out a bit because I had my French writing and I thought that I should revise thoroughly for that, but -
> Interviewer: So you made a conscious decision not to revise, because you were feeling stressed, or because you wanted to put your efforts somewhere else?
> Anna: I wanted to put my efforts somewhere else.
> Dan: And the French writing is part of your GCSE though…
> Anna: Yes, it is
> Interviewer: OK and looking back, are you happy with your decision?
> Anna: Yes, because I think I saved myself a bit of sleep; I heard loads of people saying they were really stressed about it and got too tired.
>
> [*Anna, post-intervention*]

This appears a remarkably mature decision to prioritise her time in this manner and rationalise her decision to peers. Table 6.1 indicates that Anna's anxiety at the start of the study was high at 2.5, signifying that she identified with a high (but not very high) state of anxiety about her physics. Post-intervention, her anxiety score was 2.0, signalling that she overall disagreed with statements about physics causing anxiety. It also relates to a salience hierarchy (Merolla *et al.,* 2012) in terms of deciding what to focus her efforts on: [*the exams*] 'didn't mean anything'. It could be postulated that this is a display of resilience in choosing to make a decision that will consciously remove her from a toxic situation.

6.4 Self-efficacy

Self-efficacy [SE], the context-dependent domain-specific belief that one may successfully execute a task (Bandura, 2006) has been shown to increase in the IG cohort, and the positive SE gain in all focus group students is shown in Table 6.1. This is exemplified in passages below, such as being able to choose and use equations, as Anna explains, or in the ability to, for example, describe something in a sequence (Charlotte):

> Interviewer: So, would it be fair to say that once you've learned to do the calculations things, then you're more confident with those?
> Connor: Yes
> Anna: yes, the calculations are normally easier.
> Charlotte: It's when you come to describe something…
> Interviewer: It's putting together explanations?
> Bridget: Yeah
> Billy: Yes
> Interviewer: So, is it explaining why, or is it putting something in a sequence?
> Charlotte: It's like describe what happens when this happens, or describe the process of this, or explain what happens when…
> [*T1 student group, pre-intervention*]

and further articulate this with the role that verbal feedback has had:

> Interviewer: And do you feel that it [*the feedback you have received*] has pushed learning on?
> Anna: Yes, because if you come to a question on a test and you've done that previously in your book and it's been marked wrong in purple pen then you know that's the wrong answer and that's the right one
> Bridget: It helps you rethink it.
> Interviewer: That's interesting Bridget, it makes you rethink it how?
> Bridget: Say if you read the question wrong or something and you wrote down the answer to the question thinking you knew what it meant, and then miss describes it to you and you wrote down the wrong thing and now you can write down the right thing and in an exam, the question you can get right is a purple pen bit.
> Interviewer: And does it help you remember that you've had that practice time with your purple pen of progress?
> Anna: Yes.
>
> [*T1 student group, pre-intervention*]

In the T2 student group, the students were focused on SE issues caused by the application of mathematics:

> Interviewer: OK thanks Delia, so is it using the equations, putting the numbers in and working it out, or is it working out which one you've got to use?
> Edward: yes, all
> Delia: yeah and memorising the right one for which equation.
> Dan: To be fair, once you've learned them, if you get the answers given to you, it's pretty simple. You get given the numbers and you just have to times them together which is pretty simple. But some of them get more complicated.

The specificity of the examples in these exchanges are typical of test items that seek to measure student SE (Pajares, 1996). It is also apparent that a student may have a high SE with respect to a particular task, but a lower SE in response to a different task. Returning to survey items such as 'I can compare and contrast conduction and convection in heat transfer' in the post-intervention state renders them susceptible to memory loss over the time elapsed since they last reviewed the topic. However, the IG cohort returned a self-possessed increase in SE overall compared with the CG cohort.

Post-intervention, the T2 students seem clear in their own minds that specific feedback on specific approaches or topics would enable them to increase their own SE; that is, effectively perform. They describe the feedback as not helpful enough, and whilst acknowledging that perhaps the goal of T2 is to make them more independent in their learning, they retreat to a position of ‘needing to understand’ their physics more:

> Interviewer: What would be helpful for you?
> Felix: Well I think he needs to go a bit more in-depth what he says, because it’s like simple, he just highlights the bit that’s wrong, like maybe he gives a lesson where we can speak to him about what we’ve done wrong, and then we can improve for the future.
> Interviewer: So, what do you feel you get tripped up on then?
> Felix: it’s just little things, equations and symbols, and stuff that I get the wrong way around.
> Interviewer: And that’s the sort of thing that once you’ve cracked it, you’re confident?
> Dan: Maybe more helpful feedback to tell you what to do.
> Interviewer: So, when you say helpful, what feedback are you wanting to help you?
> Edward: Tell us what to do to actually fix the problem, not just circle it and tell us that there’s something wrong here, work it out yourself.
> Interviewer: OK. So maybe what he’s doing is trying to grow you as more independent learners?
> Eloise: [*pausing*] Yeah.
> Alfie: Yeah.
> Delia: Yes, I get that but I kind of need to understand a bit more to be able to do that. It’s maybe the way he’s done it, it’s quite hard to understand. Like, letters on the board to describe all the x rays and gamma rays and he’ll put [*draws in the air*], or a ‘p’ [ρ] for momentum, and I’m like, why can’t you help us out and put M? And when he writes in our book, he uses those symbols, and we have no idea, what’s ρ? I have no idea what that is.
>
> [*T2 student group, post-intervention*]

As shown in Chapter 7, T2 was indeed motivated to develop his students as independent and resilient learners, however communicating this intent to the students and/or the students internalising this intention did not at this point seem to have been successful. For example, Delia shows that she cannot access the language of physics and finds it incomprehensible, and thus frustrating. T2 strove to get the students to think more for themselves, and as Table 5.4 shows, T2 had the highest amount of Process/strategy-related feedback of all the teachers at 30.0%. However, the strategy may indeed have been slowly working on the students:

> Edward: If he described what it was that he wanted us to do when we're stuck. say if I hadn't done a question for a few weeks because I don't understand it, he could put, come and see me, or something like that, or come across to me, something like that.
> Interviewer: So, are you wanting him to spot something that you keep tripping up on?
> Edward: Yes, so that he could cover it with me properly.
> Interviewer: Do you know at the time that you're tripping up on it, or do you think that you're doing it OK?
> Edward: No, it's just that over time I've realised that I've done things wrong, and I've just done the same thing over and over again, and he could help me with that.
> Interviewer: OK. So, do you have to work out yourself sometimes what it is that you've done wrong?
> Edward: Yes, he'll say, 'improve this' or 'sort this out' or whatever.

Here, Edward is admitting that he has worked something out for himself, as was T2's intention, recognising that this is deeper learning than simply telling the student. The student, however, feels slightly resentful that his teacher is not recognising when he is struggling and coming to his rescue by 'telling him', and here we see in action the difference between *instrumental* and *executive* help (Hattie and Timperley, 2007). T2 understands that self-efficacy is not something that is 'gifted' to students, but that it must be earned oneself, and is trying to promote student agency (Bandura, 1999).

Both the students' wish for 'better, more focused' feedback and the difficulties accessing physics nomenclature indicate that SE in certain topic areas was problematic for these students, and the group appeared quite negatively focused at this part of the discussion [*field notes*], despite the survey results indicating a general trend of increasing SE over time.

There also seems to be miscommunication between T2 and the students regarding the purpose, nature and quantity of feedback given by T2, and perceived and enacted by the students, which is further discussed in Chapter 7. T2 believes that the VF deployed is useful for building the students as agency-driven learners (Bandura, 1997, 1999, 2018). In this instance, T2 students indicate that they are not experiencing this agentic action, particularly with respect to the fourth core property of agency: self-reactiveness, including processes of self-management and self-motivation, the lack of which can undermine self-regulation, the learning ideal. Here, the students are showing that they want to be *told*, and this indicates that the *discrepancy* and Process feedback is not being received.

6.5 Mindset

One of the key self-belief themes investigated in this study is the contested notion of ‘mindset’ (Dweck, 2006, 2013; Brown, 2017; Li and Bates, 2017; EEF, 2019), and the apparent ‘fixed vs growth’ dichotomy with which students may approach both tasks and feedback received. Chapter 2 outlined Dweck’s (2013) exploration of students’ concept of intelligence, as incremental theorists or entity theorists. Both schools involved were ‘Growth Mindset schools’; that is, the schools had adopted a policy of promoting growth mindset, which may have had an impact on how the students answered the mindset items on the survey. Reviewing the pre- and post-survey results in Chapter 4 indicated that there was little movement on the mindset items, but that generally, the IG cohort tended to a slightly more growth type.

Zander *et al.*'s (2018) study of 580 (Higher Education) students found that academic SE was positively related to growth mindsets, using validated mindset questions drawn from Dweck, 1999, as did this study. The student voice presented below has some emergent aspects of SE arising, as well as features of motivation, self-agency, and in the post-intervention condition, an emergent maturity in recognising that whilst individuals may have different strengths and weaknesses, a willingness to ‘work hard to get it’ indicates increasing self-determination.

At the start of the study, as T1 students discussed the outcome of a recent test, Alfie (very high IT, high SC and SE, low anxiety) exhibited annoyance at not beating his last score, whereas Billy (low SC, higher anxiety) displayed ET beliefs about physics; that one had to be born to be good at physics. In the same extract, two of the girls expressed a more IT belief that it could be learned, even though one of these, Bridget, had indicated a pre-intervention state of low SC and SE, high anxiety and a mindset value of 2.0 which put her on the ET scale:

> Alfie: …for me I got good grades on the past [*previous*] papers, so I wasn't very happy about that [*his last score*].
> Interviewer: Do you always want to beat your last score?
> Alfie: Yes, I do
>
> Interviewer: So, do you think you have to be born good at physics then?
> Billy: Yeah
> Anna: I think you can learn it. I think that on specific topics you need to be able to get your head around it.
> Bridget: Yeah [*agreeing with Anna*]

6.5.1. On 'getting good grades'

At the end of the study, the students had again undergone a recent test. Bridget, as above, exhibited a lower estimation of her abilities than the reality warranted, whereas Alfie displayed irritation again with his mark, this time not because it did not exceed his expectation, but that it had narrowly missed the upper grade boundary:

> Interviewer: You've recently had a physics exam; how did you do?
> Bridget: Alright. I was expecting to get a U, but I got a C, so I was impressed with myself.
> Alfie: I was quite happy with my grade because I predicted about a C, well, a high C, but I was one [*mark]* off an A, which then annoyed me.

When asked their opinion about why they get good grades, T1 students were emphatic about their ownership of the result, and there was no suggestion that T1 was responsible for any perceived failure, and that the result was the consequence of the effort they had expended:

> Alfie: When I get good grades, that's when I've revised, I've put plenty of effort into it, but in my… when my bad grades come in, it's the opposite.
>
> Billy: Well, when I get good grades its mainly because I've revised more but then when I do badly, it's down to that I haven't done any revision.

Anna and Bridget both explained that they felt that a change in their attitude to physics learning had resulted in good grades:

Interviewer: Anna?
Anna: Hmm, I think because I've taken my time with it more. I've stopped trying to catch up with Zara, and I start to understand it more if I just do it my own way.
Interviewer: OK, tell me about catching up with Zara?
Anna: Well, Zara's really fast and she just knows all of it, and it's hard to do quickly, if I don't understand it, I can't just brush it off, so I've started to do it my own way.

Bridget: I feel good when I get good grades because it means I've improved, and I can understand the things in the test better.
Interviewer: is it always about the tests then Bridget?
Bridget: well no, even if we're just doing questions in class and I get those right, like on my own, without having to use Billy, I feel good, like, I don't know, I feel more independent, more able to do it.
Interviewer: OK. And do you think that's changed over this year?
Bridget: Yes, because I used to always ask Billy, or just copy from him, but now I can do it
Interviewer: what do you feel has changed this year that makes you feel you can do it more independently?
Bridget: I don't know. I do feel able to ask for help more, and she's given me more helpful feedback that I can use.

In both of these interactions, Anna and Bridget are demonstrating an increase in the salience of physics in their hierarchy (Merolla *et al.*, 2012); Anna in not comparing herself negatively with the friend she perceives to be more able [faster at physics] than herself. In a similar fashion, Bridget feels that she is less reliant on Billy, upon whom she would previously rely (and copy) and endorsed herself in being able to ask for help and acknowledging the efficacy of the feedback that T1 has provided.

In examining their test success in physics, T2 gave a range of answers from understanding [*that section*], working harder, not being as confused as at other times, but also crediting T2 with ensuring that the class re-focus on topics that had been poorly attempted, contradicting earlier comments that the feedback could have been better:

Interviewer: When you get good grades in physics, why is that?
Edward: I've understood that section
Eloise: I've worked a bit harder maybe…recently I've been paying a bit more attention because if I don't get it it's my responsibility as well. I want to understand it as best as I can and then I've done as well as I

> could have done. The reason I've been doing that a bit more, so that when I get good marks, then I've tried for that and it went well.
> Felix: Well some topics don't confuse me as much as some others, so I do better, but I'm generally bad at exams.
> Dan: Maybe again, just work better on certain subjects, but when I go back over things…I don't know, if we all do badly on a topic, then we spend more time on it instead of going onto other topics, so he focuses on making sure that we've actually understood it.

During the duration of the study, there appears to be a move towards more student ownership of test success, from both IG cohorts, and notably from T2 students who had previously directed their concerns about their physics attainment and understanding to a perceived lack of action from T2.

6.5.2. On 'doing better'

Post-intervention, when T1 students were questioned as to whether they could have achieved higher in the recent physics test, all students but Charlotte [high attaining but low PSC] answered 'yes'; she subsequently conceded that she could both study more, and also study in a more effective way:

> Interviewer: So, if I ask you now as a group, could you have done better [in the exam]?
> Anna: Yes
> Charlotte: No
> Bridget: Yes!
> Dan: Yes
> Interviewer: So, Bridget you said yes quite emphatically?
> Bridget: Yes, I could have revised more and done a lot better.
> Interviewer: Did you not revise then? At all?
> [*Bridget shakes her head*]
> Interviewer: OK, I'm going to read you some options. if you think you could have done better, was it down to one or more of these?
> The test was unfair.
> I didn't go about studying in the right way.
> I didn't study hard enough.
> I wasn't smart enough.
> Do you want me to say them again? [*repeats them*]
> Alfie: Second and third [*I didn't go about studying in the right way. I didn't study hard enough*]
> Billy: Second one [*I didn't go about studying in the right way*]

Bridget: Second one [*I didn't go about studying in the right way*]
Anna: I'd say third out of all of them
Interviewer: You didn't study hard enough?
Anna: Yes
Dan: You didn't study at all!
Anna: Well exactly!
Charlotte: Second and third [*I didn't go about studying in the right way. I didn't study hard enough*]

In each of these exchanges, the personal agency is declared. No student felt that the test was unfair, or that they were not smart enough. However, for T2 students, who showed a wide variety of marks over their last test, one boy did assert that he wasn't smart enough; Felix, who exhibited very low SC and SE, low anxiety and inclinations towards an entity intelligence theory ('fixed'). Other students acknowledged that they hadn't studied hard enough, however Delia articulated a different view in specifying that her study strategies were not effective, but that she didn't know what to do to change this:

Interviewer: so if I ask you what were the reasons you got what you got, if I give you different choices, if you felt you didn't get as much as you would have liked, which of these were the reasons? The test was unfair, you weren't smart enough, you didn't study hard enough, or you didn't go about studying in the right way (and that's different to not studying hard enough)
Delia: Last one, I didn't study in the right way
Felix: I'm not smart enough
Eloise: I don't think I studied
Interviewer: You didn't study hard enough?
Eloise: Yeah.
Edward: I didn't study hard enough.
Dan: I didn't study hard enough.
Interviewer: So, Delia, when you said you didn't go about studying in the right way, what do you think you could do to change that?
Delia: I don't know.
Interviewer: What did you feel was the wrong way?
Delia: Well, not doing it properly like…when I revise for other subjects I copy multiple different things and do past papers, and when I do physics revision I just read over my book and hope for the best, whereas with the others I put a bit more effort into it.
Interviewer: More active strategies?
Delia: Yeah, that would do me a lot better, particularly something I struggle with.

All T2 students but Felix felt that the agentic action was their responsibility, whereas Felix blamed his perceived lack of intelligence.

6.5.3. Learning from the test

In exploring how the students might have thought and learned about preparing for, and taking the physics test, the interviewer gave a range of options and asked them to identify with the option(s) they felt most mirrored their future approach:

> Interviewer: And thinking about going forward, again if I give you some options: I would spend more time studying for the test, I would work harder in this class from now on, I'd spend less time on this subject from now on, I would try never to take this subject again.
> Bridget: First and second, yes, those two [*I would spend more time studying for the test or I would work harder in this class from now on*]
> Interviewer: Charlotte? [*pause*] shall I read them again? [*repeats*]
> Charlotte: the first and the second [*I would spend more time studying for the test* and *I would work harder in this class from now on*]
> Interviewer: The first and the second, thanks Charlotte. Billy?
> Billy: The first one
> Interviewer: Alfie?
> Alfie: The first one
> Interviewer: Anna?
> Anna: The second one.
> Interviewer: You'd work harder in the class? OK. so, do you think that putting the effort into the day to day class stuff would give you a better grounding than revising?
> Anna: Yes, because I think that you're not kind of set in your knowledge more, so when you're revising, you're probably revising a particular thing, but if you work harder in class you're probably going to learn more than when you're trying to revise something.
>
> [*T1 cohort, post-intervention*]

Here, Anna is making a distinction between personal study away from school and expending more effort in the physics classroom as she sees the latter as more of an opportunity to *learn* rather than 'going over' material already covered. The two other girls choose to both study more for the test and work harder in the class without making this distinction. Anna seems to make a more thoughtful decision about what conditions will grow her as a learner, and intimates that the teacher, or at least the classroom presence will do that. Alfie and Billy indicate only that they would spend more time studying for the test, and in the absence of a discussion of study strategy, it is not a certainty that this is the best learning option.

When asked the same question, nearly all the T2 students chose the option of working harder in class. This intention to agentic action contrasts with the students' earlier assertions in section 6.4, in which they appear to want T2 to give more feedback, as well as *tell* them [the answers] when they don't understand. This might indicate that in a post-intervention state, they are beginning to have more of an understanding of their own role in their learning; equally, this may just be sign of maturing. As previously however, Felix indicated that he would rather step away from the subject than work at it:

> Interviewer: If you think about the outcomes from your physics test, I'm going to give you several options again. The options are: I'd spend less time on this subject from now on, I would work harder in this class from now on, I would spend more time studying for the test, I would try never to take this subject again. You can choose more than one. Do you want me to say them again? I'd spend less time on this subject from now on…no-one? I would work harder in this class from now on
> Eloise: Yes, I think I would do that
> Edward: Yes [*agreeing with Eloise*].
> Dan: Work at it.
> Interviewer: You'd work at it. Felix?
> Felix: I wouldn't want to do it.
> Interviewer: You wouldn't want to do it? So, you'd spend less time on it?
> Felix: Yeah.
> Interviewer: And/or you'd try never to take the subject again?
> Felix: I'd try never to take the subject again.
> Edward: I would work harder in class.
> Delia: I'd probably work harder at it, though I probably wouldn't want to. I mean, I know I should, but I probably wouldn't want to.
> Interviewer: So, does it feel like hard work because you don't want to?
> Delia: Well when it comes up to my GCSEs, I'll put so much effort into it cos obviously I want to get it, but, I don't know, it's putting my head down to do it
>
> [*T2 cohort, post-intervention*].

With the exception of Felix who states an intention of disengaging from physics, the students appear to be motivated by general GCSE success rather than stating liking for physics, or a need to improve in physics *per se*.

6.5.4. On hard thinking and working

What do the students think of the 'working harder makes me smarter' mantra adopted by schools utilising a growth mindset approach? As discussed in Chapter 2, Dweck (2017) asserts that praise (or feedback) solely on effort is wrong, and that the feedback should instead focus on how the effort has created learning progress. However, students bring their own complex self-beliefs to learning situations and do not always recognise (or want to acknowledge) that using higher-order thinking skills is beneficial to their learning if they are struggling, lacking motivation or through low self-concept, are convinced that they are 'just not good' at the subject. The interviewer posed the question 'I like schoolwork best when it makes me think hard':

> Billy: No!
> Bridget: Well in some ways…
> Alfie: Well if you finish it off and do it right, then yes
> Billy: it depends to what extent how hard you have to think…if you can't do it then it doesn't really make it as fun, as fun as if you succeed.
> Interviewer: OK, do you think you've got to be naturally good at something to work hard at it?
> Billy: Not necessarily. If you've just got the interest in it then…
> Interviewer: do you think you've got to have the interest to make yourself work hard at it?
> Billy: Yeah, but also the determination
> Interviewer: Do things come easier when you work hard at it?
> Billy: Yeah
> Interviewer: Anna, what about thinking hard?
> Anna: Sometimes…I think it's quite nice if you do get the hard stuff right. but when its easy you get to do a lot more.
> Interviewer: So, which is it more satisfying to do, hard stuff or easy stuff?
> Anna: hard, but I don't like it.
> Interviewer: OK that's fine. Bridget?
> Bridget: I don't like it.
>
> [*T1 cohort, post-intervention*]

Billy's immediate reaction is to reject the idea that schoolwork is best when it makes him think hard, however he then moderates this with 'it depends *how* hard you have to think', and adds a qualifier that if you can't do it, then it's not as fun as if you can. This is based in aspects of both self-concept and self-efficacy and shows some personal movement in self-belief for Billy from the start of the study. In section 6.5, he states [pre-intervention] how he thinks you have to be born good at physics; here he is prepared to concede that one doesn't have to be naturally good at something to work hard and have interest in it, but that determination is then necessary. He does agree that 'things come easier when you work hard at it'. Both Anna and (to a less stated extent) Bridget see the value in thinking hard, but don't *like* to do it.

The same question was asked of the T2 students:

> Interviewer: How about this: 'I like schoolwork best when it makes me think hard'.
> Eloise: No!
> Edward: Again, it depends on the subject
> Interviewer: OK Eloise, your first reaction was no, you don't like to think hard?
> Eloise: No not hard, I like it when it comes naturally, when you just know what you're doing. I don't mean this in a weird way, but I've never had to think properly hard. I don't know, what do you mean by hard?
> Interviewer: It's like when you're puzzled by something, and you've got to really think about it to work it out.
> Felix: When you think about something too much it gets you bored
> Interviewer: Edward, you said it depends on the subject
> Edward: If it's something I don't enjoy then I won't put as much effort into thinking about it. If it's something like DT or drama, then I think quite hard about it because I want to get a good grade in that. I'm more determined to get a good grade in that than a lesson that bores me, or I just don't understand.
> Dan: I'm the same really. I'll work hard in subjects like computing, but things I don't really care about, like English, which I've despised for ages, I won't work in that, not hard.
> Interviewer: OK. Felix? Do you like thinking hard?
> Felix: No, it bores me
>
> [*T2 cohort, post-intervention*]

On review, this seemed a surprising comment from Eloise who hitherto had provided responses which indicated an inclination to appreciate what working hard would yield, as shown in section 6.5.1above. Although not directly explored with Eloise at the time, it raises a possibility that Eloise might distinguish 'working hard' from 'thinking hard', for example, in terms of increasing the *time* spent on a subject, but not increasing the depth of mental effort expended on it. She does not seem to have clarity about differing cognitive demand. Felix makes his feelings about thinking hard very clear, and this supports his assertion in 6.5.3 that he wouldn't take subjects he found difficult further. Edward and Dan both indicate that they are operating a subject salience hierarchy (Merolla *et al*., 2012), and this was pursued by the interviewer:

> Interviewer: OK, so do you have a pecking order of subjects that you prefer?
> [*all boys agree*]
> Interviewer: Do you work harder in the subjects that you find easiest?
> [*all agree*] Yes
> Edward: And more enjoyable.
> Dan: Yes, and more enjoyable.
> [*all agree with this too, Felix gives example of computing and DT*]
> Interviewer: Do you work hard in those Felix?
> Felix: Yes, I do
> Interviewer: because you like them more?
> Felix: Yes [*emphatic*]
> Interviewer: So, are all of you saying that you work harder in subjects that you not only enjoy more, but you feel you're better at?
> [*all agree*]
> Interviewer: Why do you think you're better at them?
> Because we work harder
> Because we enjoy them [*split answer, all saying these together*]
> [*All pause, thinking*]
> Interviewer: It's a vicious circle isn't it?

In this extract, Felix indicates that he is prepared to work at subjects that are higher in his salience hierarchy, reasoning that he is both better at them, and enjoys them more. It would appear from these exchanges as well as his survey responses that physics is much lower down this order, and that to an extent he is disengaging from it, reasoning his efforts should be put into more enjoyable and potentially higher grade-yielding subjects. Dan and Edward display similar thinking, but not to the same disengagement extent as Felix, and their SC and SE are higher, and their physics anxiety lower. Dan had a very high incremental mindset score (4.0) and Edward a high score at 3.0. Felix changed slightly through the survey from an entity theory score of 2.3 to 3.0 (incremental), however his physics SC and SE were low, though improved from the start of the study. Despite his comments here, he conceded responses to the anxiety items in the survey indicate that he was indeed worried about getting a poor grade in physics.

Dweck (2010, 2017) reported that a common response from ET students was that they did not like to work hard, because they felt that it should 'come to them naturally'. As highlighted above, working hard and thinking hard may have been conflated, yet may mean different things to the students and in the absence of student definitions of these, it was necessary to question further:

> Interviewer: When you work hard at your schoolwork, does it make you feel like you're not very smart?
> Alfie: Well sometimes, if everyone else is getting it and then you're just there like, how do I do this?
> Bridget: Yeah.
> Interviewer: Anyone else?
> Connor: Yeah, the same
> Anna: The same. If other people are getting it and you don't
> Interviewer: So, it's by comparing yourself to someone else?
> Anna: Yes. You're not just comparing your knowledge, because some things you can do that other people can't do, it's a bit down heartening.
>
> [*T1 cohort, post-intervention*]

In Alfie's reference to 'getting it', he appears to be referencing cognitive demand, and negatively compares himself in wondering how to address a task that others seem to be attending to with ease. Anna recognises that peers are different but concurs with others in the group using the same language in agreeing with Alfie. The interpretation here is indeed that working hard has been interpreted in this conversation as 'thinking hard', and not spending more time on it; the time implication above is more in reference to the time the students feel they are wasting in not being able to perform a task. T2 students conveyed similar thoughts:

> Interviewer: OK. Next one, to tell the truth when I have to work hard at my schoolwork, it makes me feel like I'm not very smart.
> Dan: Yeah
> Felix: you're putting more effort into it, so you feel more dumb that you're putting more effort in.
> Edward: there's other people who can do it in a minute, and then there's me that's sat there for like twenty minutes trying to work it out
> Interviewer: So, it bothers you because you compare yourself to someone who has done it quicker than you
> Edward: Yeah
> Dan: The same really, I don't want to get paranoid because someone else has finished, I'm just 'well get it done quickly' and move on.
> Felix: You feel rushed
> Interviewer: OK, Delia, when I have work hard at my schoolwork, it makes me feel like I'm not very smart?
> Delia: I agree with what they've all said.

> Eloise: I think I disagree really. When you work hard, and then you finally get, then you are smart. I think some people get things quickly, some topics…because I know that if I find sometimes find it harder, and someone gets it really quickly, then in another thing, I get it really quickly. So, I don't really mind not being quick at things at times.

Felix gives a response that is typical of Dweck's (2010) entity theory examples and explains further that he feels rushed in addressing the task. Like T1 students, most are concerned at the comparison to peers in the time it takes them to complete, or even start, a task. They are describing episodes of a lack of self-reactiveness (Bandura, 2009); a 'semi-fugue' state of momentary absence of self-regulation and self-efficacy, providing an emotional state which causes them to compare themselves negatively with their peers (lower PSC). In contradicting her friends, Eloise is a lone brave voice in recognising and articulating that there are differences in the way individuals approach or complete tasks and shows genuine appreciation of the hard thinking that ultimately leads her to success and self-congratulation.

In rephrasing the question to a self-concept context, the students' views of whether effort could make a difference was explored:

> Interviewer: If you're not good at a subject, working hard won't make you good at it?
> Delia: I disagree with that
> Edward: I disagree, if you're not good at it, putting effort in will make you better
> Felix: I think eventually you'll get good at it; you just have to go through the boring bits to get there.
> Dan: I think it will make you improve, yeah, I think that if you work hard you will get better eventually at the subject.
> Eloise: I think that if you work hard then obviously you'll improve, but if it's not something that you want to do, like a lot, then work hard to get a good grade but don't spend all your time fixating on it, cause if it's not your thing, it's not your thing.

All T2 students disagreed with the statement that working hard wouldn't make a difference in a subject one was 'not good at'; these responses are phrased in terms of time, effort and cognitive expenditure.

6.5.5. Learning from mistakes

Dweck (2017) claims that part of establishing a growth mindset culture is to celebrate making mistakes, teaching students that that is how the brain learns. This contrasts sharply with student practice observed in the classroom where in both IG cohorts, students were, for example, reluctant to commit pen to paper until they knew what they were writing was right [field notes]. This tended to occur more with the females than males, however, Billy and Edward were observed to check before commencing to write, and all the girls in the study did on occasion. This prompted the interviewer to question them about making mistakes:

> Interviewer: OK next one, I like schoolwork I learn from, even if I make a lot of mistakes.
> Eloise: I get irritated when I make mistakes over and over again -
> Delia: Yeah
> Eloise: - like when you're trying really hard to get it -
> Delia: like in maths, when you try and do something, and 'oh, you're nearly there' and you go away and do it, and you go back and they're like 'oh you're still not there, well I give up then.
> Felix: no…it annoys you, like you spend another half hour redoing it and its wrong again, it's a waste of your time
> Dan: I do this thing…they say just move on if you can't do it, and I don't because I'm so set on doing it right, but if I don't get it right it's annoying, because again, you've wasted your time.
> Edward: just the same as Dan, if I keep on making the same mistake when I'm doing some work, it will really irritate me, even if it's the smallest mistake, it will still irritate me, and I won't be able to do it properly, so no.
> Interviewer: So, it winds you up that you can't do it,
> Edward: yeah
> Interviewer: Does it make you want to beat it, or does it make you want to move on from it?
> Edward: It makes me want to throw it in the bin to be honest.

These T2 students all find making mistakes annoying and time-wasting, regarding them as counter-productive to learning. Again, this is in direct contrast to the Mindset author's recommendation that mistakes are used to push the boundaries of learning. This would seem to indicate that modelling of learning from mistakes did not take place in T2's teaching, yet the lesson observations contain many examples of T2 using language to grow answers beyond incorrect responses and explaining that mistakes help learning. Instead, as Dan says above, they are ‘set on doing it right’, which in part could be caused by the current examination-driven performance goals school culture, naturally at odds with growth mindset messages.

6.6 Student acknowledgement of feedback

Chapter 2 sought to provide topographical features of the feedback landscape, and highlighted how the feedback deployed by the teacher is not always the same as what has been received and internalised by the students, as well as considering the role that emotion plays in receiving, acknowledging and accepting the feedback as part of the learning process. Different sources have outlined the vulnerability that students may experience upon reading or hearing the feedback they have received and that its received meaning and/or efficacy could be some distance from the educator’s intention.

In exploring the potential link that the feedback intervention may have had on the self-beliefs of the IG students, it was necessary to establish whether the students themselves recognised feedback taking place, and also to enquire from the student perspective whether they felt that the teachers’ feedback had an impact on them. Student notions of feedback are unlikely to exist in a vacuum; they will have long been the recipient of feedback from many teachers, and will have formed their own ideas about the relative merits of the pedagogical, moral and relational economies of feedback outlined by Elbra-Ramsay (2019).

In a pre-test focus group, the interviewer sought details of the feedback students received at that time:

> Interviewer: do you get feedback on your physics work?
> Edward: Yes
> Dan: Yes
> Interviewer: OK can you tell me about that feedback Dan?
> Dan: it’s just how we can improve and what we did well, and then we have to write down…do what he says, the improvements we should do.

> Interviewer: OK, Dan, how does it make you feel, your feedback about your physics?
> Dan: Not too bad, erm I mean it could be better than just red pen, a bit more interactive. It could just be a bit earlier if we've done poorly on certain subjects, recap it a bit, I mean, learn the things we did badly on, and then if we learn them again we might actually know how to do it.
> [*T2 student group, pre-intervention*]

Both sets of IG students confirmed that their teachers gave them feedback, and that it was focused on commentary about what had been done, and student pointers about improvements that could be made. Although in different schools, both teachers used a model of WWW/EBI (What Went Well and Even Better If…), which the students understood. In the post-intervention state, both student groups gave a rapid and definite acknowledgement of feedback:

> Interviewer: do you get feedback on your physics work?
> All: Yes
> Interviewer: so that's an emphatic yes from all of you. What sort of feedback does your teacher give you about your physics work?
> Anna: written and verbal
> [*T1 student group, post-intervention*]

> Interviewer: do you ever get feedback on your physics work?
> [*all say*]: yes
> Interviewer: and what sort of feedback is that? what sort of feedback are you talking about can you give me an example?
> Felix: how to improve our work and stuff
> [*T2 student group, post-intervention*]

6.6.1 Verbal feedback

The classroom observations were focused on the oral interactions of the teachers to the students, and the opinion of the students on this was crucial. T1 students confirmed that the VF given by T1 was both very likely to occur in the learning instance, was tailored to the learning needs of the student, and as Kerr (2017) reported, was a 'form of focused conversation':

> Interviewer: Yes, it is, okay thank you. so, when she gives you this feedback for your work, what sort of feedback does she give you?
> Alfie: Verbal feedback
> Anna: Yes, I like the fact that she does it as a class, but then personally talks to you as well; it's not just a general thing, it's about you and what you need to do to improve.
> Interviewer: So, do you feel – and I don't want to be putting words in your mouth – do you feel she tailors it for what you need?
> Anna: Yeah, yeah, definitely
> Interviewer: OK can you give me an example of that?
> Connor: so, like, on the test, you'd probably need improving on something different that I would (*to Alfie, who agrees*), and you'd probably be good at something that I'm not…like if she was talking to the whole class; on the whole you need to work on this, but if she came to me, she'd be like 'you need to work on heat particles [*sic – a class in-joke directed at Alfie, group laughs*], or convection
> Interviewer: so, she gives you verbal feedback – a lot? enough? too much?
> Billy: Enough. I don't think she gives us too much.
> Anna: She gives us enough to take in. [*T1 student group, pre-intervention*]

In contrast, the students of T2 did not appear to recognise the VF that they received:

> Interviewer: Does she give you feedback in class?
> Delia: No.
> Eloise: not really.
> Interviewer: OK.
> Delia: We don't even get that much in our books, do we?

From these two comments, the group began to focus instead on written feedback in their books. Pupils at both schools revealed at pre-intervention that there was a policy of marking work after which the pupils were given class time (sometimes known in the UK as Directed Interactive Reflective Time [DIRT]) during which they would make corrections or additions in a different colour pen.

When questioned again post-intervention, the students of T1 are very clear about the VF that is used and can give examples of it. The focus here from the students is one of improvement; 'more perfect', 'make it better', and Alfie's response shows unequivocally that he considers the feedback to be learning-related. It is also evident that he is capable of distinguishing between the actions required arising from the feedback; straightforward corrections and 'ones at the highest level'; by this Alfie is referring to the 'six-mark' examination questions that require higher order thinking skills to create an answer to, such as sequencing or evaluating:

> Interviewer: What sort of verbal feedback does she give you?
> Anna: She tells us how to improve, how to make the answer more perfect.
> Interviewer: More perfect, I like it. What do you think Billy?
> Billy: Yes, it's OK.
> Interviewer: OK feedback, or what you do with it?
> Billy: both really.
> Interviewer: so, when you get the feedback, what do you do with it?
> Anna: We purple pen it.
> Interviewer: So, when you have your purple pen, what do you do with your feedback?
> Alfie: Well normally for the questions that we get wrong, not the ones at the highest standard, we normally just correct them. Sometimes we get homework set to do the corrections at home,
> Interviewer: Is it always about corrections?
> Bridget: Not always.
> Alfie: It's about learning from it, that's the main thing really.
> Interviewer: Bridget, what were you going to say?
> Bridget: it might be about adding to your answer to make it better.

When asked the same questions, the students of T2 are again dismissive of the feedback received. Initially, Dan hesitates to evaluate in this way and then joins his peers in agreeing that 'not much' feedback is given:

> Interviewer: OK, what verbal feedback does [T2] give you?
> Eloise: not much.
> Delia: not much.
> Felix: don't think so.
> Interviewer: OK, do you all agree?
> Eloise: Hm Hm [*agreeing*]
> Dan: [*pausing*] Yeah.

At the start of the study, it was unclear how much VF the students of T2 received, however upon commencing classroom observations and over the time period of the study, T2 was found to give a mean number of 209 feedback interactions per lesson (Table 5.4). This is a wide gulf in understanding between parties. T2, as discussed below and in Chapter 7, believes she is providing VF in the learning instance which will help move student learning on, and also is shown to provide the highest amount of Process level feedback, yet the students cannot (or will not) recognise this. This contrasts with T1 students who both recognise and attend to the feedback that they receive. Both of these perspectives diverge from Glover and Brown’s (2006) study in which undergraduate students claimed to attend to their feedback, even when they failed to act upon it.

6.6.2 Clarity and understanding of feedback

In further exploring how students receive and engage with their feedback, it is possible to build a picture of two different classroom scenarios. In the first, T1 provided a large amount of oral feedback interactions (the highest at a mean of 256.5 interactions per lesson), and the students are cognisant of this. In the second, T2 provided a large amount of oral interactions (which at 209.2 interactions per lesson is considerably more than T3 at 68.3 and T4 at 127.8 interactions per lesson), an amount which far exceeded the written feedback in student books, and the students do not acknowledge the classroom discourse as VF. This does agree with a finding of Glover and Brown (2006) in which the undergraduates indicated that they did not find their feedback from their tutors to be either helpful or plentiful. Other studies have produced similar findings (Carless, 2006; Boud and Malloy, 2013, Carless and Yang, 2013; Evans, 2013; Carless and Boud, 2018); Carless and Boud (*ibid*) describe a student feedback literacy model to enable students to understand the information, and use it formatively in future work. Aspects such as these may prove useful for other age-phase teachers as they negotiate the waters of student communication, especially with respect to feedback.

Pre-intervention, the students were questioned about what feedback they had recently received, and this prompted a question as to whether they felt this was positive and/or negative:

> Interviewer: what sort of feedback has she given you recently?
> Bridget: I don't know really, just verbal and stuff. telling me what I can improve and what you've done well.
> Interviewer: So, when she gives you the feedback, is it a mixture of positive and negative?
> Connor: Yes
> Billy: Yes
> Interviewer: And which do you prefer?
> Connor: Both
> Bridget: I like having both because then you've got what you've done well in, and if she praises you for doing well in something, and if it's something you haven't got, then she's well, you need to work on this and it's like you know which bits to work on.
> Billy: I sort of prefer the negative feedback, because it pushes me a bit more to revise a certain aspect of it.
>
> [*T1 student group, pre-intervention*]

Here, not only do the T1 student recognise that they receive a mixture of positive and negative feedback, but that they appreciate the *nature* of the negative feedback. Billy asserts that he prefers negative feedback, or *discrepancy* feedback as characterised by Hattie and Timperley, 2007 and Voerman *et al*, 2012, in focusing his efforts into a certain aspect. As these aspects were explored with T2 students, they indicated that they also could recognise the difference between positive and negative feedback, although they were slightly more focused on written examples:

> Interviewer: So, whenever you get this feedback of whatever type, would you say it's positive or negative, or both or neither?
> Eloise: Both.
> Delia: One positive, one negative
> Dan: Yeah, you do get positive feedback - well done, but you do get negative feedback to improve on I guess, sort of both.
> Interviewer: Is that negative if it's telling you your next steps? Or is it negative in saying 'you haven't done this'?
> Eloise: I think he is quite positive, cause here he's put 'very nice answer', and he's put, 'very interesting, good application of reasoning', he's put 'much improved' [*Eloise using an example from her physics book rather than a VF example]*

In continuing to talk with T1 students about the nature of the VF they received, an interesting discussion emerged wherein the students showed that they recognised that discrepancy feedback seemed to be more important in progressing their learning. They did not seem to regard what they called negative feedback in a pejorative sense, so the interviewer asked whether they minded which sort they received. Here, Alfie shows that he appreciates the recognition that he has done well, but actually wants the specificity of finding out what he has done wrong so that he can correct it, saying that is how he learns:

> Interviewer: Does it matter to you, any of you whether the feedback is positive or negative?
> Alfie: Sometimes. I mean the positive always matters because you know you've done well, and you can apply that in the future. But you always need some negative in a way because you never get full marks and there's always going to be something that you've done badly
> Interviewer: Do you always need negative
> Alfie: well to improve my work, I'd say I would.
> Interviewer: So, is that going back to the corrections things then?
> Alfie: yes
> Interviewer: Or could it be what Anna said about making it better in some other way?
> Alfie: No, I'd say it's going back to the corrections that make me learn from my mistakes.

Billy reflects on how negative feedback does not feel negative, provided it is constructive. The manner in which T1 provides the feedback influences their relationship with the feedback and consequently with T1, as suggested by the 'ribbon' model of Elbra-Ramsay (2019) in which the relational, moral and pedagogical strands of feedback are interwoven.

> Interviewer: Billy what about you? Does it matter whether it is positive or negative?
> Billy: I think like a bit of both, because you need to know what you've done well, and you need to know what you can improve on, but as long as it's like constructive.
> Interviewer: Do you ever get feedback that isn't constructive?
> Billy: Sometimes but if it's like really bad…
> Interviewer: Can you give me an example?
> Alfie: [*mutters*] heat particles…
> Billy: heat particles [*laughing, others laugh too*] I can't remember one off the top of my head.
> Interviewer: OK, Bridget what about you? Positive or negative, or doesn't it matter?
> Bridget: Both I guess, because the negative helps you improve, the positive, I don't know how to say it, makes you feel better that you've got some bits right and the negative helps you improve it even more.

The students are able to see both aspects of feedback related to learning and affect and seem to welcome both the balance and the supportive process. It led to the interviewer asking what ratio of positive to negative they felt would be most beneficial for them:

> Interviewer: So, if you could ask for a mix, say out of 100, what proportion, what percentages would you rather have positive to negative?
> Bridget: Positive 60 – no negative 60, positive 40, because the negative helps us improve more than the positive, the positive just makes you feel better.
> Interviewer: OK. Billy?
> Billy: probably the same
> Interviewer: Anna?
> Anna: Yes, I'd say 60-40
> Alfie: Yes, I'd say 60-40 too.

This exchange highlights how much the T1 students value the feedback they have received, and that *positively phrased* negative feedback enables them to take agentic action in a non-demotivational way. The students agreed that had T1 given discrepancy feedback that was more negatively phrased, they would not have put the negative-to-positive ratio this high. A wish to have a greater ratio of discrepancy to positive feedback could be construed as a very healthy attitude to learning in physics, supported by this teacher.

T2 students did not report the same level of VF but did recognise that there were positive and negative aspects to it, relating to the use of WWW/EBI, but again with the predominance of written feedback. When trying to probe the nature of the feedback, Dan refers to it as mainly commenting and corrections:

> Interviewer: So, when he gives you this feedback then, does the feedback relate to what he has planned as a learning objective, or is it feedback as to how well you've done the questions, the correctness of it all? What do you think Dan?
> Dan: Well it's just how well you've done then really. if they're right, and they are right, how well you've written it.
> Interviewer: OK so it's the amount you've got correct?
> Dan: Yes.

This led the interviewer to ask of the students (pre-intervention) what they would like the feedback to be like:

> Delia: More like, if he's looking at what I've done, if he thinks I'm struggling or missing work, he should come up to me and be like, OK, I want you to go through this with me.

Two contemporaneous lessons were observed at the time of the initial student focus group. In the first (which was shortened due to the students completing the survey), T2 used 139 oral feedback interactions with the class. 55.4% were Task-related and 41.7% were Process related; the remainder were coded as self-regulation prompts. A second lesson three days later included 205 VF comments and questions; 62.4% Task, 31.7% Process and 5.9% Self-regulation. Yet the students do not identify this classroom dialogue as verbal feedback.

In returning to the students post-intervention, T1 students further articulated not only their continued appreciation of discrepancy feedback, but were also very clear in recognising that T1 provided a greater proportion of verbal feedback in the lesson than written in their books:

> Interviewer: Would you say that you get more feedback in the books than in the lesson or ore in the lesson than in the books? Charlotte, what do you think?
> Charlotte: In the books (*pausing, not sure*)
> Interviewer: Anna?
> Anna: I'd say in the lesson
> Interviewer: Alfie?
> Alfie: I'd say in the lesson
> Bridget: definitely the lesson
> Interviewer: Billy?
> Billy: In the lesson
> Interviewer: OK, so Billy, why would you say more in the lesson than in the book?
> Billy: well I think it's like Bridget explained, she can't write everything down in your book.

At this point, Billy realised that there was more feedback provided verbally, not simply because T1 could not physically write everything down in every student's book, but that what the in-the-moment verbal feedback contained was both resonant and unambiguous at the point of learning need, as the students went on to discuss:

Interviewer: OK, so are you saying it is the *quality* of the feedback that is better in the lesson than in the book?
Billy: yeah probably
Interviewer: OK, what if we go with the *amount* of feedback in the lesson versus feedback in the book? would you change your mind then?
Anna: That depends because, if you're having verbal feedback then you can ask more questions about it because you're having a conversation about it, but if it's written, you've got a specific question to answer…
Alfie: It's a bit more concise in the book as well.
Anna: To the point
Bridget: I think I get quite a lot in both, cause like, she puts quite a lot in my book, but she gives me it in the lessons as well, like how to improve, and because of that I have actually got better at physics.
Interviewer: that's nice to hear. which do you prefer, the written or the verbal?
Bridget: I think the verbal, because it's a bit easier for her to explain it to me, it comes across easier.
Interviewer: So, feedback in the moment…?
Bridget: Yes so, I feel like I can apply it right there and then, instead of trying to do it myself and getting it wrong again.

This rich description shows that T1 students know that they are getting feedback, that the VF is plentiful and has value, and that it trumps written feedback principally in the learning instance. That 'frequency [of feedback] is no guarantee of learning' (Sadler, 2010; Brookhart, 2012] would seem to be disproved by the quality and quantity of T1's verbal feedback.

To summarise, T1 students both acknowledge and appreciate the VF that their teacher provides, regarding it as instrumental in their continued learning in physics, and the nature of it is rooted in their mutual relational regard.

6.7 The emotional space of the classroom

Teachers' relationship with their class is a complex association, and one in which verbal feedback as part of classroom dialogue must play a pivotal and influential role. One of the three key influences in physics education highlighted by Murphy and Whitelegg's (2006) literature review was the teacher-student relationship, and how personally supportive the students found their physics teacher.

From the beginning of the study, T1 showed that they had warm personal regard for their teacher, which they felt was reciprocated, and that they appreciated her competence as a teacher of physics:

> Interviewer: What do you think about [T1] as a physics teacher?
> Alfie: Very good.
> Interviewer: Why Alfie?
> Alfie: Say, when we get your work back, if you haven't done very well, like I didn't do very well last time [*group laughs*]
> Interviewer: Was this the heat particles?
> Alfie: Yes, she went through it all really thoroughly with me, and what I could do to improve
> Anna: [T1] doesn't…lose control of the class
> Bridget: [*agreeing*] she doesn't lose her temper and shout at them, she just tells them off calmly and gets on with the lesson, and they're fine, but with some teachers, they lose control of the class and then just start shouting at you for the rest of the class and you don't learn anything, and she also explains things if you don't understand it like, if you don't understand something on the board, she'll go through it, whereas in other lessons, if you don't understand it, the teachers kind of move past it and say, well you should already know it, and she, like, goes through it thoroughly.
> Billy: she goes through a lot of the stuff if you've got it wrong, and she doesn't talk down to you. She talks to you like an adult which is better than most teachers do
> Anna: [*when giving negative feedback*] she puts a humour aspect on it, she jokes, she says in a nice way you should know that, without you feeling stupid
> Interviewer: So, if she's got something negative to tell you, she'll phrase it nicely, so you're not upset by it?
> [*group all agrees*]
> Connor: It's never a put down
>
> Charlotte: but if it is really negative though, she puts a humorous sign on it
> Billy: Yeah
> Charlotte: So, it's not as hard
> Anna: So, you don't feel as bad about yourself
> Alfie: but you know you've done something bad.
> Charlotte: but she doesn't make you feel bad about it like others.
> Billy: like some teachers -
> Bridget: Hm hm [*confirming*]
> Billy: - that say you're awful.

Notably, every student contributes to this defence of T1 to give a picture of a physics teacher who fosters good relationships, creates a classroom environment conducive to learning, and crucially, one in which mistakes can be worked through in a beneficial manner for learning (Rimm-Kaufmann and Sandilos (n.d.). By contrast, T2 students did not share the same relationship or enthusiasm for their teacher, yet did not seem to fully realise how he had flexed a lesson in response to their learning need:

> Interviewer: [*to Delia*] You said it [*physics*] was boring. Why is it boring?
> Delia: it's just that there's nothing…there's no example that interests me at all.
> Edward: Yes, same, it's so unrealistic. The things that we do it's like – what? That's so pointless. It's like James throws a ball in a river – who cares about that?
> Delia: Why would we want to pay attention to James?
> Interviewer: OK
> Delia: So, you just switch off.
> Interviewer: So, would examples that meant a bit more to your real-life change things for you?
> Delia: Yes, so [T2] said one earlier, didn't he, and we were like 'what'? so he changed it into skiing, and we were OK, now I understand this.
> Edward: Yes, because he said about going up in a ski lift, and that was like potential energy, and when you came down that was kinetic energy, so I got that a bit more.

The students were not hostile, but they had yet to engage with either the physics and/or their teacher. They could say that they would like more contextualised examples to help them understand their physics but did not really show appreciation when this was done for them. Klem and Connell (2004) asserted that between 40-60% of secondary age students have become disengaged from school, and linked teacher-student relationships and learner support to student engagement and achievement. Here, the students are socially compliant but appear not to mentally participate in physics:

> Eloise: In physics its quite easy to get away with talking and not listening
> Delia: Yeah.
> Eloise: – and you miss things.
> Delia: I can't name of a single person who can sit and listen in a physics lesson for the full hour. In physics, he'll talk, and then no-one really knows what to write. He doesn't write that much on the board so we wouldn't know what to write down anyway.

From this student-centric view, the irony of the situation is unseen. Eloise and Delia are commenting here on several issues; a perceived lack of classroom control by the teacher, an overload of teacher talk, a lack of notes and thus they regard themselves as unable to write physics notes. In context, the teacher has pre-prepared the class materials in ‘Showbie’, an educational app used to reduce the amount of paper used in their classrooms.

As the study progresses, the students remain unconnected to the teacher, and the relationship never achieves that of T1 and her students. Nevertheless, they are, despite the comments on talking and not listening above, generally compliant in class, and do feel that they can ask T2 any question when they require support [field notes]. Dan’s post-intervention comments in section 6.5 above show that the students were moving to a more collective sense of responsibility, and personal agency, poignantly captured by Eloise:

> I wish I could back to the start of the year, like when I didn’t work as hard. I worked hard, but halfway through the year I really started trying at subjects.

T2 did indeed use positive feedback, but this was not perceived by the students, who perhaps considered the higher use of Process feedback a disabling factor rather than an enabling one, resulting in, not a conflictual relationship, but not an entirely easy one either.

6.8 Chapter Summary

This chapter has explored the student voice in the study, resulting in a rich consideration of physics self-beliefs and student-teacher relationships from the adolescent perspective. It is clear that the respective associations of these teachers with their classes has resulted in a different physics learning experience for each cohort over the study. Also, each class appears to have a different view of the type, nature and amount of verbal feedback that they have received; T1 students in acknowledging and using the VF to ‘make themselves better at physics’, and T2 students not recognising the amount of VF that they have received (Yang and Carless, 2013), and whilst benefiting from this in the classroom space, as shown by the survey results, still wish for more.

In the students’ language in both cohorts there was a sense of personal agency developing, which may link with the feedback they had received, but this could also be a maturation effect over the study. Nevertheless, data from the surveys together with this student voice would indicate that, in response to RQ1 and 2:

> *What are the students' self-belief systems in physics as they enter KS4?*
> *Do the students' self-belief systems in physics change during the life of the study?*

evidence has been provided that maps student self-beliefs in terms of SC, SE, anxiety and to some extent mindset, and has shown these to improve in physics for the IG cohort, but not the CG cohort.

The evidence to answer RQ3:

> *What types and proportions of feedback do teachers use in their verbal interactions with students*, and RQ4 *Has the feedback pattern changed*?

has been viewed through the lens of student perspective which shows that students are able to describe and discriminate between the various forms of feedback they receive, and under optimal learning conditions, enabled by a personally supportive physics teacher, can use mistakes to progress their learning in a non-demotivational way. They do not however, largely perceive a *change* in feedback from their teacher, although the T2 cohort in a post-intervention state, did concede more that feedback was taking place. This could yet be due to maturation.

Finally, and ironically, given that VF is a form of classroom communication, it would appear that the different way in which these classroom teachers have conveyed the message that feedback is taking place could have implications for teacher-student interactions in general, and this will be explored in the Discussion in Chapter 8.

Chapter 7. Teacher case studies

7.1 introduction

The previous findings chapters explored variations of student data with respect to the pre- and post-intervention surveys in self-belief and the student articulation of physics self-concept and feedback, as well as a presentation of the feedback styles of the teachers involved in the study. These were in relation to the research questions:

1. *What are the students' self-belief systems in physics as they enter KS4?*
2. *Do the students' self-belief systems in physics change during the life of the study?*
3. *What types and proportions of feedback do teachers use in their verbal interactions with students in physics?*
4. *Does the feedback pattern change during the life of the study in response to a CPD programme of flexing verbal interaction styles?*

A fifth research question asked:

5. *To what extent can data evidence an impact of RQ4 on RQ2?*

This longer chapter will focus on the participants who received intervention professional development during the study, and attempt to provide further data to answer RQs 4 and 5 in using the teacher perspective to explore whether the participants themselves attribute any change in feedback styles to a change in their students' self-belief systems.

7.2 The participants

Both intervention teachers were interviewed throughout the study in addition to the initial professional development meeting. They each demonstrated personal investment and engagement in the study, though each had different teacher 'personas' with their classes, and differing personal connections with the students, resulting in distinctive dynamics. T1 was an energetic, empathetic educator who enjoyed a 'bantering' relationship with the students in which even slight sarcasm was taken with good humour as the intent was clearly benign. The students in the class worked well together even when moved around. T2 did not have the same closeness with the class although the intentions were always positive and professional. In this group, there were some student factions, and less interaction between the members of the group as a whole. This contributed in some way to a less settled learning environment.

The interviews, or coaching conversations are used chronologically within this section to portray teacher development. Table 7.1 below shows when these interviews took place both within the academic year and the timeline of the study.

Interview	Month	Month of the study	Teacher
1	February	2	T1, T2
2	March	3	T2
2	April	4	T1
3	June	6	T1
4 (T2=3)	July	7	T1, T2

Table 7.1. Interview calendar of IG teachers

7.3 Teacher 1

T1 was an experienced teacher of physics who had been at the school for a number of years, teaching science to some of the current class in previous academic years. Beyond some in-service CPD at the host school, T1 had not received other professional development for some time, and this contributed to the motivation to participate. Several features emerged from successive discussions with T1; her notions of, and intentions regarding taking part in this feedback study, the deep and informed nature of her ‘reading’ of her students, and their learning needs, her efforts to flex her ‘default’ feedback style against a background of engrained habit, and the reflections she shared on her own development.

7.3.1 Notions and intentions of feedback

In the intervention CPD (Appendix 3), drawing on Hattie's (2012, personal communication) commentary that feedback appears to be seen as a unidimensional notion understood as the same thing by all, the interviewer asked T1 to develop her own definition of feedback for the purpose of the discussion:

> I would say the feedback is an analysis of what has been done and it's quite circular until it is a discussion rather than a statement often for me. So, I would say it's an analysis what has been done and what's been done well, and what needs to be done to improve on anything on a particular activity or in a thought process or in a way that, um, you know, an approach or an analysis [*T1, initial interview*].

The interviewer did not immediately evaluate this response, and introduced a slide showing feedback statements from Wiggins (2014) in which the author had offered apparent feedback statements, later asserting that none of the statements were feedback, describing the middle two as advice, and the first and last as merely expressing a liking for the work. T1 was asked her opinion of these as feedback:

> I would say they're all feedback of some description. I would say some are more useful than others and it depends on what you're trying to achieve. This [*pointing to first one*] is feedback because it's telling them that they've done a good job, it's improved boosting them because feedback, I mean in terms of like electrical feedback, you know, it's just something that goes back to them and gives them something back. And so, I would say all of these are a form of feedback. This is telling them less about what they did [*the second one*]. I mean it's something about, you know, it's feeding back in terms of telling them what they did, but this [*the last one*] is the one that's least appropriate because you just simply stating something. So, whilst you're still feeding back cause it still going backwards and it's still a circular motion, you know, actually they're not going to gain very much from that apart from the fact that you love ants and therefore you've bought into the kind of their story. But I would say all of them were feedback of one description or another [*T1, initial interview*].

The interviewer gave the background to the statements and then revealed the amendments the author had introduced, asking again, but with respect to the definition produced by T1, which she now considered to be feedback:

> Interviewer: So, if I asked you, you again, which of those are now feedback? If we go by your definition of feedback, which of those now qualify?
>
> T1: In terms of our school in using What Went Well/Even Better If, some of these would fit in perfectly. You've said what they've done well [*pause*]. Um, but this one [*first one*] doesn't give an 'even better if', this one doesn't give them a way to move on forwards. This one [*second one*] sort of says what the problem was, and that just says the problems and doesn't tell them how to fix it. Oh no. Yeah, no, that doesn't tell them. It just says make it clearer, which is fairly weak, I'll say. But I don't know. I'm not sure any of them would fit my idea of what is required by feedback, you know, what is expected from me in terms of feedback, which is a WWW/EBI. But, um, some of them are saying this went really well. Others of them were saying this went, but not very well. But it's not really sort of saying this, this one [*third one*] is the closest where it says, you know, your visual is more, more polished and supportive of the teaching, but it doesn't really say *how* that would happen.

Initially, T1 assessed these amended feedback statements as agreeing with her personal and school policy ideals. As perusal of the text continued, T1 re-evaluated, deciding that these went some way to providing feedback statements [WWW], but not EBIs, i.e., next steps prompts for students to take action on their learning. Through this reflection, T1 shows an ability to distinguish between different levels of both commentary and constructive criticism.

In a subsequent coaching conversation, the interviewer explored T1's motivation in what she was wanting to achieve with the feedback she used:

> T1: I'm wanting to understand how they can progress, and also to make them better scientists and believe in themselves, and ultimately, um, ultimately to help them progress. Which can be as much of a confidence builder as anything else.
>
> Interviewer: Okay. So, what does success at progress look like for you then?

> T1: Partially it's partially them getting their expected grades. [*motions inverted commas*]. A lot of these guys are getting or exceeding their expected grades actually at the moment so it's more for this lot about getting them to feel like they can actually do it, and also finding it easier because a lot of them struggle with certain things. And they don't, they don't always find it easy, and so it's about getting them to the next time you come to it, find it easier, and to understand it better, and to be able to explain it better, and to be able to try it more easily, and to be able to not to give up quite so quickly.

Here, T1 makes several interesting points. As discussed in section 7.3.4 below, T1 experiences a tension between the performance goals [*exams grades*] of her institution, and the learning goals she espouses arising from her own values. T1 recognises the importance *to* the students of experiencing self-efficacy in how they approach their own learning, as well as the associated, hopefully positive motivational aspects they experience during the attempt. In asking the same question at the final interview, T1 articulates an even stronger arc:

> T1: I wanted to improve my feedback so they could get more from it about how to make better progress themselves and to make them have faith in themselves because I think a lot of whether they do well or not is their confidence, and their ability to think they can do it, and that they *want* to do it as well, because they get something out of it. And that's across the board. And I like to be a positive feedbacker [*sic*] and I suppose that's what I think I've found out, that I do a lot of positive backslapping, but possibly for the wrong reasons, and that might have a negative effect on some of those students that I don't think I'm doing…I don't *realise* the negative effect it might be having sometimes.
>
> Interviewer: When you say the wrong reasons, what do you mean by that?
>
> T1: Well sometimes I might appear to be praising the fast worker, the neat worker *over* the likes of Zack, who can think out of the box, or Sam, who asks good questions but shouts out, and doesn't write anything down. So, he gets minor praise for asking the questions, and being involved in the lesson, but also gets a negative kind of response sometimes. But then, when I'm going around, he doesn't get the positive, 'oh you've finished, oh that's neat, it's pretty, you've done well, you've got it right, because he doesn't do any of those things.

From a student-centred perspective explored further in the section below on Teacher knowledge of student learning dispositions, T1 immediately highlights the importance [to her] of improving her feedback to make a difference to her students' learning. Secondly, she highlights that it is about these students being able to manage this academic progress for themselves. A third and fourth point references their confidence/self-efficacy as well as their motivational likelihood to make an attempt. Lastly, T1 has recognised a long-standing tendency to give a lot of undifferentiated praise within the class, which whilst contributing to a favourable learning environment can also have unintended negative consequences in the way in which the students receive and internalise this generic feedback.

In reflecting and engaging more deeply with her feedback repertoire, and acknowledging that students may have acquired a different message than had been intended, T1 showed that she herself had learned how to use her verbal interactions in different ways, even as she recognised that she was not yet doing this to the extent that she would have wished.

7.3.2 Knowing the learning dispositions of the students

T1 demonstrated a very able knowledge of her students…. In the discussion on 'mindset' in the intervention CPD, she exemplified how some of her students attribute their success:

> Like Zara was saying that, 'why doesn't Ms Cooper like me?' And she's internalized it. She very much internalizes everything. It's all about, she was the one who was asking which teachers… She's already worried in the middle of Year 10; which teachers will be teaching what at A level so that she knows what to concentrate on at GCSE because she doesn't want to do it. Those teachers might *leave*. You've got to make choices about yourself. Whereas Charlotte is an external, 'I was lucky'. She's an external, stable specific [referring to a table in the PowerPoint] which is probably the worst you can be [*T1, initial interview*].

This led into a discussion in which the self-belief questions for the pre-intervention survey had been answered by the class. T1 recognised that they were likely to choose to respond in a way which would show (for example) their mindset in a more favourable (to them) way, yet their day-today attitudes and speech may convey otherwise:

> They're quite bright and they're quite clever and they know what they're supposed to write. They know you're not supposed to say that intelligence is fixed. So, they will all have answered probably mostly, ‘No you can't. You can't say that intelligence is fixed at birth. You can help yourself’. Even if they don't believe or believe in *that*. [*T1, initial interview*]

This became a useful episode in which T1 considered which students could form a focus group of mixed presentations of self-belief in order to both map potential SB changes, as well as provide richer student voice on the learning and feedback process, such as the inclusion of Charlotte, above.

As the academic year progressed, T1 continued to grow an awareness of the learning needs of the students. She was able to identify strengths and areas to develop in individual students and relate this to a distinction between being able to competently perform (e.g.) a calculation, and then using that same process in a different context:

> Interviewer: Okay can you give me an example of something they are struggling with?
>
> Well they are struggling with this [*indicating that day’s work on acceleration*]. Anna in particular struggles with the rearranging and things. So, they struggle with maths. And they struggle with applying it to new situations. So, whereas I can go, 'well you've just done that calculation', they're like, well it's a completely different calculation, it's got different numbers.' So it's about getting to see that they can do it and do it over and over again, in different ways and basically getting them more comfortable with the stuff, so that when they come to the question they can look at it and go, 'oh yeah I can do that', or at least have a go at it, and want to fight their way through it and find what the answer is, errm, which they sometimes don't.
>
> Interviewer: Does this mean that they're looking at something that is very context specific and they don't take the process and apply it to a different context?
>
> Yes. Yes, I think they really struggle with that. And then I can help them with the process in a particular context and they're like yeah okay I get

> that now, and then they can't move that process over. They look like they're doing really well, but actually if you look at their processes and their learning and their actual understanding, it's not there yet. [*T1, second interview*]

In a subsequent meeting, T1 commented generically on her class [Year 10] being able to see more connections within the subject, and specifically on Bridget's [a member of the student focus group] increasing confidence:

> See, the other thing is that I think they tick over, part of going from Year 10 into Year 11 is that they tick over and become more confident anyway, because they start to see the big picture, and they start to see how things inter-relate…so Bridget came in, amazingly, because she usually does *no* homework, but she came in and we talked about proportionality last lesson, or the lesson before, and she was like, 'you know that fish you drew? That proportional fish? We did that in maths today', And it makes you realise doesn't it, that I'm expecting them to realise what proportional means at the beginning of Year 10, and they haven't met it yet. And she was saying, 'yeah, I knew I could answer that question, because I knew that we'd done it in science which was really good. They should be able to bring more maths with them [*to KS4 science*] to my lessons than they do. I think there's an element of they just compartmentalise it.
>
> Interviewer: Could it be a blockage for them?
>
> Yes, they don't take things between lessons, so I think that was a really big thing for her [Bridget]. [*T1, second interview*]

As discussed in chapter 6, a girl in the focus group (Anna) admitted to not revising for 'the exams' to save herself some 'stress'. In the final interview, T1 and the interviewer were discussing how the students had approached the summer examinations:

> T1: A lot of them have taken these exams really seriously and done really well. Zara, bless her, has worked really hard. I'm very impressed, because she's always been good, but I'm really glad she got her [high mark] that made her day. For some you do see the diligence coming through Well for her that's the diligence, because Anna didn't work for the exams, she chose not to.
>
> Interviewer: Did she say that?
>
> T1: Yes.

Interviewer: Why did she choose not to?

T1: I don't know. She said, that she didn't really think that they were very important and that she wanted to see how she could do, and there wasn't a lot of point, and the thing is, she's done OK in physics, but I don't think she's done very well in chemistry, I think she got a C in chemistry, and I think she got a U in the biology. And when I said this to the biology teacher, she said I did wonder why she didn't look too worried about her U. She's basically come out and gone, well I didn't revise anyway. But I made a big show of saying well, the biology paper was actually quite easy, it was not a hard paper, it had two grade C/D part questions on it. You should have been getting 100% on those questions and you got 2 out of 11. I didn't say it like that, but I said I think [teacher name] was quite disappointed.

Interviewer: Well it will be interesting to see what she says, because one of the questions I'm going to ask is 'could you have done better?'

T1: Yes, that will be interesting.

T1 appeared frustrated by this student's decision, which Anna had rationalised to herself. At this point, T1 was not privy to Anna's decision-making process, explained by the student in section 6.3 in the way she deemed herself to have escaped attendant examination anxiety by choosing this course of action. Anna also went on to say in 6.5.1 that she had stopped comparing herself to Zara, who T1 saw as working very hard for the examinations. Here, both T1 and the biology teacher appear personally, perhaps even emotionally invested in their students' success. T1 continued by discussing the progress that Billy had made both in terms of personal motivation, which whilst not universal, did appear to have moved in a positive direction, in modifying his process of demonstrating his 'working out', discussed in an earlier interview:

Billy did some work for them. I've been badgering Billy about showing his working, and he has been, then the other day, he didn't show his working, and I was like, oh no, he's regressed, but he doesn't see the point in it. But he's been working very hard, he's been quite focused. I think he's bored, to be honest, but he rushes through work.

Maybe he needs to learn to check?

It will be interesting to see what he says [*to you*], because I think he's very quiet, and if you don't get on him, he doesn't push himself forward. I think he's a bit bored, but I can't quite put my finger on it.

The student whose progress had impressed T1 most was Bridget, who had been identified via the pre-intervention as low SC, high anxiety, and tending towards an entity theory of intelligence or ‘fixed mindset’, shown in Table 6.1:

> Because I think Bridget worked hard for them. I think Bridget has come a long way this year actually, she’s making links between subjects, and things, you know she was talking about proportionality. She was saying you know that little fish thing you did, well we did it in maths and I knew what it was, which was great. It was exactly the boost that she needed, and she has really come on. She did well in her physics exam, unfortunately I don’t think she’s done quite so well in her chemistry and biology, which is a shame, because they do all get wrapped up into one a little bit, but she did well in physics, and I know she wanted a B. So, she was getting Ds at the beginning of the year and now she’s getting C and Bs. I said something to her along those lines, I said you are right on that cusp, that’s brilliant, you have really proved that you can do it, but I think that I need to incorporate more of those statements, and particularly the [specific] affirmation ones, because my usual ones are ‘well done’, and ‘you have done well’, and they’re like: ‘well, *what* have I done well? *I’ve* got a C and *they’ve* got an A-star.’ [*T1, final interview*]

The teacher assessment here correlates strongly with Bridget’s own personal reflections in 6.5.1 in which she attributes a positive change in her attitude towards physics, and an increased sense of independence, as well as effective feedback from T1, to a higher grade in the summer examination.

This section draws together some critical data to synthesise a response to the questions:

2. Do the students’ self-belief systems in physics change during the life of the study?
4. Does the feedback pattern change during the life of the study in response to a CPD programme of flexing verbal interaction styles?

and hence begins to develop some evidence that the feedback CPD intervention would appear to have had some impact on student learning and attitude according to:

5. To what extent can data evidence an impact of RQ4 on RQ2?

Overall, T1 showed an extraordinary level of knowledge and understanding of both the attitudes and academic attributes of the students in the class, displaying a personal investment in them as individuals as well as learners, and developing notions of how to increasingly *specify* the nature of the positive feedback she employed. This will be further explored in the next sections theme.

7.3.3 Flexing the use of language to calibrate feedback

The purpose of the study was to professionally develop teachers to be able to vary their verbal interactions with their students to purposefully employ more process and self-regulation feedback at the learning instance, with the stated intended outcome of measuring whether this could have an impact on their self-belief systems in physics. This has been referred to throughout as *flex* to distinguish it from a simple change. Through the professional development background of the researcher, it was not anticipated that this would be either a swift or indeed easy process, hence the requirement for a longer period of study.

During the first meeting, in which the feedback intervention was introduced, the concept of *performance* goals versus *learning* goals was considered (Reay and Wiliam, 1999; Dweck, 2000, 2006), and consequently how students may experience motivation differently:

> T1: Some of the kids in this class will be obsessed with the grades.
>
> Interviewer: Because [*under our current educational system of ranking schools according to student outcomes*] *schools* are obsessed with the grades and they send a very clear message that 'we value you according to your target grades'.
>
> T1: Yeah. We're going to value you a little bit more if you've got higher target grade, you know, cause that's going to look good for us as a school.
>
> Interviewer: So, this is the message which institutionally can be enacted…
>
> T1: Yeah. Even if what you say is, you know, very *nurturing*.

This developed into a dialogue about how such interactions might be received and internalised by the students, and the behaviours or responses that may come from these, centring firstly on how such nurturing talk can often appear under the banner of praise, colloquially deemed to be ‘a good thing’, but is in fact undifferentiated feedback (Dweck, 2000), and may produce a response of muted effort. T2 was shown some examples of phrases teachers or parents might say as ‘praise’:

> Interviewer: So, these are the sorts of things which would come under the banner of praise, but actually they're undifferentiated. They're non-specific. Sometimes they're just fibs. Like, ‘you're a fantastic swimmer’, but actually you've just gone across the breadth of the pool there, Rufus.
>
> T1: So better is wow, ‘your swimming is really coming on. You couldn't do that last week’. Yes. Yeah.
>
> Interviewer: And a lot of this, there is a take home message that goes along with it that, that Dweck points out when she says our language tells people what we value. So, when we say to students ‘oh well done, you did that really quickly’, then the take-home message is –
>
> T1: -do it really quickly
>
> Interviewer: ‘Miss values speed, so I’d better speed up on this or she won’t value *me’*.
>
> T1: Yeah
>
> Interviewer: Or well done, you’ve got everything right. ‘So now I’m only get praised when I do everything right. And Carol Dweck says, what should we do when someone has got everything right? She says this is a time when we are sorely tempted to give praise. Well done you’ve got everything right. Instead she says, you should apologise.
>
> T1: Yeah for making it too easy
>
> Interviewer: [*agreeing*] You should say, I clearly need to give you something more challenging to do.
>
> T1: Yeah
>
> Interviewer: And she says we shouldn’t be praising them and giving them tasks, which shows them that low effort is right.
>
> T1: Yeah, yeah
>
> Interviewer: And that’s a really hard message.
>
> T1: It’s a gut thing, isn’t it? We say well done you’re finished.

> Interviewer: Yes, we say ‘well done’ all the time. All the time. Because it’s easy, and because we’ve been brought up on it. And it’s difficult to come away from it.

The training then enlarged on the other levels of feedback, as outlined in Table 2.1, which likened the ways teachers use Task feedback as a comment, criticism, clarification or confirmation on the product or artefact, or as, simply, ‘the work’. Process feedback however related to cognition, or how the work had been ‘done’, whereas self-Regulation interactions were associated with motivation, persistence and effort. The preliminary analysis of the first observed lesson was shared to introduce a baseline for T1 of the proportions of levels that occurred:

> Interviewer: So, the breakdown on this particular lesson, your particular breakdown was 60% Task. But this is where the *context* becomes quite important. This lesson was about calculations. You were helping them *work something out*. And I think we saw a shift because of the context to a higher proportion of process sorts of feedback, which was 28%. You did about 2% sort of undifferentiated and 5% Regulation, which is actually *higher* than on those studies.

A slide taken from a National Strategies (2009) was used to promote an examination of positive versus negative feedback (Figure 7.1), in which different aspects of ‘negative’ were explored; as undesirable, or adverse, and in the sense of the ‘closing the gap’ or *discrepancy* feedback (Kluger and DeNisi, 1996; Black and Wiliam, 1998a; Hattie and Timperley, 2007; Shute, 2007, 2008; Voerman *et al*., 2012, 2015). The argument was made by the source that the most effective feedback was always positive and specific, and the iconography accompanying the assertion could be argued to be leading opinion of which type was ‘best’, using examples of feedback to sort into the relative parts of the crossed continua:

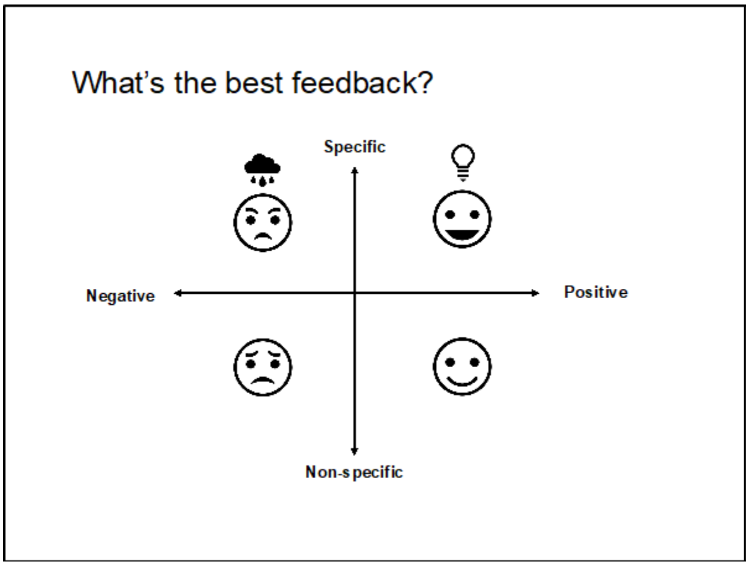


Figure 7.1 Slide 18 CPD materials. Graphic modified from DCSF (2009)

Interviewer: So, what would you say is the best sort of feedback, positive, specific, positive, nonspecific. If you had to rank those, what order would you put them in?

T1: It depends what you're trying to achieve and who it is. So, in terms of what will help the student most, generally assuming that they will move forward with positive feedback…. So, if they would move forward with a next step, then that would be the best one. But probably, it's a combination of the two together. So, you know, you did this really well, but this is how you would approve it. I would say that would be the best option. Um, I think if you're going to give praise specific praise is helpful. I think the non-specific negative, there's no good at all, ever. That's just not worth having, you know, 'you're rubbish', 'You didn't do that very well', kind of thing. Unless it's hugely sarcastic. Like 'don't use heat particles' [*laughing, said self-deprecatingly*], but um, you know, like kind of, I don't think this, I don't think this is very useful at all [*negative non-specific*]. Then there are times when kids, some kids just need a boost and they just need to be told they're good [*positive non-specific*]. And, but then I would say in terms of actually moving forward, a combination of those two is the most important [*indicating negative-specific and positive-specific*]. And in terms of moving them forward, that's probably the most important [*positive-specific*].

Interviewer: Well, teachers talk about feedback being kind, helpful and specific. and it seems the argument is almost overwhelmingly that it should be positive, it should be specific, and it should relate

> to the success criteria. But some teachers have said to me, well *negative*, because actually we're reporting on a *gap*, you know, because you talk about closing the gap so you're reporting on a negative and they can take that as a negative. But I think here, this is more to do with how you use your *language*.
>
> T1: Yeah. Reading *that* [*a negative specific comment*], I'm thinking, that's not what I mean by next steps. Yeah. That what I mean is *this*, with a 'now to improve you do *this'*, whereas this says you didn't do X in your presentation. Which is, you know, kind of not, not helpful at all.

Here, T1 is very clear that her mental model of negative feedback is that of discrepancy feedback, and that it should be couched in such a way as to enable the student receiving it to be able to independently take their learning forward by having clarity about what it is that they need to do. She also talks about 'when kids just need a boost' in terms of non-specific good feelings. When this is related to the passage in section 6.6.2 in which T1 students assert a ratio preference of 60% 'negative' feedback to 40% 'positive' feedback, largely due to the *kindness* of the language T1 uses, it would appear to support her decision to feedback in this way as successful.

There followed some training in which T1 was asked to give feedback on an activity but could only use process and regulation feedback statements and prompts. The teacher struggles with this, as she was more used to giving feedback on the finished, rather than developing artefact. Hattie and co-authors (Hattie and Timperley, 2007; Hattie and Masters, 2012; Hattie, 2012) have suggested that feedback should be calibrated relative to whether the learner is a novice, somewhat proficient, or competent, and that the feedback should range accordingly from Task to Process to Regulation. This is contradictory to Shunk's view (2007) that regulation feedback is most beneficial in the early stages of learning but should shift to process feedback if the learner is trying, but not succeeding, in order to avoid becoming demotivated. T1 captured it thus:

> So, my job over the next week or two is to try and incorporate and shift my feedback to being more about this um, sort of resilience and effort

> and process? What they're doing well, in terms of their *input* as opposed to what they're doing well in terms of their *output*. [*T1, first interview*]

T1 then continued to work on this and recorded lessons. A coaching conversation was scheduled for two months after the intervention CPD. A lesson was observed, preliminarily coded and field notes documented; these were subsequently used as a basis for discussion in the coaching conversation, where T1 looked at the verbatim feedback and considered how this could have been modified to stimulate increased effort, or support the processing of a response:

> T1: I've just seen this one [*points to comment made during observation which said: Billy, are you working it out? Boys responds 'no'. Are you even trying to work it out?*].
>
> Interviewer: Yes? Okay, so what would you to turn this to?
>
> T1: How are you..., so Billy, are you working it out (no), so how could you work it out? What do you need to do to work it out? [*thinks*] I would turn this into, I *should* have turned this into: 'So, what do you need to do to work it out, how are you going to work it out? Explain to me where you would start?
>
> Interviewer: Okay. And how would you shift it into effort? Because you are talking about process there, weren't you?
>
> T1: Yes.
>
> Interviewer: How do you shift it into, him rising to the challenge?
>
> T1: Yes! So, kind of like, come on Billy, you know how to do this, have a go. Why don't you tell me what you think, or, at least have a look at it, because I think you can do this, yeah.

The suggestion was made that T1 could write down some similar prompts prior to the lesson and position them on the side of the desk to remind her to look at them, and further lesson comments were scrutinised to enable practise of flexing the feedback away from instructional or task interactions. In particular, the notion of depersonalising the feedback arose:

> Interviewer: Sometimes it's the way that it is phrased [*pointing: you're making it too complicated…*] *You're* making it too complicated. How would you feel with someone saying that to you?
>
> T1: Yes [*nodding*]

Interviewer: What could you say instead?

T1: [*Murmuring to self, 'you're making it too complicated'*)... It's not as complicated as that, errm, it's far easier, you can do this,

Interviewer: Yes. Do you see how that takes away the personalisation?

T1: Yes

Interviewer: So instead of it being *you*, that it is some fault inside of you, that you're making it harder than it need be, some way of shifting it into a less pejorative, less judgemental way perhaps, of doing it?

Yes, yes... It's like with my year 13s, they did really well with a little set of questions, and the first thing I said was, well, I didn't particularly choose hard questions. And what I meant was, not that they were really easy so that was a rubbish mark. What I meant was great, that's really good. Now we need to move onto the next step. But that's not what I said I said that they weren't particularly hard questions. So, no I can see that, and that's it. It's an automatic, yeah, it's not as complicated as it looks would be better.

Interviewer: Or it's challenging, but I know you can do it?

T1: Yes!

Interviewer: And then you're citing your self-belief as well, and the more you talk about self-belief, and that you believe in them, the more likely they are to believe it themselves.

T1: Yes. Now you see, interestingly, I said that [*tapping page*] a lot. I think I said it's a really challenging piece of work, but I want us to have a go at it, I think you can do it. I think I said that a few times, when I honestly thought it was a challenging piece of work. And I think when it's not that is challenging, I just go, come on it's easy, it's not that difficult you're making it more difficult than it needs to be. So, what I need to do is...say, yeah...

Interviewer: Perhaps when it's not that challenging, is when you recognise the effort?

T1: Yeah.

Interviewer: You could say, I'm so proud of you, you worked really hard on this, or you really got your head down and look at what you've changed.

T1: Yeah, you're really working hard to solve this problem, let's see if we can make it a bit more straightforward. Yeah okay.

Here T1 recognises two different issues. Firstly, she uses language such as 'difficult', 'complicated', 'hard' when internally she is thinking about challenge, and the way she expresses this to students could reinforce negative self-concept (Murphy and Whitelegg, 2006; Rattan *et al.*, 2012). Secondly, in the way the interaction is presented to the student, there is an aspect of personalisation which could reinforce the student's 'take-home' message that they as individuals are struggling with physics, where peers may not. This perception and comparison may of course be false, however in considering the notion of self-concept as 'the reflected self' (Tice and Wallace, 2003), students may be disposed to make comparisons to their peers in such a way. The conversation continued to examine the potential student interpretation of verbal interactions:

> T1: This was Bridget. 'I've tried it, but I can't do it.' Me: 'You've not had time to try it yet.'
>
> Interviewer: Because that response was immediate wasn't it?
>
> T1: Yes, it was.
>
> Interviewer: And you're right, because she *hadn't* actually had time to read it properly…
>
> T1: No [*laughing*].
>
> Interviewer: But again, turn it into, something that's about rising to the challenge...?
>
> T1: Would it be something like, errm, so she sort of said I can't do it and I said, have you tried? And she said I have tried it, so I could've said, what have you done so far? What do you think you need to do? Good start, but you need to… [*pause*]
>
> Interviewer: With someone like Bridget, who we have identified as someone who will walk away from the problem rather than rise to it, if we say things like, you haven't tried, you need to try harder -
>
> T1: She will just back away from it.
>
> Interviewer: Yes. Instead we might need to say something like, which particular bit are you getting stuck on? You identify where you're getting stuck, for someone like Bridget, you could shift it to more into the process or something that she can do, or strategies that she can try?
>
> T1: Yes [*writing*]
>
> Interviewer: And then when she has succeeded, that's when you say…?
>
> T1: Well done, that was really good.

> Interviewer: - or how about, look how you really stuck with that, look at what you achieved.
>
> T1: Yeah. I know this, I say this, it comes out of my mouth without my having processed it at all. It's kind of partly a joke, come on you've not even had time to. Because then, I went over and talked to her. What I did was say was, right then, let's start, we're doing this, what's this? What's that?
>
> Interviewer: Yes, starting to break it down...
>
> T1: Yes, but then, *I'm* breaking it down for her. Yep okay. So, what I need to do is flip it so that *she* is breaking it down…So break it down for me, Bridget. Let's start at the beginning, read the question, tell me what you can see from the question.
>
> Interviewer: You might even try giving her that little bit of thinking time? 'So, can you identify where you're going wrong, which bit is the trickiest for you? I'll give you a few seconds to think about it and come right back'.
>
> T1: Hmmm [*writing*]
>
> Interviewer: And then she knows that she's got something to do, instead of you launching in and helping her.
>
> T1: Because the other thing, is that when she then did the entire sheet, I went past and said well done or that sort of thing. That was it. That was the sum total despite... She struggled at the start, but that was the sum total of the praise that she got.
>
> Interviewer: Yes, or the recognition
>
> T1: Yes, the recognition of her effort. Yeah, yes okay.

In quickly responding with 'Well done, that was really good' as a suggestion for flexing feedback to effort, T1 typifies how easily she slips into generic praise, which she has described as 'back-slapping'. When the interviewer proposes 'look how you really stuck with that, look at what you achieved' as a shift away from task-generic praise towards effort and persistence, T1 acknowledges that 'it comes out of my mouth without my having processed it at all'; here she is in her *default* feedback mode. However, unprompted, she considers what she did say and do to support Bridget, self-identifies that she herself is doing too much for Bridget, and needs to challenge the student to take ownership of the learning. In disclosing what she then went on to give Bridget as feedback upon completion of the work, she reflected upon both the repeat of generic praise, and a lost opportunity to genuinely remark upon the resilience Bridget had demonstrated, and build the confidence of a student who demonstrates low self-concept in physics. To practise more effort feedback, another interaction was selected:

> T1: Which one was this? 'Do you want me to do it?'
>
> Interviewer: Was that Connor? Something that he hadn't finished? And he said, ‘oh the last question?’
>
> T1: He said, ‘do you want me to do it?’ And I said it might not be a bad idea. Yes, because he just hadn't looked at the back page of his ISA. I think, errm...
>
> Interviewer: [*prompting*] And again, it's about finishing the job isn't it?
>
> T1:’It would be a good idea, because you've already put all this effort in? It would be a shame to lose out.’

As the conversation progressed, T1 would occasionally ask for possible responses rather than continue to modify them herself. This therefore resembled a mentoring relationship, or interaction, rather than a coaching one:

> Interviewer: This was to Charlotte [*pointing to quote*]. ‘that's beautiful, I like your speed’, what was she doing? I just wondered what it was that was beautiful?
>
> T1: What she done, was that she'd made notes, she’d drawn in the speed, 30, 40 miles, she'd done it all in colour, so it was beautiful. It was kind of like, I like your notes.
>
> Interviewer: Okay. So, the subtext of that was you saying that you admire beauty?
>
> T1: Exactly! It looks pretty! [*said with irony*]
>
> Interviewer: You admire a perfectly presented book. You admire the quality of the appearance of the product rather than the learning itself. [*also said with irony*]
>
> T1: Yep, yeah. [*both laughing*]
>
> Interviewer: Ok, so what could you shift that into? Because she’d gone extra there, hadn't she?
>
> T1: So Errm, ‘you've copied all the notes down, and now were going to use that. How are we going to use that in this question?’

> Interviewer: You could, you could…
>
> T1: What were you thinking? You've obviously got something in mind!
>
> Interviewer: I'm supposed to be coaching! I'm not supposed to be telling you the answers.
>
> T1: Errm [*T1 pauses*]
>
> Interviewer: Do you want an example?
>
> T1: Yeah.
>
> Interviewer: I would be wanting to play it; in going above and beyond in what she's actually done, you could recognise the effort: ‘I really like the way you took the trouble to lay that out, to make it clear to yourself.’
>
> T1: Yes, yes
>
> Interviewer: And actually, that's a double whammy, because you're saying, look that was good effort, but also look how you've processed it further down the line.
>
> [*T1 Writing*]
>
> Interviewer: Maybe that's something that you could practise saying? ‘I really like the way you had taken the trouble to do… this... To make it better for your learning.’ And you're saying this as an educator. It might look beautiful, but always bring it back to the learning. That's where you're making your progress.
>
> T1: Yeah, I see what you mean. Rather than just a statement. Because you're right, what I meant was that she done a good job of copying all the information down.

Here, the recommendation was being made that T1 construct some useful phrases that could be practised and deployed during the lessons as recognition of effort. This led into a discussion of the dialogue being context-dependent; a debate lesson on wind-turbines might necessitate different language than a lesson focused on application of calculations, as this lesson had been. When Zara was questioned on showing her working, there were universal groans from the students. The suggestion was made that T1 might consider proximate praise of someone who was showing the steps of the calculations, as well as recognition of the effort of someone who was clearly showing their thinking during their working out.

As the study progressed, T1 showed some development in flexing her verbal interactions, but still continued to use positive, non-specific praise to the students:

> Interviewer: OK, when Zach said the equation: half x mass x v squared [*you said*] brilliant answer! Can you see that that is the non-specific-?
>
> T1: Yes, yes, I could have said, how did you come up with that? I looked it up in a book, he would have said, and I could have said, good, that's a good place to look for it, because you're modelling the fact that it's a good thing to do.
>
> Interviewer: Yes, good effort for bothering to do that. Do you see?
>
> T1: Yes, yes I do.
>
> Interviewer: How do you find v from v squared? I think this was Molly, and she said really hesitantly, take the square root, [*you said*] well done.
>
> T1: Yeah, so the same again. Erm, [*pause thinking*] I don't know what to say there.
>
> Interviewer: 'Well done for having a go, because you're spot on – '
>
> T1: Because she was hesitant, yes, recognising that…Right, that's the bit…in hindsight, that's much easier to see, yes, she was hesitant.
>
> Interviewer: Was this Bridget here, and here? She did something didn't she?
>
> T1: Yes, she did the proportional [*sic*]
>
> Interviewer: You said, excellent, well done, that's good. And then she said, 'and I did it all by myself as well', and you nodded and smiled at her.
>
> T1: Hmm [*agreeing, but pulling face*]
>
> Interviewer: What could you have turned that into for Bridget?
>
> T1: 'I'm really proud that you did that on your own, and you tried hard, and you really persevered with that, that's excellent [*laughing, seeing the difference*]
>
> Interviewer: And Bridget is one of our target ones

By this point, T1 was more confident in suggesting alternatives during the coaching, though occasionally still requiring some answers to be given. She was increasingly able to see the impact of her language choices on the students, which allied with her deep knowledge and sense of them as learners:

> Interviewer: So, the last one I wrote down here, our friends at the front; 'I'm not seeing much work from you'. This is perhaps more about positive framing, but you were utterly right to challenge them.
>
> T1: Yeah. And the thing is with Billy, it's funny, because he always gets stuff done quickly, he needs challenging, but he won't accept the challenge, he's lazy. He wants to do the bare minimum, but is actually bored, because he can do it, so he's sort of going through the motions rather than doing it properly. So, I want to kind of get him to move on…
>
> Interviewer: So, then you talk to him about your expectations –
>
> T1: Yes
>
> Interviewer: And that he needs to challenge himself, because he is capable of that challenge.
>
> T1: Yeah – so that one [*pause, thinking*] when you're focusing, you're working really hard today. Or if they're not, yes, I need more effort from you. You're not putting the effort in; you're not doing the work – no that's personal: I need to see the effort now.

The point was made that often students will think of work as the product or the amount of writing in their book, so if there is constant reference to the word effort, then this will assert what T1 values:

> T1: Yes, if they're hesitant; well done for giving it a go. I'm really proud of you, you've worked hard.
>
> Interviewer: it's the positive framing and the *relentless* emphasis on effort and sticking with things.
>
> T1: I like that [*writing*] emphasis on effort and resilience. I need to keep a little quote on my desk. I need to keep that in my head because I don't think I've quite shifted over. Whilst my gut is…that is what I value, it's not what I value if you see what I mean when my language doesn't show that's what I value.

Here, T1 reflects that her intentions do not always correspond with the way her interactions are verbalised, and that she recognises that her students may interpret her language to indicate that instead she merits busy-work rather than learning-work.

By the end of the study, T1 demonstrates that she has learned sufficiently from this intervention that she is able to flex her interactions with other classes:

> T1: So I sat down and had a chat with one of my Year 12s and said, look you tried really hard with this, and it might not pay off, the exam might not pay off, what do you want to do, I really appreciate the fact that you've moved yourself forward with this, you've had a go, you've found it difficult, and we talked about that, and I was saying that the thing you struggle with is this, and that we can work on it like this, and he was like, oh yeah, that's really made me feel better. And personally, I think it boosted him to know 'look a t where you've come from at the beginning of year 12',
>
> Interviewer: That's really interesting. You gave him feedback relating to the hard work that he'd put into it, but because he was working hard and it wasn't really working for him, you turned it into strategies that he could try further. And that was the calibrated feedback that he needed at that time.
>
> T1: Yeah
>
> Interviewer: Because to be told you're working really hard when you're failing…
>
> T1: Yes, that's really not going to help you. 'You're trying really hard, but let's face it, you're not going to get there' [*makes face*].

This was a confident episode in which T1 was very clear about not only how she had shifted or calibrated the feedback to the student's learning need, but also resulted in increased confidence and self-efficacy. This theme continued as she spoke about the current Year 10 class:

> T1: So I say [*things like*] I'm really proud of the way you tried that, or I really appreciate the approach you've used there, or you've worked really hard to try to complete this question, and you've got most of the way through, and this is where you need to go next, and the bit you missed out was –
>
> Interviewer: And when you have done that, what has their response been?
>
> T1: [*pause*] Not greatly different. I think that if there's the time, they're quite good at having a conversation with me, I think on a one-to-one basis, they're quite good at being able to talk to me, and say OK, I can see how I would do that, or yeah, I find that hard. But they need time to come out with that. So I spent quite a lot of time with Lorna, and she was

> still, I don't know how to do that, and I was like, well, have a look at this, you have a go, and when I came back a few minutes later she was still saying, I don't know how to do this. And I said, well you've only just started here, and I did think that it gave her a boost. You know, they sit up a bit higher, don't they, and then they just give it a go next time. Molly's doing a lot more of that actually, I think she's gained a lot of confidence over the year. It would be interesting to see whether her survey bears that out actually.

This marks a contrast to the way in which T1 required coaxing through the creation of different feedback earlier in the study. Here, she identifies that she is able to give effort and process level feedback in an affirmative way that can result in students 'sitting a little higher', gaining in confidence and 'having a go'. T1 admits to 'working very hard in the classroom', often feeling the most hard-working person in the room, and this was borne out by the constant monitoring and support offered to the students. This last passage seems to mark a shift from feedback as a *gift* into feedback as a *dialogue*, and also underscores how readily the students engage with her about their learning process. T1 did however acknowledge that she continued to bestow her undifferentiated 'back-slapping', and that she would continue to try to give this more specificity to learning and self-regulation. Nevertheless, T1 was able to suggest her own modification:

> Interviewer: So, it's moving away from the generic praise, and you feel that you have shifted in that?
>
> T1: Somewhat, and that's what I think I need to work on more.
>
> Interviewer: Do you hear yourself say -?
>
> T1: I hear myself say 'well done' each time yes, and I hear myself say that generic phrase, and Ugh! [*sighs*] And I don't know what to say. I hear myself say it and know I should be saying something else, but I don't know what to say, so part of it is about analysing that, and then thinking, what could I have said there?
>
> Interviewer: It does take time
>
> T1: Ask them maybe? Maybe I should ask them? 'I've said well done, what would help you instead of well done? What do you think I've said well done about?'

> Interviewer: Why not? That's really powerful. When they're saying to hear what they need to hear, or what they need help with, that would help you with the targeting.
>
> T1: Yes. I might try that. Especially with that lot, they are quite good. They're quite good at giving you feedback. A lot of them have taken these exams really seriously and done really well. Zara, bless her, has worked really hard. I'm very impressed, because she's always been good, but I'm really glad she got her [high mark] that made her day.

As noted in Table 5.1, at a total for 2.3% overall for Self/Praise feedback (from feedback types only), T1 was much higher than the other teachers. Of the total interactions recorded, 15.3% were coded as 'Other' and much of this could be characterised as 'banter' with the students. To the observer, both the praise and the repartee were large contributors to the low-risk, purposeful learning climate that T1 had created. Although T1 shows some exasperation here in saying she hears herself say 'well done' each time, the arc of these discussions highlights a recognition of and reflection upon practice, resulting in challenge and change to that practice, and the teacher's own reflections are examined in more detail in the next section.

7.3.4 Teacher 1 self-reflections; ceding control, developing independence and self-belief

T1 showed good self-awareness of both her qualities and abilities as a teacher; during observations she worked tirelessly to support the students' learning. She acknowledged that after a number of years on the classroom, she considered herself to both quite 'set in her ways' and very controlling of the classroom. In attempting to challenge and change her own practice, she felt some trepidation about whether she would be able to, and was frank about exposing vulnerability:

> T1: The investment in my classroom is immense. You know, it's kind of always been a bit of a kind of cautionary tale back in my mind but trying to let go is quite difficult.
>
> Interviewer: Well that will be interesting to map your own response.

> T1: I'm petrified. I'm hoping I can tell you that. I can see this big mountain in front of me because I feel very set and I know that I'll be a better teacher if I, you know, these are things I need to change and, and that's where my growth mindset comes in. I always think you can, but like at the same time, there's this part of me that's like, well, that's who I am. That's what I do. So yes.
>
> Interviewer: That's really honest of you.
>
> T1: Yeah. I need to try. And it's small steps. It is. Yeah, it is. It's about noticing it as well, isn't it? [*initial interview*]

This demonstrated an openness to change to 'become a better teacher' and despite the vulnerability, did not indicate helplessness. This demonstrated strength in revealing a state she described as 'petrified'. She mitigated the risk of opening herself up to exposure in her own classroom by self-reassurance; 'I need to try. And it's small steps. It is.'

In the first coaching conversation, T1 showed a deep knowledge of her students, and could indicate where they encountered difficulties (section 7.3.2 above). In shifting the conversation to T1's role in remediating these issues, the interviewer sought to explore T1's understanding of herself as an enabling practitioner, and here T1 demonstrated that she had learned from the intervention CPD in thinking beyond the Task-level of feedback and learning in focusing on Process:

> T1: I think partly it's about showing them the process rather than the context, so fixing on the process within the context more. So, I did some revision with a girl yesterday which was much more on the process and then said try that here, which was brilliant, and it worked perfectly on a one-to-one basis. And that's it. It's fine to do that, but it's much harder to do it on a class- basis, because they all need different pointers, don't they? So, it's more about pulling the process out, and that's what I'm not quite doing. I'm doing you need to do this to do that problem, instead of you identify the process in order to do this question.

T1 recognised that on an individual basis, she is able to develop the cognitive aspects of learning by focusing on the processes or strategies involved rather than the physics context, but that this is more challenging to do for the whole class, assuming their different starting-points. She demonstrates that she is aware of the difference to student learning, but acknowledged that 'she's not quite doing this' since she is *telling* the students how to proceed on a problem, rather than training them to become aware of and identify what processes are involved to take ownership of their learning themselves. This is a transition from the start of the study.

> Interviewer: So how could you turn it around so that they are owning the process more? What might you have to do to enable them to do that?
>
> T1: Get them to explain it to me, more than me explaining it to them.
>
> Interviewer: What sort of level of explanation do you think you provide?
>
> T1: I think I provide quite a good explanation [*sic*]. When they're struggling, I try to get them to explain it to me, but I think I jump in too quickly. I finish their sentences, I jump in too quickly to help them, I given the answer too quickly. Therefore, that will be hindering their ability to process. And partly because, if I ever try to sit back, they say why aren't you giving us the answer?'
>
> Interviewer: Is that because you've created the expectation...
>
> T1: *I've* created the expectation; I create in them a need... I create in them a feeling of they can always get the answer from me.

In accepting that she not only tells the students (too much), T1 conceded that that she also moves too quickly to rescue them from struggling. By doing this, she hijacks the cognitive dissonance phase of learning, and recognises that this action will be detrimental to them developing the ability to not only process for themselves, but ultimately self-identify and employ a suitable strategy for themselves. Furthermore, it highlights a student expectation that T1 will ultimately tell them the answers, which they ask for in anticipation rather than solving the problem. She is *creating* dependency in her students, an aspect of her practice which she wanted to change:

> Interviewer: Do you think they depend on you?
>
> T1: I think they do depend on me a bit. I think they get better and I think there are positives to the way... I know my issue is the lack of independence I give them. I am too controlling of my classroom, of my job, of the things that they do and the way that they do it. Therefore, they lack their independence. What I do think I prefer; I think I generate over the two years a group that are quite confident in themselves who will start to become more independent. But what I don't generate in the beginning is very independent learners. That's something that I need to

> work on, and I've been working on it for years [*laughing*] and failing miserably. But yes, I think I jump it too quickly and give them the answer.
>
> Interviewer: So, how could you in create situations where that independence is brought to the fore?
>
> T1: Possibly by pointing out where *I've* done it wrong...and sort of say, well you could've done it this way, have you thought about trying it that way? You know that sort of thing might help. I think yes, just drawing my attention to it, and this is the problem I'm having with this, when I do a lesson I am on automatic pilot, I’ve been doing it long enough, I just automatic pilot it and that's why changing a habit is so hard to do.

T1 makes several points in this passage. Firstly, she believes that the dependency of the students lessens over time as her supportive teacher inputs generates confidence in themselves and their learning. Secondly, she recognises that her excessive control of the classroom learning environment is a contributing factor to the lack of independence, certainly at the start of the course. A third element is that she has known about this aspect of her teaching for some time and has been unable to change it. Fourthly, she considers ways in which she might be able to break this habit, whilst acknowledging this automaticity in teaching and difficulty in breaking teacher habits. Having watched an in-school video recording of a lesson, T1 went on to reflect on both change to classroom behaviours and recognition *in* the teaching-instance of the emerging conditions of the classroom:

> T1: When you first start teaching, you change and develop so much, because you're having to think about it whereas I, when I watch that Iris video of me I was like, oh my God, so that's what I do in the classroom. That isn't at all how I felt the lesson was going, it was going much better according to the video then I felt it was going. I felt there was more gaps in it, I thought it didn't flow, I thought kids were sitting around doing nothing for periods of time. Whereas when I watched it, I thought oh no, I actually am all over the place, I noticed I was all over the place.
>
> Interviewer: Do you think you work hard in the classroom?
>
> T1: I work *very* hard in the classroom. Far harder than most of them. And there’s that element of I jump in too quickly.

This informed a discussion about how T1 might, in terms of challenge and developing incremental mindsets, start to let the students take more time on problems to enable them to experience the cognitive conflict longer, and start to think of ways in which they might be able to help themselves more:

> Interviewer: So, if we think what I've spoken with you about before, that the best learners are internal unstable, so that 'it's up to me and that I can do something about it'. What we need you to do is to create those situations where they can feel that.
>
> T1: Yeah. That's what I'm not doing at the moment.
>
> Interviewer: Let's look at some of the things you've said in this last lesson observation. So here at 15 minutes when you started going through things on the board, I wrote a question here to myself 'what would have happened if you had got some of them to explain it to the class?'
>
> T1: When we were going through the sheet?
>
> Interviewer: Yes. What would have happened, do you think? Pros and cons?
>
> T1: Okay pros because I think I asked them to explain it but then they do it verbally, I don't put it up on the board visually because I think the visual is really important, because if someone explained it to me I don't always understand it. I think the pros are that a) they would get into it a bit more, because it's them that is doing it, and that it would allow some of them to show the class that they know what they're doing. It would get them into the habit of explaining things to other people, because it's easy to do but not so easy to explain. So, I think there's a lot of pros... The cons, the drop-in pace of the lesson, and that's why I don't do it and the lack of control. And the fact that something that should take a few minutes takes forever, and you also get errors going up, and the kids copy things down.... and that's why I tend to avoid it. But I don't think this is the biggest problem; the temptation is 'just get it done'. It's the *end* of the process isn't it, when in actual fact it's *part* of the process, and I should be... the explaining it and marking of it can be just as much part of the process... And what I've treated it as is not the process, I have treated it as, right there we are, we've done that now, let's just mark it. Haven't I?

In asking the question of T1 as to what might have happened had she ceded some control to the students to empower them to answer the problems themselves on the board, T1 articulated what she considered to be both pros and cons of the approach. She summarised these with a realisation that her practice tended to take this feedback route so that there was not a decline in lesson pace, lack of teacher control and minimising lesson time on the task. However, this culminated in recognition of allowing them to provide answers at the board themselves as part of the learning process, which was both valid and effective learning and teaching. Here, T1 was realising that her decision to force the pace of the lesson was detrimental to developing the cognitive independence and self-belief of her learners:

> Interviewer: It's up to you, to make a judgment decision about whether the drop in pace is outweighed by the quality of learning they display and the way they are modelling process making, to themselves and to others
>
> T1: To others yes [*agreeing*]. Yeah thought so actually going through the work and getting them to explain the work, yeah, ok.
>
> Interviewer: Could you create the expectation that this is what it's like going into year 11, this is the journey that they need to be taking?

In summarising the themes of the coaching conversation, T1 discussed next steps in terms of how she might prepare to flex the feedback, keep the focus on effort and process, and declared the intention of listening to the recordings herself to judge how she was doing:

> Interviewer: So, for next steps going forward, I see that you scribbled some things down for you. What are you going to go on to do?
>
> T1: So, I need to look at increasing the identification of skills, and giving them the opportunity to show those skills, so like, using some of those and getting them to feedback and explain how they think they did, rather than just go through it. I need to put in opportunities to do that, and I need to think about, probably have this out [*indicates prompt sheet*], I mean to have this out, and think possibly, almost pre-visualise using these in lessons, where I might get them in, and think about what they, moving it forward into processing and effort, and self-belief and the recognition of skills.
>
> Interviewer: And positive framing?
>
> T1: Yes, and positive framing, and not making it personal, depersonalising, okay?
>
> Interviewer: I think that's a really good suggestion.
>
> T1: Maybe I need to listen to the recordings and see if I can spot some like this [*tapping page on 'how are you doing girls?'*]. Where I can pull out something slightly different... 'how are you doing girls?' What I meant

> here was ‘are you doing ok? Do you need help?’ Make it kind of more of a ‘explain to me *what* you're doing,’ or ‘explain the *way* you're doing it’, ‘you seem to be working very hard, well done’, ‘Explain to me what you're doing.’
>
> Interviewer: It's a generic question that could actually cover an awful lot, so this ties into really being quite precise, quite calibrated about what your feedback or your interrogation says.
>
> T1: So, it’s using more precise…I’m very generic, I’m very...just standard throw-outs, it’s about being more precise with the language that I am using to try to see if that makes a difference to how they’re progressing. Hmm, I think I can see the benefits, and I can see how it’s good. And it’s…you get through the door, and it goes out of the window because you just go onto automatic pilot.

In this passage, T1 has additionally identified how a generic phrase such as ‘how are you doing girls?’ can be interpreted in different ways; *are* you doing it? (Task), *how* are you doing it (Process), are you *able* to do it (Regulation), and that if there is greater specificity in the way she expresses the question, she will be able to focus their thinking in the preferred direction.

Two months later in the next coaching conversation, T1 is planning lesson episodes in which the students model their thinking, such as after an open choice differentiated homework:

> T1: They’re all going to come back with different questions, having done different questions probably, so my thinking is that at the beginning of the lesson, we’ll go through, get each of them, one of them to come up and put the equations on the board…they could model through the answers. I suppose ask questions like ‘what were you thinking about there? Or ‘how did you know that that was what you were supposed to do?’, so we can be talking about this sort of thing, ‘what else could you put?’
>
> Interviewer: Yes, good idea
>
> T1: Or, how could you make this a better answer? What was your thinking while you were doing it? How did you reason it?

T1 continued to muse on this, realising that additionally, finding out which of the questions the students had chosen to do could inform her about their willingness to engage with challenge, the instruction to do any six out of nine questions for homework. She continued to phrase challenge as 'hard' and 'difficult' however:

> T1: So, a couple of the questions are a bit more difficult, harder questions, and there'll be kids who won't have done the harder questions…we'll see. It could be that all of them will have chosen to do the first six and no one did the last three, which will be interesting
>
> Interviewer: Yes, that would be an interesting exercise in itself?
>
> T1: Yes, noting who does what.
>
> Interviewer: Who goes for a challenge, and who chooses not to.
>
> T1: [*writing*] What I can do is have a run round, because I'll have to check whether they've done their homework anyway, so as I go around, I can check and say which ones did you do? It will also be interesting [to see] who will do all of them. Gina will do all of them. I think Zara will do all of them.
>
> Interviewer: Actually, that's quite an interesting message to send to them.
>
> T1: 'I'm interested in which ones you did' -and how hard you pushed yourself.

T1 appeared to be taking the opportunity to analyse the choices of questions her students had made, whilst obliquely sending them a communication of 'I am checking to make sure you have done it, and I am checking to see whether you have made an easy or more challenging choice; I am a noticing teacher.' She is confident in her knowledge of the students, in that two of them would have completed all. Overall, this passage marked a change for T1 in ceding control to her students and enabling them to demonstrate their developing independence in physics.

On reflecting upon her use of language in classroom practice, T1 explained her thoughts about why Task-related feedback often forms the largest proportion of verbal interactions (Gan, 2011; Hattie and Masters, 2011; Bergh *et al*, 2013). She reasoned that it was the easier path, both in terms of producing feedback on the (e.g.) classroom artefact, but that by the performance goals of the school, the artefact itself could be representative of effort and engagement, and perhaps be an actual measure of it, even if the feedback has not commented explicitly on this:

> T1: There's an element of what you're expected to produce, and it's an easy way of measuring effort and measuring engagement, so yeah, my language is very work-oriented, even though I have always said that I would much rather have, and I say this to all my classes, it's a much bigger deal to have people who have tried their best than people who have coasted through and done it in ten minutes, and got it all right. I have far more time for the likes of Molly and Lorna who go, well I don't really understand that, and have pushed their way through it and asked a question than Billy who will just wallop through it and not think about it at all. But that's the way it comes out, because Billy gets the well done you've finished now move on to this. Whereas Molly gets the attention, but for something that she probably deems a much more negative thing which is I'm going over to help her finish something off.

Here, T1 expresses the dichotomy between the Task feedback she gives, which is easy to do, and what she internally values, which is her students trying and displaying self-regulatory behaviours. Hattie and Timperley (2007) asserted that task feedback became more effective when combined with self-regulation feedback, however Glover and Brown (2006) indicated that this does not often occur.

By the final interview, T1 had used some recordings, including a lesson captured by IRIS, an in-school video technology system, to reflect upon her feedback and teaching. It enabled her to track the learning behaviours of two students, one of whom was in the student focus group (Bridget), decide on a different lesson strategy to engage them more, and also consider the language she had used in speaking to them subsequently:

> T1: They were working like Trojans, they were absolutely brilliant, Bridget and Sam, and I thought I'd seen them working, and they showed me their work, and I could see what they'd done, and it looked really good, but again, I think I might have said something along the lines of, 'oh, you've done lots of work', as opposed to 'I really like the fact you've done this'. And I kind of walked away from that and went off. And then when I re-watched it, the last twenty minutes of the video, they are just mucking about, they do nothing. At all.
>
> Interviewer: So, having shown you what they were doing, they then put their feet up?
>
> T1: Basically. When I watched it again, I realised what had happened. Another teacher had come in, and they [Bridget and Sam] started a

> conversation, and then never got back to work. And I watched that bit first and didn't understand how they'd done the work. Where had the work come from? And then I watched it again and realised it was when this teacher walked in and they just went pens down and stopped. And I talked to them about it, but I realised that I was focusing more on the fact that they had stopped, because that was what had attracted my attention, and I'm saying it's really weird because you've done all of this work, and I don't see when you did it. So, I focused on the negative rather than the positive, so then I tried to go back again, and I'm not sure how I did…
>
> Interviewer: Dd you realise when you were saying this -?
>
> T1: Yeah. And I tried to turn it around and say, 'but you did so much work in that lesson, you did really, and actually I understand what disturbed you and why you put your pens down'. For me, what I needed to do was to stop them and move them onto something else. So, for next lesson, I did a different structure, because whereas previously I had set them off on something and said, get on at your own pace, and I will come around and help you, and they got on really well, and then stopped. Actually, at that point, I needed to change the focus of the lesson and do something different. So, for the next lesson we did more kind of directed questions, where I gave them something to do
>
> Interviewer: How did they respond to that?
>
> T1: OK really, better. What became apparent when you watch the video, is that Bridget got her head down, and Sam followed. Sam kept sitting up and trying to catch her attention, and then when she wasn't listening, he just went back to doing some work, so it was Bridget who was the driving force there.

This vignette indicates that T1 is open to reflection and challenging on both her language and practise through use of this video technology. From the intervention and coaching conversations, she is able to deliberate on differing strategies so that her students gain more independence, and she has tried to relinquish some control in not leaping into action to support them when they encounter problems. Additionally, this shows Bridget applying herself and refusing to be distracted from her physics work, which correlates with both her survey results showing an increase in self-concept and self-efficacy and a decrease in anxiety, as well as Bridget's own voice in Chapter 6.

The discussion then focused on how students apply themselves in different ways, and how attitude to learning may be damaged by performance goals:

T1: And it’s things like that. It’s not having a go or being afraid to have a go: I’m only going to get it down in the book if it’s perfect, because I want a perfect book. But that’s not growth mindset. No, not at all. And they really struggle with that. My sixth formers do as well with their lab books. They don’t get lab books [*meaning they don’t understand using them*], it’s got to be right. They don’t get the whole point is that you cross things out, you learn as you go through, it develops,

Interviewer: Yes

T1: I think we have a lot of fixed mindset in this school, a lot, and also what I notice is the jump between primary and secondary, when the Year 6s came in [transition event], the difference is that the Year 6s will have a go in a way that once they become secondary school, they don’t. There’s an element of hard-working, wanting to do well, that’s not ‘cool’. Charlotte’s a prime example of that; she doesn’t want to succeed because she doesn’t want to be seen as someone who’s succeeding in case that’s uncool. When you looked at the Year 6s, they were just giving things a go, they were happy to show their enthusiasm, and almost as soon as they come into Year 7, they lose that.

This then led into a discussion of student change that T1 observed as an experienced teacher:

T1: They change. They change in style, they change in their attitude to learning, it’s interesting because I’ve taught a number of students right through from Year 9 to A levels, and they really change. But you notice the change between Year 9 and Year 10, and 11 and 12, and they change personality, they change the way they work, and sometimes it’s for the better, and a lot of the time, it isn’t. They get stuck and they can’t make that change.

Interviewer: And why do you think that is having watched that so often?

T1: I think that part of it is about the level of ‘cool’, as you get into A level there is a different level of cool. In the opposite way, they struggle to make that jump because they lose confidence, then they shut down and sort of stop trying.

Interviewer: And would that shutting down be fixed?

T1: Yes, that would be fixed, they don’t want to try because ‘I might fail’.

Interviewer: So, what can you do to help them pass that hurdle?

> T1: [*pause*] give them feedback which encourages the development of their growth ideal.

Here, T1 seems to have come full circle. She recognised the different ways in which students mature, and the impact that a maturation effect has upon their attitude to learning, but also indicates that she believes that the teacher has a role in helping to manage that maturation take effect in the way the teacher uses language to the students. In closing the interview, T1 was asked what she was going to do going forward:

> T1: I think that I'd like to keep this at the front of my mind. I tell you what we've got to do. I have to do six hours of independent CPD next year. But I think that what I want to do is keep moving the feedback forward. So, keep recording, and use some of the six hours to review my recordings, and have a think about how I can move it forwards.
>
> Interviewer: That's an excellent idea.
>
> T1: Because otherwise it slips. I mean, this is at the forefront of my mind, and it's still slipping to the back of my mind when I do it, when I'm in the middle of a lesson it goes, so I do need to keep it at the forefront, and get that to shift. Maybe do more reading about it.
>
> Interviewer: Perhaps having some really tried and trusted phrases –
>
> T1: Those phrases I've got in my file [*from previous coaching conversation*] are really useful. I do use them, I try to.
>
> Interviewer: So, you're using them more?
>
> T1: Yes, I'm using them more, and I have them out on the desk a bit more.

In conclusion, whilst T1 began the study with a clear conception of what constituted effective feedback, and what she wanted her feedback to achieve, she had been unaware both of the different levels of feedback, and how students might receive these different forms. Although she found it extremely challenging to change her feedback 'style' (Schneider and Randel, 2010; Wiliam, 2010; Berg, Ros and Beijaard, 2014, 2015; Heitink *et al*., 2016), her acute knowledge of her students and the way in which she strove to provide prompts as next steps (or EBIs in their classroom vernacular) firmly contributed to her efforts to flex feedback styles into more regulation and process, despite her acknowledge retreat into default modes. Finally, her honest reflections throughout the study showed a developing change to relinquish some control, and grow the independence of her students, whose self-avowed increase in self-belief support their view of her as a cherished teacher.

7.4 Teacher 2

Teacher 2 was relatively new to teaching and was a mature converter to education. He was new in post in the school and was motivated by the nature of the research to take part. The main themes which emerged from discussion with T2 were a mapping of the successive attempts to shift feedback to become more positively framed, a teacher-led intention to increase the independent learning of his students particularly through the increased use of Process feedback, and some honest reflections on practice, although with some elements of disavowed knowledge.

In terms of the amount of oral interaction in class, T2's practice was 75.0% dialogic by the end of the study; this is the amount of feedback-related interactions as a percentage of total interactions. This rose from 65.0% at the start of the study; T2 had the highest of the group. T2 set up *chains* of interactions with students, also the highest amount of process feedback (which came back to bite him with the resistance of the students) see table 5.6

7.4.1 From personalisation to positive framing

A theme that emerged early in the study through observations and subsequent discussions with T2 was the personalised language used in interacting with the students, captured both through field notes and analysis of the lesson recordings. This raised a concern that the student perception of this might result in a way to negate the intended outcome. In the first coaching conversation, examples of these were shown to T2 and there was some consideration of how these could be more positively framed so that the 'take-home' message of the feedback could be better received:

Interviewer: So, I'm going to share some examples of things you said to the students and we'll have a think about how these might have been differently phrased.

T2: Okay.

Interviewer: 'Am I interested in your artistic abilities? No. I'm interested in your conveying the concept.'

T2: Ah, I just meant 'don't take a lot of time drawing a bicycle, we have to concentrate on the forces on the wheels.' [*pause*]. Yes, that didn't sound great, did it?

Interviewer: So, what you've just said there, can you see that that has a different feel to it? You've turned a 'you' into 'we' and an instruction to save time and focus learning?

T2: Yes, I should have said it like that in the first place [*writing*]

Interviewer: 'You haven't done this; I want this done.'

T2: Erm…This piece of work needs finishing off?

Interviewer: Yes, so again, you've taken out the 'you', and you might want to follow that up with, 'I'll be checking for it later' to ensure it gets done?

[*T2 writing try not to say 'you'*]

Interviewer: Ruth, Paul, where are your books….so why didn't you pick them up when you came in the door?

T2: where are your books…why didn't you…Erm, Ruth, Paul, the exercise books are in that box by the door?

Interviewer: Yes, good, it's then more of a friendly reminder that they *need* to get their books, and that you have spotted that they haven't yet and are onto them – without saying that.

T2: [*writing*] I see, sending them a message…

Interviewer: How about this one: 'Are you just guessing at the numbers?' How might you have phrased that differently?

T2: [*winces*] Ah. Who was that? Oh, Faith, yes, she was just guessing at the multiple choice instead of working it out. Erm…let's see which ones it can't be? Erm…what information do we already have, what do we need to know?

Interviewer: Yes, and then you've shifted it into her taking an *action* or strategy or process for her learning.

In the second interview, T2 expressed his belief that he still did not feel he was providing enough positive commentary, but that he was cognisant of the Process feedback that he was able to deploy; lessons 6-10 indicated a gradual rise in Process feedback (see Chapter 5).

> Interviewer: How are you have you been doing with the feedback in the lessons?
>
> T2: Good, I think. I still don't think I'm doing enough of the positive, the encouraging sort of feedback, the right kind of feedback. I'm still working trying to inject it.
>
> Interviewer: So, the next steps as we left it was to do more positive framing in terms of the sort of things that you have been saying to them, and to do more of it?
>
> T2: it helps that I have the prompts in strategic places, but I need to make them a lot larger.
>
> Interviewer: Do you find that you can see them? Do you have time to look at them?
>
> T2: Yes, I just need to get those reprinted so that they are staring me in the face a bit more.
>
> Interviewer: And might it be a case of practising some of them? As cheesy as it sounds, that you hear yourself saying it.
>
> T2: no, I think that's a very good idea, because I want to have a library of stock phrases, and you want to change the stock phrases that you previously had.

Here, Teacher 2 recognised that he has to replace what he terms his 'stock phrases' with new terms that reflect process and regulatory ideals, and that these routinised phrases will need dislodging (Eraut, 2004). He acknowledged that he 'had got very used to saying things' in a way that became traditional to him, and that 'to start saying something else, or to put a spin on it, you have to put the practice into that just as you would want them [*the students*] to practise something.

> Interviewer: Yes, So, if there are particular things that you want them to do in a processing sense, or a strategy sense, the way in which they are approaching it, then you practise those sorts of feedback prompts and

statements. Whereas if you're wanting them to stick at something, they're the sort of phrases that come a lot harder, because we don't tend to say it anyway near as much.

T2: So, phrases for sticking at it…Erm, you're on the right track, keep at it, keep going,

Interviewer: Yes, yes

T2: I probably do say more of those…but I might not…you're the one with the recordings, but I do feel that I say more of those than the positive framing, the depersonalised positive framing stuff, hmm. Although there's the question of catching them in the middle of something, rather than catching them at the end of having done a good job of something.

Unlike Teacher 1, T2 did not appear to have listened to the audio-recording of his lessons to reflect upon his use of feedback. Remaining with the theme of shifting oral interactions into more positively framed discourse, T2 expressed some frustration in the lack of application of the students in what he described as a 'rather verbose class':

T2: Maybe its summer term, I don't know. There's a lot of cat herding going on; the lesson I recorded this morning, the amount of times I heard myself saying, I asked you to do this once, I've asked you to do this twice, three times now. Why aren't you getting on with this?

Interviewer: And how could you shift that sort of thing?

T2: But when a kid has not started the stuff that they had been asked to do, I mean, it was just about everybody towards the end of the lesson. I stopped all of them and said, this is what you were asked to do. No one seems to have done it, and yet you all seem to think that you're done. Where do you get that impression, that I've asked you to do A, B and C, and when you guess at C without doing A and B, you've not done the work that I've asked you to do.

Interviewer: OK

T2: So, I'm not sure how to positively frame that. [*showing a need for mentoring rather than coaching here*]

Interviewer: It's interesting, isn't it, because despite being this time of year, they're clearly making a decision not to do it?

T2: Hmm.

Interviewer: Which means they're taking a step back from the personal responsibility of doing it.

T2: Yes.

Interviewer: So, can you put it in those terms? It is about what's coming up, but it's also about good habits, isn't it? It's good habits of learning.

T2: Hmm.

Interviewer: And I think that perhaps instead of saying that 'I have asked you and asked you', you might say 'my room is a place where we get on with the job. Because the job is really important'. And then the emphasis is on that work, rather than you, I asked you to do it. Put the emphasis on, back to the effort.

T2: Right yeah okay. [*T2 writing a note*]

Interviewer: I think that it is tricky to do, on the spur of the moment, I think some of these are things that you might have to practise saying to yourself, and that's a bit difficult. But I think it is always about trying to find a way to phrase it so that it depersonalises it, so that it puts the focus on the job in hand, rather than the person that's doing that job.

Yes [*saying/writing down depersonalising, job in hand not person*]

In this section, T2 still expresses a lack of confidence in knowing what to say when opportunities for depersonalising the oral interaction arose, as well as flexing the nature of the feedback to reassure students that it is 'alright' not to know something in physics, and finding out was part of the learning journey. He also needed reassurance about what to say to the students, as the relationship here shifted more into a mentoring conversation ('telling, guiding') rather than a coaching one ('questioning to support developing own route'). Consistent with other studies indicating even committed teachers struggle to integrate training material into their classroom approach (Guskey, 2002a, 2000b; Kwakman, 2003; Darling-Hammond *et al*., 2017, 2019; Schütze *et al*., 2018), T2 knew *theoretically* what the purpose of flexing the feedback was to achieve, but in practice, this was not occurring:

Interviewer: And for those who maybe feel that they didn't have a very successful lesson because they didn't know, to say, it's alright not to know –

T2: Hmm [*agreeing*]

Interviewer: That this is about finding out. And there's lots of times in science when we don't know, so you shouldn't feel silly when you don't know because this is all part of us finding out together. And then you give that reassurance that they're not stupid, because some might self-label to that extent.

T2: Yes

Interviewer: So, again, it's bringing it back to a safe place isn't it, where it's not about them, it's about the work.

T2: Hmm Hmm [*agreeing*]

Interviewer: So, what else could you do, do you think?

T2: [*Pause*] I don't really know really

There was for Teacher 2 here a sense that he knew what he had to do but was unsure about how he was going to get there. That this mirrored the purpose of prompt and process feedback was somewhat of an irony, and the learning need that T2 demonstrated at this point was indeed more indicative of a development process requiring mentoring rather than coaching at this point. He did realise that there was a need for rehearsing the language outside of the classroom (as T1 had done) to gain more familiarity with certain phrases:

Interviewer: What are you going to think about for next steps?

T2: Practise the language,

Interviewer: Yes?

T2: More visuals around the class to remind me to use the language. Bigger ones. Well that's pretty much what it is because I've got those ones on the side of the filing cabinet, but you can't exactly go [*mimes bending down and squinting at text*] just let me look at that before I feedback to you – how did that make you feel?

Interviewer: If I come back in again and listen to you –

T2: Of course.

Interviewer: - and then we'll look at the language and consider how you might flex it, should that be needed. But this isn't going to be an instantaneous thing at all

T2: No

Interviewer: And that is why this is running over quite a bit of time because you need time to get your head into changing the habits you've acquired, just as we would maybe want to shift some of their habits, about not learning in the right way, or not approaching learning in the right way

T2: Hmm hmm

> Interviewer: But they're not going to do that without the prompting language from their educators as well.
>
> [*T2 agrees*]

The next steps for T2 were therefore more directed at this point in time than for T1. Consistent with support for impact suggestions from Guskey (2002a, b), the interviewer reminded T2 of a prompt sheet that had been shared previously, and suggested also that he listen to a past recording, and what was said in response to the students, and consider how could that have been rephrased? In addition, he was asked to reflect upon how quickly he responded to the students; did he take the thinking time to consider how to flex the feedback so that it was either positively framed, or instead of a comment on the work could they be prompted to talk to him about how they had processed something?

By the third coaching conversation, T2 had increased the use of process feedback; lessons 16-18 demonstrate a peak (Graph 5.13). In lesson 19 in which the context was plotting radioactivity decay, T2 demonstrated encouraging and prompting responses to the class regarding 'not knowing', and the nature of science. He discussed not being too swift to make assumptions about outliers in their data collections, since this area was 'new' science for them, and they would not yet know what valid and reliable data might look like. Here, T2 as a knowledgeable scientist talked about the nature of working scientifically to contextualise the science content:

Once you've finished…this is a challenge for the early finishers	RpC
Tell me what the half-life of protactinium is	PpC
You can get it from the graph, but I'm not going to tell you how	PpC
Put your heads together and see if you can come up with a way to work out the half-life of protactinium	PpC
It's fine if you don't get it, it's a real challenge, right?	RsC

(Lesson 19, 6.22-6.46)

Overall, Teacher 2 seemed to experience inertial difficulty in changing the tone of his responses to more positively-phrased reactions, although over time he was able to change this, in part due to his method of not evaluating answers immediately as IRF (Sinclair and Coulthard, 1975, cited in Howe and Abedin, 2013). This momentary pause upon listening to a student enabled him to begin to adjust the way in which he framed his answers. Often, he would say 'hmmm' in a thinking tone, occasionally because the student response was incorrect, but more often to prompt further talk. This linked well to Teachers 2's primary stated goal of wanting to improve the students' own perceptions of themselves as independent learners, which is discussed in section 7.4.2 below.

7.4.2 From learner reliance to learner independence

As explored above, T2 espoused a belief that whilst it was certainly 'easier' as a teacher to employ a more spoon-feeding approach to his students' learning in order to demonstrate their progress, it was against his principles to do so. His motivation lay in wanting to enable his students to make decisions for themselves in a future where their teacher would not be available to tell them what to do [field notes]. He considered that he built challenge into his lessons to encourage them in this attitude, but did recognise that his language, and the type of feedback that he provided was instrumental in facilitating this change. In the second conversation, T2 described the dependency that the students displayed alongside the expectation that he would solve their problems for them when encountered:

> Interviewer: So, can you tell me what they were doing this morning?
>
> T2: Well, it's still the electricity unit, so they were looking at potential dividers, which sound frightening, but there just two resistors that split the potential difference into two parts. You play around with them and start to realise that the ratio of resistors is something to do with the voltage you read across each. Right at the end, a couple of the stronger students [*sic*] were starting to get a notion of the use. One of them said, at the very end, 'oh, you could use this for a light sensor'. And this was great, but I was hoping that the whole *class* would be there at the end of the lesson. They're not going through all the steps the way I would like them to, they are complaining: 'oh, this is the same thing we did the last time'. They are not quite seeing that this is repeated takes of the same notion until the *concept* is engrained.

Here, the students see practice as repetitive work rather than mastery, as they seem unable to extend this surface learning to mastery of the concept of ratios of voltage in a potential divider device. Their teacher is still having to guide their thinking as they are not able to 'play' (as T2 suggests here) with the circuit to explore the potential difference relationships within the circuit, either on paper, or by practical experimentation:

> Interviewer: Hopefully, more of them will start to see that it's those rules about potential difference that come in and underpin it. Were they doing this in a practical?
>
> T2: Yes, yes, they were. Yes, it's still doing that discover through making the circuit and testing, that kind of process, which I do think is helping them. You know at the beginning, I posed this as a question, took suggestions, then worked it out as a class. The first suggestion wasn't as good, we couldn't make this work, but the second one was better because of these reasons [*modelling*]. And that suggestion was pretty much a potential divider, so I scaffolded that into experimentation with them as a whole class sort of thing.
>
> Interviewer: So, the difficulties they're encountering, what sort of difficulties are these?
>
> T2: Well, there's been technical difficulties in the past, just getting the equipment to work properly, and making sure the circuits read correctly. But that's not any student's fault, because they don't have the background theoretical concepts to tell when the equipment isn't reading properly.
>
> Interviewer: So, they don't know how to troubleshoot?
>
> T2: Yes, exactly. So, if they get an erroneous reading on a circuit, they aren't instantly aware of it, so they can easily write down the wrong thing. So, they can have two different resistances and write down the same potential difference across each, when it should be different, and not realise, until I come around and ask oh, why is that the same? Why is that the same when you've got a big resistor and a little resistor?
>
> Interviewer: So, they're not knowing enough to know the questions to ask, and/or they're not knowing what they don't know.
>
> T2: Yes that. They don't have enough experience yet to spot when something has gone awry. It's the unknown unknowns in this case.

The students' lack of working knowledge of circuit behaviour was critical here in providing a barrier to their learning about the potential divider itself. This was partly due to not knowing the current and potential difference ‘rules’ in series and parallel circuits sufficiently, to be able to build upon these for a more complex circuit, but also in their lower self-belief in physics, and confidence in exploring and deducing circuit relationships when the technical hitches encountered produced additional barriers. Field notes indicated here that when the circuits failed to work, few students seemed able to systematically test leads (for example) to ensure electrical connection, described by T2 as ‘unknown unknowns’. In discussing this unwillingness to learn from experimentation and (managed) failure, the conversation turned to what T2 might offer in terms of teacher feedback responses to recognising when effort is being expended, even if unsuccessful, and what other strategies they might try instead:

> Interviewer: So, if you think about what you would want your feedback to do *there*, because where they don’t know, and perhaps where they know that they don’t know, they feel a bit more needy maybe?
>
> T2: Yes, I should say things like ‘how could you check, if that’s correct?’ because they do have the maths. They’re resistant to applying it but they do have the maths to actually say, ‘hey we could work this out using V=IR, using Ohm’s Law, and that isn’t what I expect. So, something must be wrong, what’s gone wrong?’ [*writing to self, how could you check results?]*
>
> Interviewer: What else might you prompt them with?
>
> T2: Hmmm. What sort of things might you check for? How would you know? What questions could you ask of the data? [*writing these down as he speaks*]

It was hoped that in considering and rehearsing these prompting phrases *outside* of the classroom demands, T2 would be able to draw upon them inside the classroom. In discussing the learning behaviour exhibited by students during practical activities (Schütze *et al*., 2018).

7.4.3 Modelling the language to promote growth mindset and independence

T2 was asked whether the students that showed more liking for science benefited more from a practical approach than other students:

> T2: No, I find that they get *less* out of the practicals.
>
> Interviewer: Why do you think that is?
>
> T2: Because for some of them, they would just rather be working it out on paper, I think. They don't seem to like the process of physically setting up the circuits, or physically setting up the equipment to gather the readings. They would just rather have 'now here's a set of data, now do a graph from it, now tell me an answer from it'. The nuts and bolts of setting up something that doesn't always work perfectly, it doesn't follow the theoretical prediction exactly, but its close enough… If you've got the science experience, you go OK, you've got a voltage of 5.6; it's supposed to be 6 volts, its close enough. For them, they want it to *be* 6 volts, because that's the number they *expect*. The *law* says it's got to be that – and it's not, it's the real world. They don't like it as much because it doesn't have that perfect output.

This section demonstrated that T2 could recognise both lack of science knowledge and self-belief barriers to his students' learning. It also mirrored a view articulated by some students that they 'only wanted to put things in their books if it was right' [field notes]. The theoretical perspectives of current, potential difference and resistance laws were, in the minds of the students, being distorted by imperfect circuit realities, challenging their acceptance of such physics laws. This led into a discussion of modelling the language to promote growth mindset and independence, especially when engaged upon practical work:

> Interviewer: Perhaps make a comment about how hard they have been working at something? Now most of the students in your class profess to have quite a growth mindset attitude, but it's in the way in which they are approaching their work, or rather not, when they are being challenged by it, is far more revealing than one snapshot survey.
>
> T2: Yes, yes…I don't think they're very growth mindset [*sic*] at all at the moment, to be honest. I'm not trying to pass the buck, they're lovely kids, but so many of them want it to be right, the first time, and it should be – come on, keep figuring it out.
>
> Interviewer: Are you thinking of anyone in particular?
>
> T2: The only ones that I'm seeing growth thinking from at the moment are Dan and Chloe. She will offer up an idea, like today when I asked how is this potential divider working?
>
> Interviewer: So, she's not afraid to have a go?

T2: Not afraid to have a go, or *less* afraid to have a go.

Interviewer: That's good to hear. So maybe you could try that proximate praise, 'this is what I like to see, someone having a go', and show that it doesn't matter if you're right, and it doesn't matter if you're wrong, it's the having a go that's important.

T2: It's having an idea and putting it out there [*writing a note*] Yes. But that's about it in the class [*referring to growth mindset students*] that are actually grinding away at problems to try and solve them, as opposed to, 'oh well, the circuit isn't working, I'll just chat until Sir comes around and asks why I'm not working, kind of attitude.

Interviewer: can you see that happening as you circulate?

T2: Hmm [*confirming*]

Interviewer: So, what do you say when you come up to them? Do they just want you to fix this problem?

T2: Yeah, what do I say to them, I'd have to listen to the recordings, it's mostly 'get back to work. You shouldn't be sitting here, twiddling your thumbs and waiting for someone to notice that you're struggling, you need to ask for help. You need to be sorting it out so that the circuit does work'. That's probably not what I'm saying, that's what I hope I'm saying.

Interviewer: OK, but then, if you go back and listen to that, you can check whether the first thing you say is, 'why aren't you doing something' or whether you say, 'let's get this sorted'. Do you see what I mean?

T2: Hmm

Interviewer: Because how you approach it, is the sub-text for them. Because if it's a why haven't you done anything yet, then it's *you* not having done the task that I have set you, whereas a 'let's not let this beat us' is a very different way of phrasing it. So, it will be interesting to listen to find out what you say –

T2: Yes, I think it will probably be the former. I need to write that one down [*does so*]

Interviewer: Get practising, say it in the mirror

T2: Well I think the practising will help to be honest.

During this extended section, T2 recognises the dichotomy between a school which champions a growth mindset approach versus the actuality of personalities in his classroom. He is aware, both through the school approach, and the intervention professional development of the attributes Dweck and colleagues associate with being an incremental theorist, yet in selecting only two students who he feels embodies these in both actions and words, he is giving his opinion that the majority of the class are not, despite the pre-intervention survey outputs. This section also highlights how, in contrast to Teacher 1, T2 does not listen to the recordings he makes of his classroom dialogue to reflect upon his responses, leading the interviewer to suggest that he does, and also practises several key phrases so that he is able to call upon them in moments of classroom action.

During the final interview, using the GROW model (Whitmore, 2002) to discuss what the goals had been and what the current situation looked like, T2 again confirmed that his primary motivation was to build his students as independent learners. He showed concern when low self-concept students self-labelled as ‘bad at physics’, tried to increasingly use process and regulation feedback to modify this, and on self-reflection, noticed that he had increased his use of *process* verbal feedback:

> T2: But I’m noticing much more when I’m just giving Process feedback
>
> Interviewer: What are you wanting to achieve with feedback?
>
> T2: I think at the moment what I want to do is encourage them as much as possible to feel that they’re more capable than they realise. Even just this morning, I had a student in another Year 10 class come in, first thing she said as she walked through the door was ‘I can’t do physics’, and I said, ‘Oh, come on’, because she does say that fairly regularly and them all the way through the lesson when a problem was put in front of her, she was just knocking it out of the park. Absolutely no problem; sixes, sixes, sixes, sixes the whole time [*cricket analogy*]. And I said to her, look, you came in saying you can’t do physics, and you’re leading the class right now, you’re working really hard at this. [*T2, final interview*]

Drawing upon his own non-specialist teacher of physics background was common for T2, and despite having extremely good subject knowledge, he deliberately used this background to articulate to students that ‘you don’t have to be born a physicist’. The classroom messages were consistently modelled to the student body as ‘it’s okay to not know, as long as we try to find out’; learning from failure so that it is instructional rather than demotivational (Bandura, 1997; Darling Hammond *et al*., 2019), taking every opportunity to learn, and valuing persistence. This was clear as he continued to talk about what he wanted to achieve with the feedback he used:

> T2: So yes, just to try to get that feeling with all students across all year groups from 7 right up to 13. I mean, the way I'm mentally preparing at the moment is to channel that 'we're all in this together', the Blitz spirit kind of thing. I mean, I don't refer to it as that with the students, but that kind of 'I don't know either so let's find out'. That kind of attitude to things. It's helped me, because some of the teaching I'm doing right now it's not in my complete area of expertise in the physics world, it's things like radioisotopes for example, where I don't have a laundry list in my head of what all of them are, and what they emit, what they decay into, and what they're used for. So, if a kid goes 'what's iodine-127 used for? Well I don't know, let's find out, let's explore that; oh, it goes like this. So, have you found it does that? OK. So more of an exchange but also trying to knock down especially in the lower year groups; teacher-tell-us-the-answer attitude that comes up quite frequently.
>
> Interviewer: so, if I could sum that up, it's [promoting] a mixture of 'can-do' and independence.
>
> T2: Yes, yes, and I've realised it now, because I had 'independent learning' on the board for about two weeks whilst I was hammering it with all my lessons [*T2, final interview*].

In expressing what success would look like in his classroom, T2 conveyed a picture of students being able to undertake work for themselves and be process-focused, rather than task-focused:

> Interviewer: So, going on from that, what would success look like to you?
>
> T2: not entirely facetiously, to be able to go, 'here's what we're doing today, I'm going to talk for five minutes, right, you go and find out about this, that and the other thing, and have them just *try* it, and be asking questions like, 'have I understood this?' as opposed to, 'what is the correct answer?' do you see what I mean?
>
> Interviewer: Yes.
>
> T2: so, 'here's some stuff that I've found', or 'here's some information I've put together', here's my understanding of it; am I on the right track? As opposed to 'is it the correct answer?'
>
> Interviewer: 'Is this right?'
>
> T2: yes.

Interviewer: do you feel that they still want to know that it's right?

T2: yes, especially with the lower year groups, but that's to be expected; they've come from primary where its right answer, wrong answer.

Here, T2 is clear that he sees his teaching role as one in which he seeks to minimise teacher talk as instruction at the front (compared to T4 in the same institution), and instead use strategies to challenge the students in engaging with thinking tasks more independently. Following the earlier comments about younger students just wanting right answers all the time, T2's view on whether age might have an impact on students' ability to reflect on developing independent behaviour was sought:

Interviewer: What would be, in your opinion, the best age-group to do this with?

[*Pause*]

Interviewer: Or don't you think it matters?

T2: I don't think it's entirely matters. It's important that it be started at the earliest possible age, I think. It's not something that you just want to introduce the A-level for example, there's then they'd have five years of other feedback models and then suddenly it's suddenly all about the process. Well it *is,* even more so at A-level than it is in years 7,8, and 9, but you want them thinking about it in 7, 8, 9, 10 *and* 11, so that when they get to 12, and you're going, well you're going to have to do a few problems now.

Interviewer: So, it's intentionally building those thinking and learning skills early on, so that they've got the capability to do it when they are challenged at A-level?

T2: Yes. From parents evening, you know I'm constantly saying well you're doing well but, well it's a common theme, you get the stereotype of the clever boy who comes in and everything is easy for him, and he's got the right answer every time, and is getting good marks in the tests, and his grades are great. And then at some point there is a cliff beyond which your natural smarts just kind of run out. And you need to work hard at that point.

Interviewer: And he doesn't know what he is doing when he does it well?

T2: Exactly. The stuff that you pick up from reading New Scientist, and you are keen about the topic, well that'll just run out at some point. You encounter something you didn't expect, or you're not aware of, all you're not familiar with, and then you're going to have to work hard. And at that

> point if you don't have the work ethic, and the keep going, and trying it again, redraft that, 'is there another way that we can put it?' kind of attitude, then you really going to fall off the cliff. As opposed to going, 'this is a bit of a problem, ah! I know what to do'. The number of times I've said that at parents evening. [*T2, final interview*].

T2 is firm in his belief in advocating a process-led, metacognitive approach in which student self-awareness of thinking skills is consistently encouraged from an early age, such that a post-16 physics student would feel more enabled to choose their own approach as the complexity of the learning demand increases. He mentions 'work ethic', nonetheless this forms but part of developing student ownership *of* and independence *in* learning. This journey is one that the teacher must take *alongside* their students in introducing, modelling and practising what independent problem-solving looks like, as students are unlikely to come to this autonomously. Coe *et al*. (2014) advise against too much enthusiasm for 'discovery learning' since they claim this is not supported by research evidence and emphasise the role of the teacher in supporting students to uncover key ideas, primarily at least through direct instruction.

7.4.4 Teacher reflections on perceived failure

In reviewing the efforts to flex his feedback over the past six months, T2 was frank in his disappointment about what he perceived to be his failure to achieve this:

> Interviewer: So, what do you feel that you have done so far with your feedback so far?
>
> T2: Oh, I feel like a complete failure, because it's not become a glowing shining success all round with all of this happening already. I do realise that it has to be built up and I have to become more practised at it. I mean, I was struggling with giving the correct sort of feedback… with the process-oriented feedback I think I am doing a bit better. But then I'm not the one coding takes to see if I'm not.
>
> Interviewer: To say that you are a failure is really strong
>
> [*T2 laughs*].

> Interviewer: Really strong indeed, and not correct.
>
> T2: Well, that has to do with my own personal standards, that's the path I took to get here, I had quite a hard time in training, and its reflected in all of my practice since.
>
> Interviewer: So, what have you learned from that?
>
> T2: I'm not sure [*pausing*].

Field notes indicate that this was a slightly emotional reflection; T2 consistently expressed aspirations for his students in physics as well as his own high standards for his professional practice. Here, recalling the challenges of his own training echoed the difficulties he saw his students experiencing in their own learning, and although vocalising this in a self-deprecating way, there were overtones of quite harsh self-judgement:

> Interviewer: This challenge to change your feedback style, you found that quite a difficult journey, and it's not, from what I've seen, through any lack of *wanting* to do it. I'm just wondering why you think you haven't done it as well as you would have wished to?
>
> T2: I think that… one thinks that they will be standing at the front of the class or talking to a student, and thinking 'ah, now I say this phrase, and now I bring out that example', so because it isn't ingrained yet, or being a reflex, I have to think about it. I suppose it's like any skill, like riding a bicycle; if you're thinking about it while you're wobbling around, once you practice then you're riding with one hand and texting with the other avoiding traffic and all of these other things [*laughing*]. So, because I don't feel I've got that… Zen-like low of 'Oh, this is how I present, and this is how I feedback verbally to get the students to go in this direction', it makes me feel that I'm not there yet, so I haven't achieved what I wanted yet.
>
> Interviewer: Does it make you think that you won't, or does it make you think that it will be a longer journey than you thought it was going to be?
>
> T2: I tell myself it's just going to be a long journey, but it's hard to avoid thinking 'I'm just never going to get this, I'm just never going to get it right.'

Over the course of the study, T2 had often expressed a viewpoint that could be considered in agreement with descriptions of 'growth mindset'. His manner of articulating both situations in class and approach to managing new learning had been to thoughtfully embrace the challenge and encourage the engagement in learning. This journey over the study from hopefulness to stoic pessimism was a sobering reminder that teachers find challenge in changing their practice as outlined in professional development literature in both generalised CPD (Guskey, 2002a; Kwakman, 2003; Darling-Hammond *et al*., 2019) and feedback CPD practices (Schneider and Randel, 2010; Wiliam, 2010; Berg, Ros and Beijaard, 2014, 2015; Heitink *et al*., 2016). In continuing to discuss this as evaluating his classroom practice, the question of automaticity was raised:

> Interviewer: Is it perhaps that in the lesson, in the heat of the lesson if you like, the automaticity takes over?
>
> T2: Oh well of course it does. And any teacher will tell you, I'm going back to what I'm comfortable with, more lecturing at the front. Other people are more comfortable, I don't know, doing practical activities. Everyone's got their own style that they fall back on. I'm trying to change a certain aspect of my style, that isn't quite as natural to me from long experience, so I'm trying to build up experience with the new style, until it's more automatic [*clicks fingers*].
>
> Interviewer: Okay, so when should teachers learn about this then?
>
> T2: Well, from the beginning I suppose is the easy answer. Everything is practice.

This reference was the clearest directive that learning to flex feedback was more challenging to do when feedback practice had been established. Since effective professional development is intensive, ongoing, and connected to practice (Darling-Hammond *et al*., 2009), there is potential for investment in verbal feedback considerations to be included within Initial Teacher Education [ITE], to empower good verbal feedback to be rooted early, and prevent downstream resistance to change.

In considering this current challenge however, T2 showed determination in continuing:

> Interviewer: So, you feel that you have somewhere to go, but you're determined to carry on doing this?

T2: Oh yes, certainly.

Interviewer: Do you think it'll make a difference then, setting your stall out at the start of an academic year?

T2: Oh yes, absolutely. Well, I tell myself that it will, but I'm sure that it can't hurt to lay out to the students. I feel more confident about doing that as well, because it will be my second year in the school. Not of spending the first few lessons thinking: 'where are the books? How do I take attendance?' But rather: 'welcome to my room, this is how we do things here.' But, I mean, it is a process, so, next year I'll be doing that from the beginning, rather than starting six months in and kind of make it my thing

Interviewer: Fair enough. So, what have been the pros and cons of doing this?

T2: I can't really think of any cons. No, because it hasn't come at the cost of any scientific rigour for example, or any kind of content rigour, it's just a different way of putting ideas to a student, or a suggestion to a student of what to do next. So, in that sense, I can't think of any negatives to it. Pros, well notionally, it should spur students to become more independent.

T2 makes several interesting points here, in thoughtful feedback reflection of his own. Firstly, despite his perceived 'failures', he deems it a worthwhile endeavour and is confident that he will continue. Secondly, he felt that using this approach would be more beneficial at the start of the academic year, rather than attempting to modify practice some way into the year; 'setting his stall out' and making the verbal feedback clear to the students. Additionally, he wanted to 'own it' and make it 'his thing' in the classroom and this implies both ownership and partisanship of the approach. Thirdly, despite his own struggles, T2 did not perceive any disadvantages to managing verbal feedback to his students in this way (beyond his own challenges) in terms of teaching and learning, and indeed, fourthly, still believed it to be of value in promoting independence in learning and showed that he was emotionally invested in it.

In the interests of full disclosure at the end of the study, it was pertinent to share information that the students of T2 had disclosed in the focus group since this involved their perceptions of feedback:

Interviewer: So, I'm going to give you something that the student shared with me yesterday.

T2: Sure.

Interviewer: One interesting thing that came out of it when I asked, 'what sort of feedback does your teacher give you?' was that they talked predominantly about *written* feedback. What emerged, was that they did not see the interactions in class as feedback. When I probed them on verbal feedback, they were much less likely to articulate what that looks like for them. So even when there's been 320 different interactions, including feedback, they're not seeing it is verbal feedback.

T2: Right.

Interviewer: And that's possibly, you going forward, to make it really explicit to all of your classes, this is feedback now [*saying it together with the teacher*]

T2: [*carrying on*] Yes: 'we're not having a chat, this is actual feedback, you need to pay attention to what I'm saying?' I've got a stamp actually that we could use when we have given verbal feedback, you don't have to sit down and laboriously write out 'please do this again with feeling', whatever.

Interviewer: It could possibly be that they are resistant to whatever it is that you *are* saying to them. So however well- formed the strategy-oriented feedback, or the effort-related feedback, *because* they're not seeing it is actual feedback -

T2: - is actual feedback [*saying it at the same time*]

Interviewer: they might not be actually internalising the (verbal) feedback.

T2: Hmm.

Interviewer: But then, that's their perception of feedback. It's been very interesting.

This section resonates with Elbra-Ramsay's (2019) study of student teachers being in the unique position as both donors and receivers of feedback, and the attendant emotion, even vulnerability that can accompany (potentially unwelcome) feedback (see also Pekrun *et al*., 2002; Eva *et al*., 2012; Gamlem and Munthe, 2014). The teacher has expended considerable effort over a period of time to change his feedback style, to be told by the researcher that his students do not recognise some verbal commentaries as feedback (Yang and Carless, 2013).

> Interviewer: So, going forward next year, what are you going to try, what different things will you try?
>
> T2: Ah, in part, starting early, and definitely being more explicit about it, as you say about the verbal feedback. So, signposting it as such. Continue with the effort and process focus.
>
> Interviewer: And making sure that they have clarity, for what the prompt is for their next step. Hattie, a prominent feedback researcher, once said that, when you give feedback to the whole class, no one listens, because they don't think you're talking about them. And I think the point to take that, is to speak to someone very directly, and for them to be sure, that they know, that that feedback loop was for *them*.

In becoming over time, the teacher who gave the highest amount of Process verbal feedback, T2 did indeed modify his VF style, despite the challenge to do this. Additionally, there was evidence of sustained verbal interactions focusing on regulatory habits which indicated that T2 valued this attribute and wanted to encourage it within his students.

7.5 Chapter summary

This chapter presented an overview of findings arising from the intervention teachers' conceptualisation of experiencing the intervention as a small case study. Four themes were identified for each of the teachers, including their self-reflections. There were differences between these teachers in terms of gender, experience, classroom presence and approach, yet similarities; both had deep and confident subject knowledge, a desire to promote independence in their students' learning, and a commitment to the professional development.

Nonetheless, both teachers experienced what the general CPD literature (Guskey, 2002a, Kwakman, 2003; Darling-Hammond *et al*., 2019) and feedback intervention studies (Schneider and Randel, 2010; Wiliam, 2010; Berg, Ros and Beijaard, 2014, 2015; Heitink *et al*., 2016) describe as the challenge transferring training content and ideas into their classroom practice. Both indicated that they intended to continue the attempt to flex their verbal feedback in the future, and outlined different ways in which they intended to do this: T1 by recording and analysing herself using the school’s designated professional development time, and T2 by reframing his verbal commentaries as verbal feedback and continuing to try to use regulatory, process and prompt feedback.

These teacher perspectives on their own practice and the development of their students’ self-beliefs together with triangulated data from other collections provide evidence to assert that it was the professional development intervention on teacher verbal feedback which enabled the teachers to modify their feedback styles, corresponding over time with mapped changes in the self-belief systems of the intervention students. The last chapter will now collate the findings in a discussion with respect to the existing literature and position the findings from this study as research contributions.

Chapter 8. Concluding discussions

This concluding chapter reviews the purpose of this research study and presents the findings interpreted in relation to the research literature. The findings both support and contradict some existing theory, and new insights afford a contribution to knowledge regarding both oral interactions and specifically, verbal teacher feedback. The chapter will thus provide a logical synthesis of the findings to answer the research questions, and opportunities and implications for teacher-practitioners considered. By reflecting on both new knowledge and good practice, recommendations for aspects of teacher professional development and Initial Teacher Education can be made. Potential limitations of the study will also be explored, and possible future research identified, concluding with personal contemplations on the process of undertaking research at this level.

8.1 Key findings as research contributions

It had been postulated that an increase in self-regulation and process feedback could have a positive impact on the self-belief systems of students in physics, comprising self-concept, self-efficacy, anxiety and mindset. Using empirical evidence gathered during the research to answer the research questions, findings support an assertion that the feedback intervention was successful in effecting a positive change in self-concept, self-efficacy and anxiety, although no effects were found for mindset. A consideration of the extant literature on mindset would indicate that this study has also failed to replicate Dweck and colleagues' experiments (see section 8.1.2 below). Given the quasi-experimental approach and mixed methods to provide triangulation for greater validity by integration of data collection, as well as the efforts to assure the internal validity of the groups, there is 'modest confidence' (Thyer, 2012) in the efficacy of the feedback intervention on the students' self-beliefs in self-concept, self-efficacy and anxiety.

It appeared through consideration of the evidence for RQ4 that becoming aware of the feedback types they used enabled the intervention teachers to modify their approach, however it did require a willingness on their parts to undergo the challenge of changing established verbal practices (Heitink *et al*., 2016; Andersson and Palm, 2017, 2018). That this change in feedback appeared to effect reciprocating influences on student self-efficacy and their sense of agency towards their own learning, as well as their self-concept in physics has implications for both current and future teachers' practice.

As hypothesised, the CPD intervention emphasising the need for increased self-regulation and process teacher feedback impacted positively on the students' self-beliefs in physics for the intervention group when assessed against the comparison students. This main finding, the purpose of the quasi-experimental approach was however underpinned by several other key findings, which are presented as summaries and subsequently discussed with respect to the extant literature.

1. A longitudinal CPD package including coaching conversations was found to have a positive impact on modifying teacher feedback to include more self-regulation and process types. However, as found in other studies (Darling-Hammond *et al*, 2009; Bergh, Ros and Beijaard, 2013; 2014; 2015; Schütze *et al*, 2017; Andersson and Palm, 2018), even committed and effective teachers experience difficulty in transferring training content to their real-time instructional practice, although change can be achieved with sustained CPD models. This therefore has implications for Initial Teacher Education in developing feedback practices of early-career teachers.
2. Analysis of the pre- and post-intervention survey data on students' self-beliefs showed an increase in physics self-concept and physics self-efficacy, and a decrease in physics anxiety over the duration of the study for the students in the intervention group compared to the groups not receiving the intervention. Survey stability and reliability were high, and the analysis and findings supported other studies which operationalise

these as distinct self-belief constructs. There was negligible effect on mindset for the intervention group.

3. Students who received a higher amount of both *prompt* and *Process* feedback, as well as *self-Regulation* feedback developed more positive self-belief systems (gains in self-concept and self-efficacy, reduced subject anxiety). These students were in the intervention test condition, supporting the findings that the CPD intervention was successful in modifying teacher feedback with the aim of improving student agency.
4. Teachers have dialogic characteristics that can be mapped across lessons and form a 'teacher profile' or repertoire; teachers vary in the amount of instructional dialogue they give to students; feedback styles as part of oral interactions vary between teachers and are a key part of their dialogic repertoire;
5. There are a large number of oral interactions made by teachers in lessons, and feedback interactions form a high proportion of this, although students do not necessarily recognise it as feedback;
6. Relationship is key, and oral interaction including verbal feedback can build relationships; feedback does not exist in a vacuum, and the context and learning environments are hugely influential; the notion of disparaging ‘praise’ as detrimental to learning needs to be nuanced;

These main findings are now discussed with respect to the contribution they make to existing research literature.

8.1.1 The impact of the intervention in modifying teacher feedback

This study found that teachers initially found it challenging to flex their feedback into other forms, despite the less ‘traditional’ CPD method of individual intervention and subsequent coaching sessions; this personalisation represents an unusual amount of CPD for a classroom teacher in the UK. Teacher development literature indicates that more traditional forms of CPD (such as external courses) do not always transfer to classroom practice (Guskey, 2002a, Kwakman, 2003; Darling-Hammond *et al*., 2019), and the much smaller field of teacher formative assessment CPD literature highlights the difficulties in supporting teachers to implement high quality formative practices (Schneider and Randel, 2010; Wiliam, 2010; Berg, Ros and Beijaard, 2014, 2015; Heitink *et al*., 2016), few of which involve teacher *verbal* feedback.

This study utilised a professional development trajectory (Joyce and Showers, 2002) containing theory, demonstration, practise and coaching. Joyce and Showers (*ibid*) found that the latter component best increased the transfer of skills into the classroom and was thus chosen as the intervention method (see also Darling-Hammond *et al*., 2017). Voerman *et al*. (2015) added a fifth component of feedback after Gabelica *et al*. (2012) who advocated feedback as a necessary aspect of teacher professional development. This was also incorporated into the coaching conversations in this study, in which the intervention teachers would receive feedback on their oral interactions after an observation and be coached through an attempt to flex it into different forms. The observation data from Chapter 5 and the teacher narrative from Chapter 7 provide thick descriptions and indicates that both intervention teachers laboured to modify their verbal feedback at first, which support other feedback CPD studies' findings (Berg, Ros and Beijaard, 2014, 2015; Zaccarelli *et al*., 2018). The intervention teachers became more practised at using Process and prompt feedback over time, although the self-Regulation type fluctuated; often 'peaks' of incidence followed a coaching conversation. In a similar, though larger and multi-layer CPD trajectory, Voerman *et al*. (2015) also saw an increase in types of feedback used, although they used a different feedback typology. In a teacher formative assessment study, Andersson and Palm (2017, 2018) described 'time' as an aggravating condition, during which new processes required special attention to embed; notably the authors reported student outcome gains in the *subsequent* academic year as their participant teachers embedded their feedback practices (see also EEF, 2018).

Both Hammerness *et al*. (2005) and Korthagen (2010) describe the simultaneous demands of both social and academic goals on moment-by-moment class time, from which intuitive routines are developed by experience (Eraut, 2004). This routinisation has the benefit of freeing 'practitioners' attention for monitoring what are often rapidly changing situations' (*ibid*:261); however in the intervention group teachers, this routinisation often produced an instant feedback response, 'rituals of feedback' (Molloy *et al*., 2012) that was more consistent with past experience. In short, teachers tend to revert to their default mode of feedback in the classroom instance. 'Given the complexity of teaching and the need for prompt reactions to situations, it is not surprising that teachers find it hard to translate theory into their daily practices' (Voerman *et al*., 2015:993). The necessity of moving from implicit and reactive stages to more deliberate practice requires reflection on current realities, planning learning opportunities and rehearsing future events (Eraut, 2004). The coaching conversation thus took this form, and the study showed in Chapter 5 that the intervention group teachers were successful over time in flexing their feedback to more procedural, strategic and regulatory forms. Though this was not to the extent that the intervention teachers had wished, both indicated that they would continue with their efforts, having been fully persuaded into the importance of this pursuit, and the intention here speaks to the ultimate sustainability and efficacy of the feedback intervention.

8.1.2 Findings about students' self-beliefs in physics

Chapter 4 presented the evidence to answer research question 1, clearly ascertaining measures of self-concept, self-efficacy, anxiety and mindset. In total, the measures for the intervention cohort indicated higher 'starting-points' in all self-belief constructs, which could be interpreted to denote that the intervention group students were more confident learners, or more positively adapted to learning in the first instance. RQ2, which asked, how did these constructs *change* was also answered in Chapter 4; there were positive gains in self-concept, self-efficacy and a decrease in anxiety for the intervention group, whilst mindset appeared stable. In the comparison group, self-concept decreased, negatively correlating with an increase in anxiety (Ahmed, 2019; Lee, 2009, Morony *et al*, 2013), however self-efficacy increased, though not to the extent of the intervention group. This contradicts findings by Bong *et al.*, 2012; Seon and Bong, 2019; Wolters *et al*., 1996 who show negative correlation of self-efficacy with anxiety, though the intervention group agrees with these assertions. Mindset in the comparison group showed a slight gain. Overall therefore, the intervention group demonstrated more improvement in positive self-belief, and the study is strong in considering four self-belief constructs, rather than one, and makes a research contribution to self-concept, self-efficacy and anxiety corpus of literature, but not mindset.

Analysis of the survey items demonstrated high internal reliability with Cronbach's alpha of both pre- and post-intervention surveys ranging from 0.782 to 0.866 (Cohen *et al*, 2017). Additionally, split-half reliability measures for all four self-belief constructs range from 0.846 to 0.873 (Spearman-Brown coefficient) and 0.802 to 0.869 (Guttman coefficient). There is thus confidence that the survey is a reliable instrument. Survey validity was considered in Chapter 3.

8.1.2.1 Self-concept

From Chapter 2, self-concept can be viewed as a more complex, stable, but more general construct than self-efficacy (Seon and Bong, 2019), and includes both cognitive (thoughts about) and affective (feelings about) judgements of oneself (Bong and Clark, 1999; see also Seon and Bong, 2019). Whilst most self-concept items showed consistent results across both groups, item 16 ('I have always done well in physics') indicated a surprising reversal; the comparison group showed higher self-concept, although both decreased on this item (Graph 4.10). One interpretation is that the phrasing of the question has resulted in the IG students reconceptualising this in the *past*, whilst agreeing with their current, otherwise high self-concept. Their *affective* feelings regarding their self in the subject, despite good grades, however they are comparing this with their *cognitive* feelings, in recognising that they are now achieving well (Bong and Clark, 1999; Seon and Bong, 2019). This may be explained by Marsh *et al*.'s (1990a) dynamic reciprocal effects model, which posited that domain-specific academic self-concept and achievement were mutually reinforcing, leading to gains in each other (see also Marsh and Martin, 2011; Marsh *et al*., 2019b)

Borrowing from Merolla *et al*., (2013) and Brenner, Serpe and Stryker (2014), and using the concept of hierarchy of salience of identity coined by Stryker and Serpe (1982), the researcher posits the concept of 'subject salience'; that students have an internal 'league table' of preference of subjects. The gain in self-concept for the intervention group could indicate that the subject has risen in some of the students' subject salience hierarchy, certainly desirable for influencing student continuation in physics to post-16 (Murphy and Whitlelegg, 2006; Hollins *et al*., 2006), since a higher position in the hierarchy indicates more relevance to the individual, which is thus more likely to be used (Merolla *et al*, 2013; Brenner, Serpe and Stryker, 2018).

8.1.2.2 Self-efficacy

Self-efficacy is described as a malleable subjective conviction by Seon and Bong (2019), and less general than self-concept, being domain- and context-specific. Both the intervention and comparison groups increased in physics self-efficacy, although there was a larger increase in the intervention group, who had started from a higher level initially (Graph 4.5). If self-efficacy is the 'foundation of human agency' (Bandura, 2011:10), an increase should enable students to realise both the capacity and propensity to take purposeful action towards their own learning (Ferguson *et al*., 2015). Demonstrating the opposite of helplessness, T1 students indicated greater *agentic action* (Bandura, 1989) during both the later observations and in the focus group conversations. These included more reference to, and ownership of personal mastery experiences (Bandura, 1997; Britner and Pajares, 2006; Schunk and Pajares, 2009; Usher and Pajares, 2008) which reciprocatingly strengthens self-efficacy further, promoting a disposition in which they are more likely to persevere (Bandura, 2018). Agency is discussed further in section 8.1.3 below.

8.1.2.3 Anxiety

Since anxiety has been show to correlate negatively with both self-efficacy and self-concept (Bong *et al*., 2012; Marsh, 2019a; Morony *et al*., 2013; Seon and Bong, 2019; Wolter *et al*., 1996), an increase in the latter two could be expected to be associated with a decrease in the former. Indeed, the intervention group did show a decrease in anxiety, and from a position of pre-intervention lower mean anxiety than the comparison group (Graph 4.3), agreeing with those authors. However, the comparison group increased in anxiety in the pre- and post-intervention conditions, which correlates with the change in their self-concept, but not their self-efficacy, which increased, apparently contradicting this relationship. This suggests that the self-efficacy reported above for the comparison group may be specific to the self-efficacy items used on the survey, and that the comparison group had efficacy for those items, but not for others *not* asked on the survey. They would consequently still experience low self-concept and high anxiety; thus, the self-concept measure is a more accurate indicator for that group.

In a focused anxiety domain, Gungor, Eryilmaz and Fakioglue (2007) reported that physics anxiety affected student achievement, as did Sahin (2014). Using the Physics Anxiety Rating Scale [PARS], Sahin, Caliskan and Dilek (2015) reported that physics anxiety was higher in females than in males, a finding also reported by Agra, Fischer and Beilock (2017). This was not supported by this study, which found no significant difference between students' anxiety score in terms of gender, similar to the Brownlow, Jacobi, and Rogers analysis, (2000).

8.1.2.4 Mindset

The intervention group showed a very small decrease in the mean value of mindset (2.97-2.96); this is a value indicating that as a cohort they identify overall as incremental theorists. The student voice in Chapter 6 displayed aspects of self-regulatory habits, and an acceptance of learning from mistakes, representative of learning goals rather than performance goals. The comparison group had a lower mean value of mindset at the start of the study (Graph 4.6), though this was still overall in the incremental theorist range, and this rose over the study to a higher incremental value. However, these changes are small in total.

Mindset interventions have been claimed to ‘lead to large gains in student achievement’, (Yeager and Walton, 2011:267) and ‘have striking effects on educational achievement even over months and years’, (*ibid*:268); Good, Aronson and Inzlicht’s (2003) study asserted effect-sizes of Cohen’s d = 1.13, 1.30, and 1.50 for attributional retraining, implicit theories, and combined interventions produced effects on girls’ math test scores respectively (see also Dweck and Yeager, 2019). These claims have been disputed by the meta-analysis of Sisk *et al*., (2018) which overall found weak effects and low correlation between growth mindset interventions and achievement.

The findings from this study would appear to echo recent failures (EEF, 2015, 2019; Li and Bates, 2017) to replicate Dweck and colleagues’ reports. It may be that this study failed to be the ‘genuine replication(s) and thoughtful replication(s) done by skilled people’ pronounced by Dweck in Chivers (2017).

A further suggestion may relate to the *validity* of the mindset survey items. With a prevalence of resilience and growth mindset discourses in schools (EEF, 2015, 2019), and the ‘deficit model’ nature of these, (Elbra-Ramsay, 2019), it may be posited that firstly, students do not want to *appear* to think that they believe intelligence to be fixed, even if they do, and secondly, that they are able to ‘game’ the responses to the survey so that they personally do not appear to be fixed mindset individuals, despite narrative observed in the classroom. Combined with the emergent criticism that previously acclaimed mindset interventions cannot be repeated, two further suggestions arise that the survey items relating to mindset may not be a valid measure of entity/incremental theory, and that a more valid (and certainly grounded in context) approach would be to observe response to challenge *in situ*. The girl observed in the T2 group who would not write in her book unless it was ‘perfect’ and ‘right’ tested as an incremental theorist, yet her attitude and language demonstrated performance rather than learning goals (field notes). In sum, if growth mindset findings and effects cannot be replicated, this may undermine the validity of surveys which purport to measure it in the first instance.

In summarising the outcomes from the surveys, the physics self-belief constructs of sixty-six Year 10 students were identified as self-concept, self-efficacy, anxiety and mindset. The two intervention classes appeared to have a higher baseline in all constructs than the comparison classes, although these differences were slight given the sample size.

A broad comparison of the intervention and comparison groups' self-belief constructs over the duration of the study indicated that not only that change in self-belief could be mapped, but that for 16 of the 20 items on the survey there were more positive outcomes for the intervention group than the comparison group, as discussed in section 4.2.5. Overall, the intervention group present as a student cohort with increasingly adaptive learning dispositions, and a greater sense of agentic action towards physics over the duration of the study.

There appears to be some tension between the 'growth mindset' message espoused by some schools and the clear message transmitted about achievement goals [aspiration/ target grades], which is counter-indicative to incremental theory, versus learning goals, which helps promote it. Students in these schools are exposed to surveys and language that shapes their judgement about which they should seek to be, even if they are not, and it may be that they become versed in recognising the 'answer they should give'. For example, a female student in Teacher 2's class (Student 12) gave survey answers indicating a growth mindset but made remarks in class that would indicate that this may not be the case – at least as far as physics is concerned. These included; 'I hate it when the teacher says, are you sure?', 'I prefer it when I can do it easily', and showing extreme reluctance to write anything in her book unless she knew 'it was right' on several occasions. In recent years, Dweck (2016) has started to refer to this as 'false growth mindset'; people saying they are, when they are not. It would appear from these survey results that teacher feedback efforts on the constructs of self-efficacy, self-concept and anxiety may bear better fruit than the seemingly more elusive mindset effect.

8.1.3 The effect of feedback types on students' agency in physics

Since the quasi-experimental research approach demands that the groups be comparable for internal validity (Bryman, 2016), a concern could be raised that the self-belief changes in the intervention group could be due to chance as the groups and individuals were not randomly assigned. The internal validity was discussed in Chapter 3 and observed that by using multiple groups which are comparable, threats to internal validity are minimised. Noting the slight difference in starting points in self-beliefs of the two cohorts, it is important to 'make some judgement about the plausibility that a selection bias exists' (Trochrim, 2007). However, the positive outcomes in the majority of the survey items for the intervention group compared to the non-intervention group seem to indicate that a change in characteristics post-intervention could indeed be attributable to that intervention.

The intervention teachers' increase in the employment of self-regulatory feedback was smaller than hoped for, however both process (first level) and prompt (second level) feedback showed a larger increase for the intervention group [the comparison group did not show positive gains in self-belief constructs overall]. Both process and prompt types would fall into the category of *instrumental* help, rather than executive (Hattie and Timperley, 2007, see below). Correlating this with increase in self-efficacy allows a causation to be modestly offered given the validity issues above, and this agrees with both correlation and causation agency literature (Schunk and Pajares, 2005; Chen and Lam, 2010; Arslan, 2012; Darling-Hammond *et al*., 2019). Hattie and Timperley (2007) highlighted the importance of process and regulation feedback in promoting both self-efficacy and self-regulation proficiencies and, recognising the reciprocating nature of these concepts, Kluger and DeNisi (1996) asserted that student self-efficacy was an important mediator in acceptance or rejection of teacher feedback. Additionally, apposite studies on the effect of feedback on self-concept place it as the highest effect-size intervention (O'Mara *et al*., 2006), and Burnett and Mendel (2010) note the particular impact of *effort* feedback, especially for older learners.

Ritchie (2018) notes that students exercise agency when they are dynamic in planning and executing their own learning processes, and that in this, there is overlap between agency and self-regulatory practices (see also Schunk and Usher, 2013). Seeking help is thus a learner proficiency; Hattie and Timperley (2007) discuss the distinction between *instrumental* help (seeking hints on how to work something out) versus *executive* help (seeking answers). Students in both intervention classes displayed an increasing ability to apply strategy to their physics reasoning and explanations; that the physics classroom is well-situated to emphasise process-type feedback in its focus of applied calculations where those contexts dominate (see also Schütze *et al*., 2017 for a situated mathematics feedback CPD study) is also an example of a condition in which the most powerful source of self-efficacy, mastery experiences (Bandura, 1997; Schunk and Pajares, 2005; Britner and Pajares, 2006; Usher and Pajares, 2008) are being constantly enacted by the students. The verbal teacher feedback received in these situations may indeed by acting on the intervention students in a socio-behavioural sense, such as suggested by Hitlin and Elder (2009; see also Ferguson *et al.*, 2015).

Whilst Kluger and DeNisi (1996) noted that both positive and negative feedback can have beneficial effects on learning, a finding supported by the Teacher 1 student voice thick descriptions, Hattie and Timperley (2007) argued that these effects depended more upon the *level* of feedback, and how it was acted upon. Disentangling the nomenclature of how terms are used within the feedback literature is problematic, for instance the conflation of ‘prompt’ (Bergh *et al*, 2013; Adie *et al*, 2018), ‘feedforward’ (Hattie and Timperley, 2007), ‘descriptive’ (Tunstall and Gipps, 1996; Gipps *et al*, 2000) and ‘discrepancy’ (Voerman *et al*, 2012, 2015) have over time all described the same mechanism, whether this was negative or not. Indeed, the Teacher 1 student cohort seemed able to relate both negatively and positively framed teacher verbal feedback to verification of their learning, and welcomed both, even indicating a preference for negative [deficiency feedback], since they deemed it gave them ‘next steps’ or prompts.

In summary, the students in the intervention group were observed to take more agentic action in their physics learning over the duration of the study, notably so for the T1 group and less so for the T2 group [field notes and student focus groups] as well as the post-intervention survey analysis indicating a positive change in self-belief constructs of self-concept, self-efficacy and anxiety. This coincided with the teacher verbal feedback they received increasing in process, regulation and prompt types as a result of the CPD intervention. Increased process, strategy or 'how-to' as well as prompt, deficiency or 'next-steps' verbal feedback would seem to be the most helpful in promoting all four self-efficacy sources (Bandura, 1997); mastery experiences being the most influential; vicarious experiences in which the (for example) calculations or explanations are modelled, focusing on strategies and approaches; social persuasions in the form of meaningful verbal judgements on student capabilities; and emotional states such as reduced anxiety becoming an enabling factor (Schunk and Pajares, 2005).

8.1.4 Dialogic characteristics of teacher feedback

Exploring teacher feedback drew upon the typology outlined by Hattie and Timperley (2007), as well as the notion of Ramaprasad's (1983) 'feedback loop', where progress and discrepancy feedback (Voerman *et al.*, 2012, 2015) were reimagined as 'statement' and 'prompt' feedback at a second level. This constructed typology included a third level to indicate the recipient of the feedback. This study contributes to existing feedback literature and extends knowledge by comparison with earlier studies from the teacher and/or student perspectives. Using four classes in a double quasi-experiment has enabled a moderate sample size of participants to gain valuable insights of feedback practices in contemporary science classrooms. Additionally, few quasi-experimental studies incorporate qualitative aspects, thus a secondary aim was to use classroom observations, and to discover qualitative awareness of teacher and student perspectives on their experience of the intervention.

Chapter 5 indicated that an analysis of teachers' verbal interactions had enabled an identification of the types of feedback that they used and established that a 'feedback repertoire' could be shown for the different teachers in the study, thus answering research question 3. These repertoires showed variation in the amount of instructional interaction that took place, as well as feedback in terms of first, second and third levels. It also highlighted the differences in 'teacher talk' within the classroom ranging from structured, teacher-led, and less dialogic practices to responsive, individualised and authentically dialogic patterns, culminating in widely varying amounts of feedback given to the students, as seen in Tables 5.5 and 5.6. Both teachers in the intervention group provided between 1.6 and 2.0 times the amount of *learning* feedback to their students as the comparison group teachers. This would seem reminiscent of Molloy and Boud's (2013) 'double nostrums', however upon consideration of the composition of this feedback, it is clear that it is not quite 'the more the merrier' (*ibid*), but more self-Regulation, more Process and more Prompt, and the implications of this were discussed in section 8.1.3 above.

An overview of the first-level feedback types was shown in Table 5.1, which indicated that the intervention teachers had broadly the same amounts of Task feedback, and this comprised the majority of their feedback at this level, agreeing with findings from previous studies (Brooks *et al*., 2019; Gan, 2011; Hattie and Masters, 2011; Bergh, Rose and Beijaard 2013). This was a smaller proportion than T4 and significantly below that of T3. The second highest amount was represented by Process feedback, again agreeing with those studies above [For Teacher 2, Process feedback formed nearly one third of their profile]. However, when comparing Table 5.1 with the studies summarised in Table 2.2, it can be seen that the intervention group teachers had lower proportions of Task, and higher proportions of Process feedback than three of the studies, but was in broad agreement with the Burgh, Rose and Beijaard (2013) report. Similarly, T1 and T2 showed much higher proportions of self-Regulation feedback; T1 more than 3 times, and T2 nearly twice as much as the next highest amount recorded. Additionally, T1, with the highest amount of

Self feedback at 2.3%, was still lower than the Self feedback reported in the other studies.

Over time, the amount of Task feedback deployed by the intervention teachers decreased, and the amount of Process feedback increased. Self-Regulation feedback varied over the course of the study and showed higher incidences following the initial CPD and subsequent coaching conversations. Whilst the amount of self-Regulation feedback was small, it represented an increase on three of the four similar studies using Hattie and Timperley's (2007) typology (Brooks *et al*., 2019; Gan, 2011; Hattie and Masters, 2011).

At the second level of feedback, Table 5.1 also indicates the relative proportions of statement and prompt feedback, and Chapter 5 discussed how the intervention teachers had lower statement-to-prompt ratios than the comparison teachers, which when combined with their higher incidences of Process interactions implies that their students were receiving more feedback related to their thinking and learning processes (learning goals) and mastery *experiences*, rather than their outcomes (performance goals). According to Bandura (1989, 2011; see also Britner and Pajares, 2006; Ferguson *et al*., 2015; Usher and Pajares, 2008), this should effect reciprocating influences on self-efficacy and their sense of agency towards their own learning, and implications of this finding for both current and future teachers' practice are discussed later. Over time, the amount of prompt feedback increased, particularly for Teacher 1.

At the third (recipient) level, the majority of interactions were targeted at individuals for three of the teachers; T4 differed in having a class-directed instructional approach in which the whole class received 56.0% of oral interactions related to learning (and 43.2% was coded at the first level as Instructional). The teachers in the intervention groups gave approximately equal amounts of oral interaction to males, females and the whole class. Preferencing of recipient was not observed in any observation.

In summary, this study found that each teacher had dialogic characteristics that could be mapped and represented their feedback profile' or repertoire, as discussed in Chapter 5. These teachers varied in their oral pedagogy, and features of this could be characterised, such as T1's energy in the classroom resulting in a very high amount of individual, class and group oral interaction compared with the other teachers; the subject mastery and focus on Process of T2; the Task-Product dominance of feedback for T3; the front-loaded whole-class Instructional input of T4 before releasing the class into individual work. These feedback styles as part of oral interactions were distinguishable between the teachers and given the prevalence of feedback interactions as a proportion of their total interactions (T1 = 60.7%; T2 = 75.2%, compared to T3 = 56.4%; T4 = 52.9%), for the intervention teachers feedback was a key part of their dialogic repertoire.

8.1.5 Verbal feedback or dualistic learning conversations? Frequency and student perceptions of teacher verbal feedback

In Chapter 5, Table 5.5 summarised the mean distributions of both feedback and non-feedback interactions by all four teachers, comprising *total* oral feedback interactions (N_T), oral interactions relating to *learning* (N_L), and oral *feedback* interactions only (N_F). It was noted that both intervention teachers had more interactions in all categories than the comparison teachers. Teacher 2 used a conspicuous amount of both interactions related to learning as well as feedback interaction, and both teachers were significantly higher in feedback interactions than their comparison colleagues. This certainly contradicts Hattie's assertion that 'the incidence of feedback in the typical classroom is very low, usually in seconds at best per day' (1999, see also Bond *et al*., 2000; Campbell-Mapplebeck, 2019). Similarly, Voerman *et al*., (2012) found that seven of forty interactions within recorded ten-minute segments were feedback of some form. A scaled amount from this would approximate to 240 *total* interactions, but only 42 feedback interactions per hour; this study shows that the feedback figure from the teacher using the smallest amount exceeds this significantly. It is important to note however that those authors were coding for positive/negative, specific/non-specific feedback and would have applied different selection criteria using different typologies. Nonetheless, this study demonstrates that oral teacher feedback can account for at least 52.9% of all oral interaction in the classroom, and in the case of Teacher 2, as much as 75.2%. It aligns more with Hargreaves' conception (2013) of assessment for learning *as* a classroom conversation about learning, and Svanes and Skagen's notion of feedback being an important part of a teacher's repertoire (2017). The reconceptualisation is thus dependent upon the analysis of the feedback types used; here, framing feedback as Task, Process, Regulation (first level) and statement/prompt (second level) enables the learning-focused facets of teacher dialogue to be magnified as feedback in a way that other typologies may not allow. On a cautionary note however, larger doses of administered feedback do not necessarily equate to feedback which has been received and acted upon by students. Nonetheless, this study does serve to open further dialogue about the role of verbal feedback as dialogic learning.

Since the students would therefore seem to be receiving a large amount of feedback dialogue, what *is* their perception of it? This may be a case for two of Molloy and Boud's 'nostrums' (2013); 'all feedback is good feedback', and 'the more the merrier'. Yet the students in the T1 focus group appeared fully cognizant of the oral feedback they received in the lesson in terms of both the amount and nature of positive to negative feedback, specifying that they thought that T1 delivered the 'right amount' of negative to positive feedback (60:40). This differs from both Shute (2008) and Voerman *et al*. (2012); the latter recommending that positive feedback should outnumber negative 3:1. T1 students specified further that they did not consider the negative to be 'bad' because of the kind manner in which it was delivered. To these students, discrepancy feedback was both motivational and instructional (see Bandura 1997). In contrast, despite being the teacher with the highest proportion of oral feedback interactions, often setting up extended chains of interaction, T2's students did not appear to recognise the dialogue as feedback at all. When questioned about their views on feedback in Chapter 6, the students spoke only of written feedback (i.e., in their books). Whilst this may reflect an institutional way of discussing feedback, it emphasises a worrying concern that many students do not recognise or value feedback when they receive it, or misunderstand its nature, as highlighted in other studies (Carless, 2006; Glover and Brown, 2006; Gamlem and Smith, 2013; Yang and Carless, 2013; Scott, 2019) in privileging written feedback over verbal (Handley, Price and Millar, 2011; Elliott *et al*., 2016). It may be the case that these students internalise the dialogic feedback as ‘enhanced learning’ such as Gamlem and Smith reported (2013).

8.1.6 Emotion as a mediator of feedback

As demonstrated in Chapter 6, many of the intervention students displayed a range of emotions related to their physics learning, and that these changed, becoming more positive over the duration of the study. Existing literature has established a link between emotions and feedback (Pekrun *et al*., 2002, 2014; Daniels *et al*., 2009; Eva *et al*., 2012; Voerman *et al*., 2014), feedback is not given in isolation (Hounsell *et al*, 2008; Brooks *et al*., 2019), and emotions have been noted to affect student achievement (Hattie and Timperley, 2007). Kluger and DeNisi (1996) asserted that students can reject goal-oriented feedback, and that only two of eight possible student responses to teacher feedback are positive. Pekrun *et al*. (2014) highlighted the differences in students’ emotions produced by receiving self-referential feedback (linked with mastery or learning goals) versus receiving normative feedback (linked with performance goals); intervention students were increasingly likely to exhibit more positive emotions associated with mastery (enjoyment, hope, pride, relief) and a decrease in anxiety and hopelessness. Even when a student exhibited anger (at being one mark off an A grade), this annoyance was self-directed indicating ownership of physics learning.

Students also vocalised how feedback impacted on their learning identities in terms of physics competence and self-worth, which ultimately influenced whether they rejected or engaged with the teacher verbal feedback. Voerman *et al*. (2014) placed emotion as an important issue within the feedback discourse and challenged what they perceived as an over-simplification of feedback about the self as not useful, and occasionally detrimental to student learning (Kluger and DeNisi, 1996; Mueller and Dweck, 1998; Shute, 2008;

Hattie, 2009; Brooks *et al*., 2019). This study has found praise, even as non-specific to learning to be helpful in encouraging positive and motivational dispositions within the emotional space of the classroom; Hattie (2012a) somewhat moderated his earlier stance in acknowledging this.

Equating *prompt* feedback with Voerman *et al*.'s model of *discrepancy* feedback, (2012, 2014, 2015), and noting that Voerman *et al*. (2014) discussed the impact of feedback on cognition, emotions and character strengths, this study agrees with Voerman *et al*.'s (2014) hypothesis that feedback on character strengths is more beneficial if it is related to performance or to the task at hand, but would further specify that firstly, these 'character strengths' are categorised as self-belief constructs, most notably self-efficacy, and secondly, that limiting the effects to performance or task would negate gains in learning processes (learning goals) and mastery experiences.

In shifting feedback studies from the teacher-as-donor to the student-as-recipient, it is necessary to acknowledge that feedback is both dynamic and socially constructed, it does not exist in a vacuum, and that there are sometimes unpredictabilities in emotional consequence arising in the recipient. *Verbal* feedback, occurring more naturally within the learning instance, is more ideally situated to avoid recipient ambiguity and potential rejection – provided the student understands it to be feedback.

This study has drawn attention to the reciprocating interactions between the language of teacher verbal feedback, the teacher-student relationship and students' self-beliefs. How language is used in verbal feedback has consequences for the way in which students internalise the language to make judgements and decisions about how to respond to the feedback. Linguistic ethnographers explore how classroom discourse enables or inhibits expression of identities, in which self-belief plays a key role. Sociocultural researchers emphasise how talk allows 'reciprocity and mutuality to be developed through the continuing negotiation of meaning' (Mercer, 2010). The purpose of this study was an attempt to prove that verbal feedback can impact on student self-belief, yet it is contingent upon the student internalising the feedback, making a decision to firstly accept it, and subsequently to act upon it, and here the mediating role of the student-teacher relationship would appear pivotal (Hollins *et al*., 2006).

8.2 Implications of the Study

Taken together, these findings provide useful information regarding the role of oral interactions, and in particular verbal feedback in supporting student development of positive self-beliefs in physics, and by extension, to other potential curriculum areas. The study provides both insights into practical authentic classroom feedback and suggestions for several different education stakeholders, including CPD and Initial Teacher Education providers, school leaders, and not least, classroom practitioners themselves.

8.2.1 CPD Providers

Both a review of the research literature on broad-spectrum teacher CPD as well as CPD dedicated to formative assessment has indicated the need for it to be content-focused, job-embedded and supported as well as being effectively modelled, including reflection and of sustained duration (Darling-Hammond *et al*., 2017), and for teachers to be able to diagnose and assess 'on the spot' and produce constructive and focussed feedback with which students can engage (Heitink *et al*., 2016). Providers should thus ensure that their professional development includes the pre-requisites above, and that practitioners are supported over time to embed effective practice. Useful approaches in enabling this over time could include collaborative practices such as coaching pairs or groups, video lesson capture, reflection and analysis and action research approaches.

8.2.2 Initial Teacher Education Providers

Given the challenges experienced by the participant teachers in modifying their in-the-moment classroom practice, it is anticipated that these findings will have implications for inclusion in Initial Teacher Education, since good feedback 'habits' formed early in a teaching career would be extremely beneficial. It is therefore recommended that content regarding both the relationship that feedback may have on student outcomes and self-beliefs as well as the different types of verbal feedback that may be employed as part of a repertoire be included in undergraduate and postgraduate training programmes. During these periods of development and being in the unique situation of being both recipients and donors of feedback themselves, student teachers are well placed to reflect upon and develop verbal feedback practices which are focused on both student academic progress and improving self-beliefs in order to foster resilience and independence in learning. Should student teachers become more aware of different feedback typologies and their impact on different levels of learning, they may be able to modify their feedback practice before routinisation becomes fully formed.

8.2.3 School leaders

Teacher CPD is often undertaken on the individual level, yet has more impact when for example, there is whole-department involvement (Voerman *et al*., 2015) and/or wider level support (Joyce and Showers, 2002). Operationalising change in teacher feedback practice over a period of time, especially at the school level requires executive leadership (EEF, 2018). Including all teaching staff, whilst requiring leadership resourcing and time, ensures a consistent approach across the institution through teacher development groups, reflection and on-going support. Individualised recommendations within this are suggested in the teacher-practitioner section below.

8.2.4 Teacher-practitioners

It is anticipated that the outcomes of this study will also be of interest to teacher-practitioners within a variety of educational settings, including age stages from Early Years to Higher Education. Future teacher action could include:

- Teachers should review their feedback practices through recording and reflection. They should be cognizant of the types and proportions of feedback they use and how this can change with context; also, that teachers use more feedback in lessons than other studies have asserted;

- Increasing the deployment of regulation and process feedback builds the self-efficacy of their students and fosters an 'I can' approach for agentic action; highlighting prompt-type feedback as next steps for action on the part of students and encouraging them to take these steps themselves;
- There is a need for honest dialogue between teachers and their students to make it clear that they are giving verbal feedback, which may be much more than the students themselves realise; also, that this verbal feedback can be preferenced to written feedback given that more of it occurs in the learning instance rather than at endpoint;
- It would be beneficial to include within the dialogue here what 'negative' and 'positive' feedback looks like and ask their students how much they want of each. Teachers should demonstrate how much (two-way) verbal feedback is valued within their learning environments, to recognise learning need and enable both progress and prompt (discrepancy) feedback.
- There is a need to nuance the discussion around the definition, nature and use of 'praise'. Perhaps due to the received wisdom from literature that praise is not useful to learning and may be detrimental (Mueller and Dweck, 1998; Hattie and Timperley, 2007; Shute, 2008; Hattie and Masters, 2011; Brooks *et al.*, 2019), Teacher 1 was disparaging of her own 'back-slapping' practice. However, since this was observed to influence a positive learning climate, as also perceived by Voerman *et al*. (2014), there is perhaps a place for the role of praise as an 'emotional lubricant' for learning in the goal to be a supportive teacher of physics (Hollins *et al*., 2006).
- As a teacher development model within schools, coaching pairs or triads could observe practice and feedback;
- The findings disagree with Hattie and colleagues' assertions that competent students mostly need regulation feedback (Hattie and Timperley, 2007; Hattie, 2012b; Brooks *et al*., 2019), since this study

> finds that an increase of regulation and process feedback has produced self-belief benefits *across* the intervention cohort. The increase of both process and regulation category deployment used to enhance self-concept and self-efficacy suggests that schools would be better to concentrate their efforts on these constructs rather than the ‘thoughtful replication by skilled people’ required to duplicate mindset effects (Chivers, 2017).

Overall, whilst there are implications for policy makers and CPD providers, the length and type of professional development trajectory employed here would be an expensive model for schools unless commissioned as in-house coaching and action research approaches. However, further corroboration of these findings could lead to a reclaiming of verbal feedback as a powerful motivational and instructional device operating fully within the learning instance.

8.3 Limitations of the Study

Limitations of the study arise through both methodological and researcher limitations, although much of the efforts expended to reduce each of these have been discussed in Chapter 3.

Although the data set obtained was both extensive and complex, the number of participants was small. The two teachers in the intervention group were not of the same experience, nor were the two teachers in the comparison group. In each case, one had considerably more experience than the other, and two were recently qualified. Whilst this might be argued that their feedback repertoire had not yet been fully established, the maturity of T2 could mean that his responses had become more fixed over time.

The comparison teachers and students were not interviewed beyond conversation around their classes; whilst this decision was influenced by needing to keep the intervention separate from the comparison teachers, their view on institutional enactment of (for example) ‘growth mindset’ might have provided additional insights.

Non-verbal communications could not be captured if the researcher was not physically present during an observation. However, field notes were undertaken as much as possible. Video, which may have captured this was not used as it was not able to do this for all teachers, and as T1 indicated, the knowledge that they were being recorded impacted on the students’ learning behaviour and attitude.

Although observations continued on the same intervention and comparison classes, the same number of lessons from each teacher were not obtained, nor did CG teachers have as many recorded as IG teachers. However, as comparison to other studies has shown (section 3.4), there has been a range of time over which lesson observations have been made, and so the relatively small number of CG lessons exceeds these.

A recording of the T2 CPD intervention failed to record all the way through, however substantial field notes were taken. Additionally, one coaching conversation recording subsequently corrupted electronically, but again, field notes had been taken, and there was sufficient remaining data to note the pre- and post-intervention conditions of T2.

Investigations involving human participants may be susceptible to experimental mortality, and this did prove to be the case in that the total number of student participants was reduced from 84 to 66, since they could only be included if they fulfilled the pre- and post-intervention states. Additionally, weaknesses of both self-report survey data, and the use of focus groups have been outlined in Chapter 3, not least that there may be a gap between stated and actual behaviour (Bryman, 2016). However, the use of SPSS to analyse the surveys enabled the stability of student responses to be detected, and the students gave considered statements showing constancy throughout. This resonated with the view of Kamberelis and Dimitriadis (2005) that focus groups are valuable when participants feel they have ownership of the interview.

Overall however, limitations have been minimised by a methodological construction that purposefully drew upon the dualistic discourse of quantitative and qualitative research strategies to mine the qualities of both. A ‘macro’ quantitative approach provided ‘hard’ reliable number data and structural generalisations, and a contextualised qualitative ‘micro’ point of view incorporated rich, deep data in a natural and authentic setting (Bryman, 2016).

8.4 Recommendations for future studies

Geertz (1973) devised the term ‘thick description’ to denote detailed accounts of a social setting that can form the basis of general statements about a culture or its significance (Bryman, 2016). Lincoln and Guba (1985) argue that a thick description provides others with a ‘database for making judgements about the transferability (also comparability, *my insertion*) of findings to other milieu’ (Bryman, 2016:384). In Hargreaves’ (2013) longitudinal study on nine pupils’ perceptions of their teachers’ feedback, she asserted ‘the study can provoke in the minds of researchers, teachers and policy makers with an interest in AfL [*sic*], further scrutiny of existing traditions of feedback’ (p231). External validity consideration discussed in Chapter 3 explored whether the results of this study could be ‘generalised beyond the specific research context in which it was conducted’ (Bryman, 2016:691), and when one considers Bassey’s (1999) articulation of ‘fuzzy generalisations’ quoted in Chapter 3, there is reason to believe that this thesis comprises a ‘powerful and user-friendly summary which can serve as a guide to professional action’ (Bassey, 2001:5).

This study arose from an interest in encouraging students to consider post-16 physics as initially a teacher of physics in secondary school, later a professional development leader enabling non-physics-specialist science teachers to teach the subject with greater proficiency, and subsequently as a lecturer in science education and Initial Teacher Education. The value of positive ‘can-do’ language has long been important in facilitating these roles, as outlined by Ponchaud, Murphy and Whitelegg (2006), who argued for, amongst other factors, an expectation that anyone can do physics and for students to feel supported in their learning. If this study is indeed to serve as a guide to further action, there are a number of ways in which further questions might be explored.

One such question might seek to challenge the perceived supremacy of written feedback and study students' perceptions of whether they had learned more from their teachers' verbal or written feedback.

This study has raised an interesting dichotomy of the perceived values of 'negative' versus 'positive' feedback which has contradicted extant literature. It would be noteworthy to investigate how language plays a role in how this is received by students, and indeed where their preferences lie as to which type is privileged. A similar question would be to explore what ratios of prompt (discrepancy) feedback to statement (progress) feedback lent greater leverage to learning. Content analysis of extended teacher-student verbal feedback interactions could be considerably enhanced with subsequent teacher and student interviews exploring and analysing each of their positions and comprehension.

A natural extension to this study would be to attempt to replicate it involving Initial Teacher Education student teachers to firstly explore whether they are able to develop higher proportions of process and regulation feedback early in their teaching career, and to encourage them as teacher-researchers to explore themselves whether they were able to influence students' self-beliefs in the classroom. Additionally, might there be different outcomes with different age groups? The researcher is the lead for science education at both primary and secondary stages in a university ITE programme, and so might draw upon student teachers of diverse ages.

Since the focus of this study was an attempt to improve self-beliefs in physics, in which self-concept, self-efficacy and anxiety are considered critical issues, an extension to STEM subjects, and indeed other curriculum areas might generate further benefits. There would appear to be a case for questioning the prevalence, even ubiquity of 'growth mindset' as an embraced culture of many primary and secondary schools, and one wonders whether efforts directed towards self-concept, self-efficacy and anxiety self-beliefs may yield more fruit.

Finally, the paucity of literature considering feedback from the student perspective indicates a demand for an increased research provenance from there. If the complicated feedback landscape is to become better mapped, we need a greater understanding of how our recipients of feedback perceive and act upon it, since it is, after all, purportedly for their benefit. As Wiliam (2018:1) asserts: ‘feedback research is likely to be more effective if it places greater attention on the cognitive processes that are involved in learning (the micro level) and on the social situations within which feedback is given and received (the macro level).’

8.5 Personal reflections

Like many other PhD scholars, this has been a very long journey, both temporally and privately. I have moved from a background of Advanced Skills Teacher being involved with educational research at Leeds University with the exceptional and sadly missed Professor Philip Scott, to a Head of Faculty in a school, then on to experience being a Professional Development Leader at the National Science (now STEM) Learning Centre for many years. There I was able, with an amazing team of colleagues, to really understand how research can evidence-inform teacher practice and create professional development that can challenge and transform how teachers teach, and how their students learn. It seems a long time ago that I decided that my voice might have something to contribute, and that I might actually add something to the discourse via an original investigation. During the journey, I changed occupation (and houses) twice more, finally leading on science in Initial Teacher Education for the School of Education at York St John University at both primary and secondary level: a dream job. None of this journey would have been possible without the advice from a wise old teacher to ‘seize opportunities when they are presented to you, and never stop learning.’ My old friend, I am still learning every day.

To me, the greatest difficulty initially was *being* a researcher; I experienced imposter syndrome whenever methodology was mentioned. I may have moved from researcher novice to merely journeyman, but now I embrace the hard learning that comes with exploration rather than fearing it. There was much to recognise about personal resilience in a study which went seeking it. Consequently, I feel more able to support my own student teachers in this area and feel that I can influence their learning experiences positively. I find that I use more ‘can-do’ language to undergraduate and postgraduate students who have become removed from their prior science learning, and who report low science subject knowledge and high science anxiety. They report that they are developing an appreciation for science and a gradual belief that they will become better teachers of it; a positive step in reclaiming science as an impoverished core subject at primary level.

I have also learned much more about the complex and contested nature of feedback; through literature exploration and refining my own practice, I have been able to reconceptualise feedback form the perspective of the recipient rather than the donor. I intend to continue to contribute to the feedback discourse as a researcher, supporting colleagues and building capacity in my role as Learning and Teaching Lead for the School and of course, as a teacher-practitioner myself. As I learned more about the difficulties of flexing feedback, so we have been able to share with the student teachers at an earlier stage in their development. It has been rewarding to watch them absorb with what this implies for their own practice. Some have exhibited a desire to record themselves to analyse their feedback interactions, and it makes me both proud and humble that the insights in this research have enabled the next teacher generation to engage with reflective practice so early in their career.

Appendices

Appendix 1. The Observation Schedule

Teacher Number: ______ **Date Observation:** ______ **Focus/Context:** ______________________________

Class reference: ______

Time interval (min)	Feedback level, type & recipient	Field comments
0.00		
1.00		
2.00		
3.00		
4.00		
5.00		
6.00		
7.00		
8.00		
9.00		
10.00		
11.00		
12.00		
13.00		
14.00		
15.00		
16.00		

Event Recording System v1.4

Appendix 2. The Survey Tool

Physics Questionnaire

All information supplied will be kept strictly confidential

Name:		Date:
Age:	Group:	Male/Female (circle one)

PLEASE READ THESE INSTRUCTIONS FIRST

This is not a test – there are no right or wrong answers, only what you feel about Physics

Read each sentence and decide how much you agree with it, then tick <u>the most relevant box</u> in each row. Please don't tick on a line.

	Strongly agree	**Agree**	**Disagree**	**Strongly Disagree**
1. I learn things in Physics quickly.				
2. Your intelligence is something about you that you can't change very much.				
3. I often worry that it will be difficult for me in Physics classes.				
4. I can help classmates with Physics problems.				
5. You can learn new things, but you can't really change your basic intelligence.				
6. In my Physics class, I understand even the most difficult work.				
7. I worry that I will get poor grades in Physics.				
8. I can use calculations to work things out in Physics.				
9. I can compare and contrast conduction and convection in heat transfer				

Please turn over

	Strongly agree	Agree	Disagree	Strongly Disagree
10. I feel helpless when doing a Physics problem.				
11. No matter who you are, you can change your intelligence a lot.				
12. I get good grades in Physics.				
13. I get very tense when I have to do Physics homework.				
14. I can calculate the speed of an object.				
15. You have a certain amount of intelligence and you can't do much to change it.				
16. I have always done well in Physics.				
17. I get very nervous doing Physics problems.				
18. I can interpret information in graphs to describe something in Physics.				
19. I am just not good at Physics.				
20. I can use Physics concepts to explain real-life situations.				

Thank you for your time!

*Question 9 for the other IG cohort was ‘I can compare the different types of waves used for communication’

Appendix 3. Feedback Continuing Professional Development Intervention

Flexing Teacher Verbal Feedback

Learning Outcomes

- to become aware of the different self-beliefs students possess
- to appreciate what the different levels of feedback are;
- to consider where these different levels of feedback are focused;
- to identify how students receive, comprehend and use feedback in their learning

What is feedback?

"*The notion of feedback is fascinating – whilst feedback is referenced in almost all studies of learning, it sometimes seems as if it is a unidimensional notion understood by all.*"

Hattie & Masters, 2012

What do you think feedback is?

Create a definition of feedback for how you use it in your class.

Which of these is feedback?

"Nice job on the project, Sheshona!"

"Next time, Sam, you'll want to make your thesis clearer to the reader"

"The lesson would be more effective, Shana, if your visuals were more polished and supportive of the teaching."

"You taught about ants, Stefan? I LOVE ants!"

None of these is feedback.

http://grantwiggins.wordpress.com/2014/04/15/what-feedback-is-and-isnt/

Feedback?

"Nice job on the project, Sheshona! *You answered the essential question in great depth, with lots of illustrative examples, and your oral presentation was polished and informative.*"

"*I found it very difficult to grasp your main point. At the start, it seemed that you were arguing against mining coal, but in paragraph three you focused on the need to provide healthcare to all workers.* Next time, Sam, you'll want to make your thesis clearer to the reader"

"*Your spoken delivery was clear and your account of the topic was a helpful and interesting summary: most students were engaged. Alas, the supporting materials you supplied looked unfinished and rough; 5-6 students were confused by them.* The lesson would be more effective, Shana, if your visuals were more polished and supportive of the teaching."

"You taught about ants, Stefan? I LOVE ants! *However, the task was not to please me; the task was to make students ignorant or afraid of an animal to become interested in them. Yet, you began as if the students already shared your interest in ants instead of helping them overcome their distaste and become more interested in them.*"

When feedback is only half there.

Fifth graders were busy writing acrostic poems on small posters. One girl wrote a school spirit poem, with the first letter of each line spelling out the school name: S for "super," N for "nice," and so on. She even drew a picture of a bobcat (pictures were not required) that was a spot-on replication of the school mascot.

A boy wrote an acrostic poem with the first letter of each line spelling out his name: A for "agressive" (unfortunately spelled incorrectly); N for "nutty"; and so on. No picture.

So what kind of feedback did the teacher give?

Her comments gave students the impression that the girl's poem was perfect and that the boy's poem was not so good, mostly because of that one misspelled word and the fact that his lines sloped downward on the poster.

The girl's work was a skilful replication of things she'd seen before. All the words were simple, the school spirit theme was a common one, and the point of her drawing was to duplicate the school mascot.

This assignment was a giant missed opportunity for both students.

(Brookhart, 2008)

Why wasn't it quite enough?

The girl needed to know that her work was proficient—but she also needed to be challenged to work with more originality when writing poems. She only received half that feedback.

The boy's work was more original. Although the poem was only five lines long, it gave readers a real sense of who he was—or, at least, how he saw himself. He needed to know that he had used a prescribed format creatively—but he also needed to be challenged to check his spelling and use a ruler to make straight lines of text on posters. He, too, only received half that feedback.

Consider the impact on these students' learning and self-beliefs. What might be the consequences?

Calibration of feedback to suit student's learning need in that instance.

Position within AfL paradigm

1. Clarifying and sharing learning intentions and criteria for success.
2. Engineering effective classroom discussions, questions, and learning tasks.
3. Providing feedback that moves learners forward.
4. Activating students as the owners of their own learning.
5. Activating students as instructional resources for one another.

How do students view success and failure?

Carol Dweck & her colleagues found 3 themes running through student responses:

1. **s/f due to internal or external factors, eg how it was *personalised.***
2. **s/f due to factors that were transient or long-lasting, ie the *stability***
3. **s/f due to the *specificity* of the attribution, eg some students over generalise their success or failure = *global*)**

Self-Theories (2000)

Wiliam (2011) *Embedded Assessment.*

Attribution	Success	Failure
Personalisation	*internal*: I got a good grade because it was a good piece of work *external*: I got a good grade because the teacher likes me	*internal*: I got a low grade because it wasn't a good piece of work *external*: I got a low grade because the teacher doesn't like me
Stability	*stable*: I got a good grade because I'm good at that subject *unstable*: I got a good grade because I was lucky in the questions that came up	*stable*: I got a bad grade because I'm no good at that subject *unstable*: I got a bad grade because I hadn't reviewed the material before the test
Specificity	*specific*: I'm good at that, but that's the only thing I'm good at *global*: I'm good at that, so I'll be good at everything	*specific*: I'm no good at that, but I'm good at everything else *global*: I'm useless at everything

What do you think about 'intelligence'?

"You have a certain amount on intelligence and you can't really do much to change it."

"You can learn new things, but you can't really change your basic intelligence."

"As much as I hate to admit it, you can't teach an old dog new tricks"

"Your intelligence is something about you that you can't change very much"

(And so on…measured on a Likert scale)

Questions are adapted from Dweck, C., (2000) Self-Theories – Their Role in Motivation, Personality and Development. Taylor & Francis Group: Philadelphia.

Notion of 'mindsets'

Educational psychologist Carol Dweck

Entity Theory (Fixed Mindset): Intelligence is a fixed trait or entity that you are born with and cannot change.

Incremental Theory (Growth Mindset): Intelligence is a malleable quality; a potential that can be developed

Why do people react so differently?

Some people think that failure measures you.
Some think that it makes you.

Some students regard achievement situations as a test of their intelligence. They have performance goals.

Others regard the same situation as an opportunity to learn new things. They have learning goals.

Take a look at how some people were found to react to feedback…

What happens when you give goal feedback?

The feedback you intend in not necessarily internalised in the way it has been intended…

Student response to feedback	feedback indicates performance exceeds goal	feedback indicates performance falls short of goal
change their behaviour	exert less effort	*increase effort*
change their goal	*increase aspiration*	reduce aspiration
abandon the goal	decide goal is too easy	decide goal is too hard
reject the feedback	ignore feedback	ignore feedback

How do you feel about challenges?

Interestingly, we do not often know what people are until they face a challenge.

For it is in their response to failure that their attitude to their intelligence is often revealed. Many skilled students ironically blame their ability when they fail at a task.

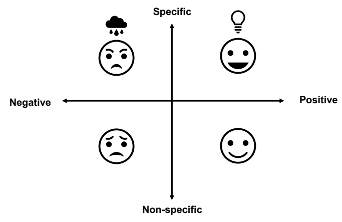
If they are confident in their ability to achieve within the task, they will attempt it. If not, they very often 'can't be bothered' and disengage – seen in classrooms across the country.

Feedback in these situations is then often viewed as rubbing salt in the wounds.

Problems with praise?

1. You must be a natural
2. That was brilliantly done
3. You did that so quickly
4. Good job!
5. Praising students who get 10/10 on a task or test.
6. 'High ability' students praised for being clever
7. 'Special needs' students given extra praise to boost their self esteem

What's the best feedback?



	Hattie & Masters (2011)	Gan (2011)	Van den Bergh, Rose & Beijaard (2013)
Level	18 secondary classes	235 peers	32 teachers in primary school
Task	59%	70%	51%
Process	25%	25%	42%
Self-Regulation	2%	1%	2%
Self	14%	4%	5%

What teachers see as feedback

- Comments
- Clarification
- Confirmation
- Criticism (pros & cons)
- Content development
- Correction

[All of these tend to relate to how the student is doing now]

What students see as feedback

- Give advice on what to do next
- How to improve
- What to do better next time but also what you are doing well
- "you do something and the teacher tells you whether you are going in the right direction"
- How to make adjustments to make it better
- To have time to make it better

[all of these are to do with 'where next']

Student perceptions of feedback…

Feedback is:

•too late

•too vague

•too cryptic

•inconsistent

•didn't relate to learning criteria

•demotivational

Practising flexing your verbal feedback

Researcher to scribble a shape and then turn it into a picture.

Your job is to encourage the drawing of the developing picture, using process or self-regulatory verbal feedback.

Bookmark of helpful prompts available!

Three Fundamental Principles of Feedback

Feedback should cause thinking

1. Be more work for the *recipient* than the donor. Students should be expected to have time to do something with it.

2. Be focused: 'less' can be more. Feedback should be accurate *and* helpful

3. Relate to shared learning goals (not performance goals). Verbal feedback flexed so that it is calibrated for the student's learning need in that instance.

[Adapted from Dylan Wiliam Embedded Formative Assessment (2011)]

Practical strategies…

- Involving the students in setting and understanding criteria
- Conscious use of problem-solving strategies
- Represents skills as acquirable
- Recognising & encouraging effort and persistent behaviours
- Avoiding labelling
- Modelling beliefs in the potential for success
- Not shying away from challenge

Effects on different groups…

'Strategies' training:

Low Achievers – emphasises that effort and strategies are the way to become more successful and overcoming failures.

High Attainers – Encourages them to seek challenges, react better to setbacks, avoid feelings of helplessness and not give up.

High Achieving Girls – Same as above plus it makes them less vulnerable to underachievement, increases self- belief in their ability

Dylan Wiliam

"an assessment functions formatively to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have made in the absence of that evidence"

Embedded Formative Assessment (p.43).

Appendix 4. The Support Bookmark

Process and Self-regulation
statements

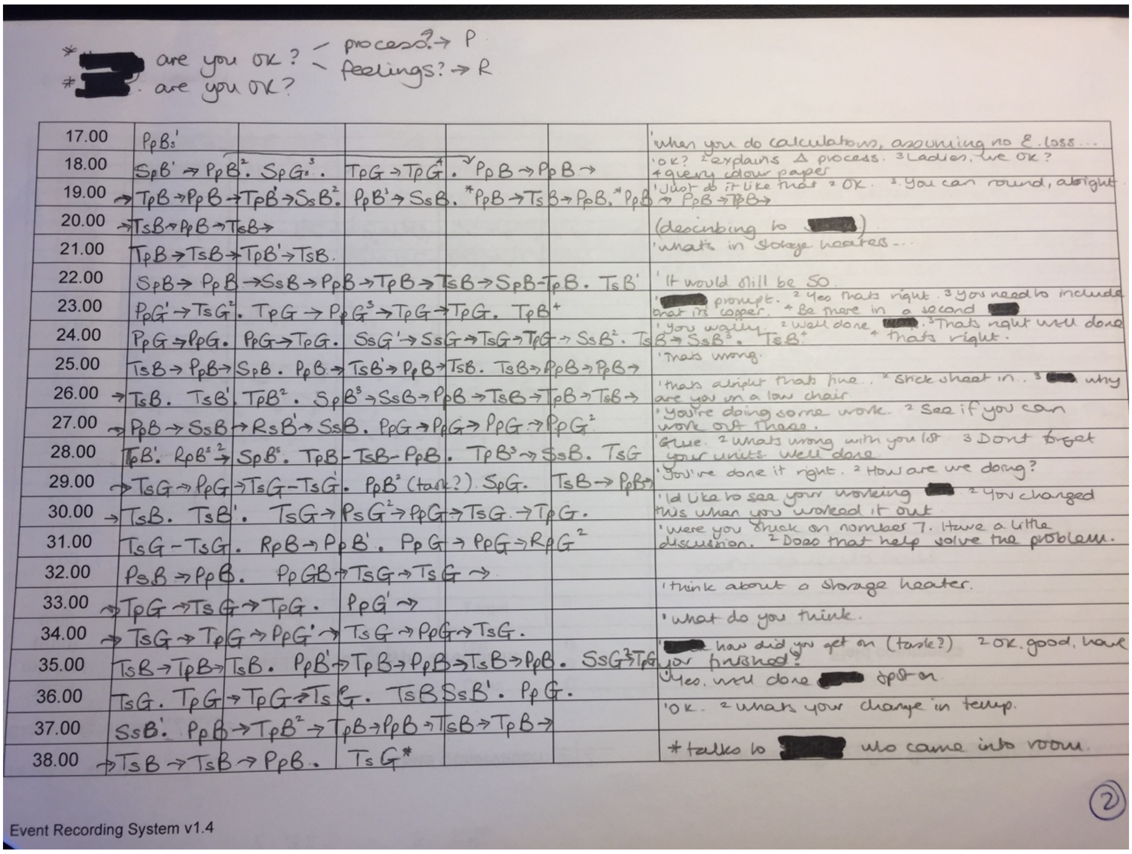
You interpreted/explained that well
You applied your knowledge well there
That was a good analysis
That was a good strategy to use
Clear summary of the method
You communicated well with each other
Excellent modelling of skills there
That was good effort there
I liked the way you kept on at that
You worked really hard at that…
I was impressed with how hard you worked
You've put a lot of thought into how that would….
Good reflections on your learning!
You really persevered with …
Well done for carrying on when you found it challenging

Process and Self-regulation
prompts

What would happen if…?
What was your starting point?
How could you break that down?
What else could you add?
How could you develop this further?
What more would you like to find out?
What does this lead onto?
How did you do it/create it?
How do you know that?
What kind of data would support…?
What is this an example of?
What is another way to explain…?
Tell me more about…
What were the reasons you chose to do it that way?
Did you change anything?
Why is that important?
How does…affect…?

@bloom_growhow

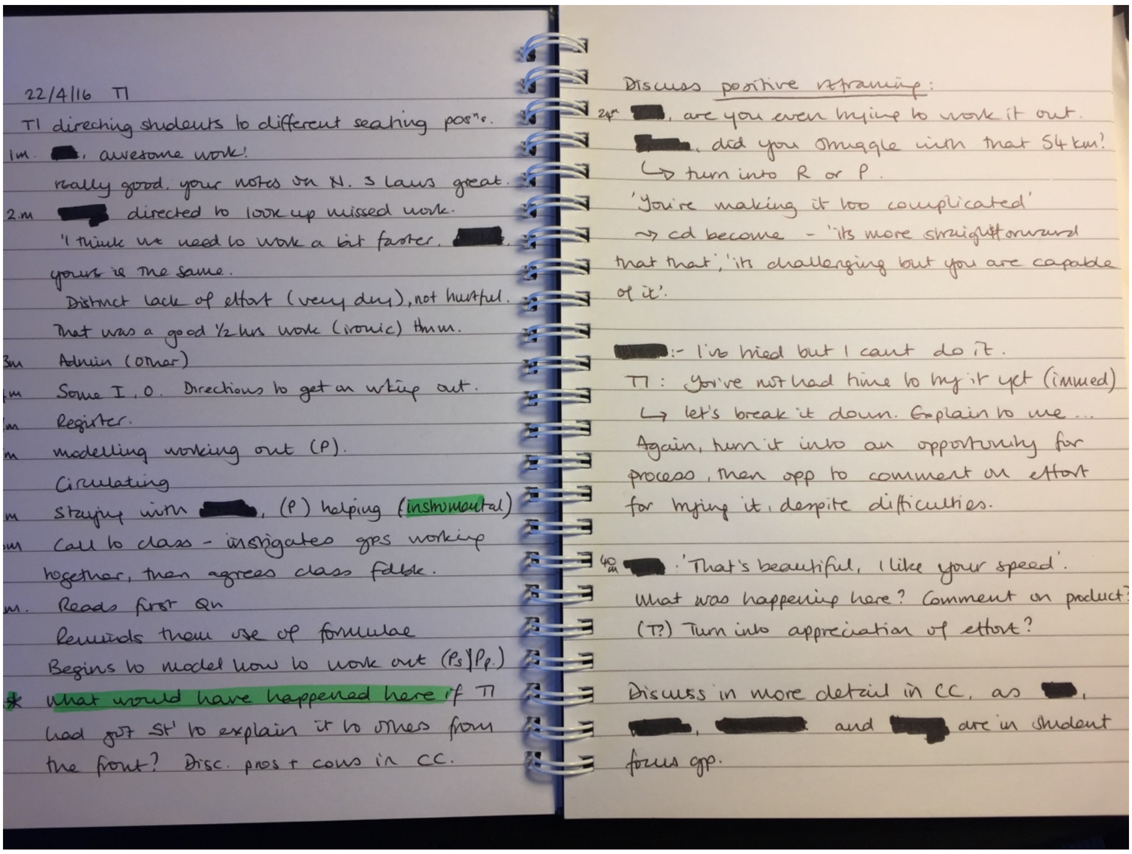
Appendix 5. Example of field notes



* [redacted] are you OK? – process? → P
\# [redacted] are you OK? – feelings? → R

17.00	PpB¹	'when you do calculations, assuming no E loss…'
18.00	SpB' → PpB². SpG³. TpG → TpG. PpB → PpB →	'OK? ²explains Δ process. ³Ladies, we OK? ⁴query about paper
19.00	→ TpB → PpB → TpB' → SsB². PpB' → SsB. *PpB → TsB → PpB. *PpB → PpB → TpB →	'Just do it like that ²OK. ³You can round, alright
20.00	→ TsB → PpB → TsB →	(describing to [redacted])
21.00	TpB → TsB → TpB' → TsB.	'what's in storage heaters…'
22.00	SpB → PpB → SsB → PpB → TpB → TsB → SpB – TpB. TsB'	'it would still be 50.
23.00	PpG' → TsG'. TpG → PpG³ → TpG → TpG. TpB⁺	[redacted] prompt. ²Yes that's right ³you need to include that it's copper. ⁴Be there in a second [redacted]
24.00	PpG → PpG. PpG → TpG. SsG' → SsG → TsG → TpG → SsB². TsB → SsB³. TsB⁺	'you wally. ²well done [redacted]. ³That's right well done ⁴that's right.
25.00	TsB → PpB → SpB. PpB → TsB' → PpB → TsB. TsB → PpB → PpB →	'That's wrong.
26.00	→ TsB. TsB'. TpB². SpB³ → SsB → PpB → TsB → TpB → TsB →	'that's alright that's fine. ²stick sheet in. ³[redacted] why are you in a low chair
27.00	→ PpB → SsB → RsB' → SsB. PpG → PpG → PpG → PpG²	'you're doing some work. ²See if you can work out these.
28.00	TpB'. RpB² → SpB'. TpB – TsB – PpB. TpB³ → SsB. TsG	'Glue. ²what's wrong with you lot ³Don't forget your units. well done
29.00	→ TsG → PpG → TsG – TsG. PpB² (task?) SpG. TsB → PpB	'you've done it right. ²How are we doing?
30.00	→ TsB. TsB'. TsG → PsG² → PpG → TsG → TpG.	'I'd like to see your working [redacted] ²You changed this when you worked it out
31.00	TsG – TsG. RpB → PpB'. PpG → PpG → RpG²	'were you stuck on number 7. Have a little discussion. ²Does that help solve the problem.
32.00	PsB → PpB. PpGB → TsG → TsG →	'think about a storage heater.
33.00	→ TpG → TsG → TpG. PpG' →	'what do you think.
34.00	→ TsG → TpG → PpG' → TsG → PpG → TsG.	[redacted] how did you get on (task?) ²OK, good, have you finished?
35.00	TsB → TpB → TsB. PpB' → TpB → PpB → TsB → PpB. SsG → TpG	'Yes, well done [redacted] spot on
36.00	TsG. TpG → TpG → TsG. TsB SsB'. PpG.	'OK. ²what's your change in temp.
37.00	SsB'. PpB → TpB² → TpB → PpB → TsB → TpB →	
38.00	→ TsB → TsB → PpB. TsG*	* talks to [redacted] who came into room

Event Recording System v1.4



22/4/16 TI

TI directing students to different seating pos^ns.
1m. [redacted], awesome work!
really good, your notes on N. 3 Laws great.
2m [redacted] directed to look up missed work.
'I think we need to work a bit faster. [redacted] yours is the same.
Distinct lack of effort (very dry), not hurtful.
That was a good ½ hrs work (ironic) thnm.
3m Admin (other)
m Some I.O. Directions to get a whip out.
m Register.
m modelling working out (P).
Circulating
m staying with [redacted], (P) helping (instrumental)
m call to class – instigates grps working together, then agrees class fdbk.
m. Reads first Qn
Reminds them use of formulae
Begins to model how to work out (Ps/Pp)
* what would have happened here if TI had got St' to explain it to others from the front? Disc. pros + cons in CC.

Discuss positive reframing:
24m [redacted], are you even trying to work it out.
[redacted], did you struggle with that 54 km?
↳ turn into R or P.
'You're making it too complicated'
→ cd become – 'it's more straightforward that that', 'it's challenging but you are capable of it'.

[redacted] :- I've tried but I can't do it.
TI: you've not had time to try it yet (immed)
↳ let's break it down. Explain to me…
Again, turn it into an opportunity for process, then opp to comment on effort for trying it, despite difficulties.

40m [redacted]. 'That's beautiful, I like your speed'.
What was happening here? Comment on product?
(TI) Turn into appreciation of effort?

Discuss in more detail in CC, as [redacted], [redacted] and [redacted] are in student focus grp.

Appendix 6. Student Semi-structured Focus Group Questions

Pre-intervention questions

Do you ever get feedback on your physics work?

What sort of feedback does your teacher give you about your physics work?

(Is that school policy?)

Does the feedback you get tend to relate to the task you have done, or the learning intentions of the lesson?

Do you work to success criteria?

When you get feedback, what do you do with it?

What about feedback in the lesson?

When your teacher gives you this feedback, how does it make you feel about your physics?

Does it matter whether the feedback is positive or negative?

How do you do at Physics?

[Probe – Triple Science]

self-concept question [How do you think you do at physics?]

self-efficacy question [do you know how to <complete task related to recent topic>]?

Post-intervention questions

Do you ever get feedback on your physics work?

What sort of feedback does your teacher give you about your physics work?

When you get feedback, what do you do with it?

What about feedback in the lesson?

When your teacher gives you this feedback, how does it make you feel about your physics?

Does it matter whether the feedback is positive or negative?

How do you do at Physics?

When you get good grades in Physics, why is that?

Recent physics exam – how did you do?

Probe: [could you have done better? How?]

Possibilities?

- I didn’t study hard enough.
- I didn’t go about studying in the right way.
- I wasn’t smart enough.
- The test was unfair

Outcomes?

- I’d spend less time on this subject from now on
- I would work harder in this class from now on.
- I would spend more time studying for the test
- I would try not to take this subject ever again

If have time, resilience questions

- An important reason why I do my schoolwork is because I like to learn new things.
- I like schoolwork best when it makes me think hard.
- I like schoolwork that I’ll learn from even if I make a lot of mistakes.
- To tell the truth, when I work hard at my schoolwork, it makes me feel like I’m not very smart.
- If you’re not good at a subject working hard won’t make you good at it.
- The harder you work at something, the better you’ll be at it.

Appendix 7. Ethical Approval from the University Ethics Committee

Performance, Governance and Operations
Research & Innovation Service
Charles Thackrah Building
101 Clarendon Road
Leeds LS2 9LJ Tel: 0113 343 4873
Email: ResearchEthics@leeds.ac.uk

UNIVERSITY OF LEEDS

Kathryn Bloom Milner
School of Education
University of Leeds
Leeds, LS2 9JT

**ESSL, Environment and LUBS (AREA) Faculty Research Ethics Committee
University of Leeds**

1 December 2015

Dear Katy

Title of study: **The role of self-regulation feedback on students' self-belief systems in Physics**
Ethics reference: **AREA 13-103**

I am pleased to inform you that the above research application has been reviewed by the ESSL, Environment and LUBS (AREA) Faculty Research Ethics Committee and following receipt of your response to the Committee's comments, I can confirm a favourable ethical opinion as of the date of this letter. The following documentation was considered:

Document	Version	Date
AREA 13-103 Ethical review application checklist.docx	2	04/04/14
AREA 13-103 Ethical_Review_Form_V3KB doc.docx	2	04/04/14
AREA 13-103 fieldwork-assessment-form-low-risk-2013.doc	1	04/04/14
AREA 13-103 KB subject consent form - Parents.doc	2	30/11/15
AREA 13-103 KB subject consent form - students.doc	3	30/11/15
AREA 13-103 KB subject consent form - teachers.doc	3	30/11/15
AREA 13-103 Physics Beliefs Infm Sheet - Parents.doc	2	30/11/15
AREA 13-103 Physics Beliefs Infm Sheet - Student.doc	2	30/11/15
AREA 13-103 Physics Beliefs Infm Sheet - Teacher.doc	2	30/11/15
AREA 13-103 Physics Beliefs Infm Sheet.doc	1	19/04/14

Please notify the committee if you intend to make any amendments to the original research as submitted at date of this approval, including changes to recruitment methodology. All changes must receive ethical approval prior to implementation. The amendment form is available at http://ris.leeds.ac.uk/EthicsAmendment.

Please note: You are expected to keep a record of all your approved documentation, as well as documents such as sample consent forms, and other documents relating to the study. This should be kept in your study file, which should be readily available for audit purposes. You will be given a two week notice period if your project is to be audited. There is a checklist listing examples of documents to be kept which is available at http://ris.leeds.ac.uk/EthicsAudits.

Yours sincerely

Jennifer Blaikie
Senior Research Ethics Administrator, Research & Innovation Service
On behalf of Dr Andrew Evans, Chair, AREA Faculty Research Ethics Committee

CC: Student's supervisor(s)

Appendix 8. Student Information Letter

Faculty of Education, Social Sciences and Law

UNIVERSITY OF LEEDS

School Information for Physics Research Project

Dear Student

I am writing to ask if you would like to be involved in a University of Leeds research project on young peoples' attitudes to Physics. This project will be conducted at school over the next few months. Physics is a subject which can be a considerable advantage in the jobs market, and is a preferred subject for entry to many universities. But its uptake at A level is uneven: it's the 4th most popular choice for boys, but the 19th for girls. Yet the girls that take it achieve really well.

A project involves doing a short survey which aims measure your attitudes to, and beliefs about physics at the start and end of the study. This survey will have a space to put your name at the top, in case the researcher would like to talk to you about any aspect, but no-one else will see what you have answered on the survey. You may be asked to do a short interview (10-15 minutes) at lunchtime with the researcher after this survey, which will be audio-recorded but not videoed. Then the researcher may be present in your physics class occasionally, doing observations on your teacher over a period of 4-6 months. Sometimes the researcher may ask if she can focus on you during an observation to see how you respond to the teacher's comments. The lessons will not be videoed, but the teacher will be wearing a microphone so that the researcher can check what has been said afterwards. The researcher will be writing observations of what the teacher has said on sheet during the lesson. No student names will be noted by the researcher. Towards the end of the project, you'll be asked to do the survey again, and another interview if you did one before, again audio-recorded. There shouldn't be any disruption to your learning, and the only lesson time involved will be doing the survey, for about 15 minutes at the start and end of the project.

If you take part, any data collected about you is kept absolutely private, and does not get shared with school. When the data has been analysed, a report will be made, which will be shared with you, and this will have information in it, based on the results of groups of students, who do not get named. Only Year 10 students whose parents agree, and who have themselves agreed to participate will be involved in the study. You and/or your parents can withdraw permission at any time during the study by telling the researcher that you don't want to take part any more. Any audio-recordings and surveys belonging to a student who chooses to withdraw will be removed and destroyed. Statistical information provided by their survey responses would remain in the analysed data, but would be unidentifiable. There are no known or anticipated risks to participation in this study, which has ethics clearance by the University of Leeds. However, the final decision about taking part is up to you.

Faculty of Education, Social Sciences and Law

UNIVERSITY OF LEEDS

We would appreciate your agreeing to participate in this project, as we believe it will contribute to furthering our knowledge of students' learning in physics. Please complete the Student Consent forms for both observations and interviews, and return it to school by [Date two weeks from receiving]

Yours faithfully

Katy Bloom
Associate Principal Lecturer
Science Lead Tutor
Institute of Childhood and Education
Leeds Trinity University

Professor Jim Ryder
Department of Education
Faculty of Education, Social Sciences and Law
University of Leeds

Appendix 9. Student Consent Form (Research and Interview)

UNIVERSITY OF LEEDS

Physics Education Research Project

Student Observation Consent Form

	Tick to Agree
I have been told enough about the project	
I understand how I will be involved in the project	
I understand that I will not be named in project reports	
I understand that nobody in school will be told what I say	
I understand that I can drop out at any time without giving a reason	
I agree to take part in this project	

Student

Name: ....

Signature: ....

Date: ....

Researcher

Name: ...Katy Bloom...

Signature: ... KBloom ...

Date: ...date you send...

UNIVERSITY OF LEEDS

Physics Education Research Project

Student Interview Consent Form

	Tick to Agree
I have been told enough about the project	
I understand how I will be involved in the project	
I understand that I will not be named in project reports	
I understand that nobody in school will be told what I say	
I understand that I can drop out at any time without giving a reason	
I agree to take part in this project	

Student

Name: …………………

Signature: …………………

Date: …………………

Researcher

Name: …Katy Bloom…

Signature: … ……

Date: …date you send

Appendix 10. Guardian Information Letter

Faculty of Education, Social Sciences and Law

UNIVERSITY OF LEEDS

School Information for Physics Research Project

Dear Parent/Guardian

I am writing to ask your permission for your child to participate in a University of Leeds research project on young peoples' attitudes to Physics. This project will be conducted at school over the next few months. Physics is a subject which can confer a considerable advantage in the jobs market, and is one of the preferred subjects for entry to the Russell Group of Universities, yet it's uptake post-16, particularly for girls, remains low, despite female achievement in the subject.

A short survey looking to measure aspects of their attitudes to, and beliefs about physics will be administered at the start and close of the study. This survey has a space for the student to enter their name; this is so that the researcher can follow up individual responses with some students if necessary, and analyse responses, however named responses will never be viewed by other people, and will be kept strictly private. Classroom observations will capture information about how the teacher interacts with the students regarding their physics learning. The focus is on the teacher not individual students, and the teacher will be wearing a microphone and audio-recorded so that their interactions can be checked and analysed subsequently. No student names if used will be noted by the researcher.

It is possible that your child would be asked to meet with the researcher for a short lunchtime interview to gather more detailed data about their beliefs. This would be audio-recorded, but not videoed. The researcher may ask your child for permission to observe them during their physics lesson. In total, your child will be asked to do a survey, possibly be interviewed, be in a class where the teacher is being observed over a period of 4-6 months, potentially be the subject of an observation themselves, do another survey, and possibly be interviewed again. The project in which your child is invited to participate has however been designed to minimise disruption to their learning. The researcher is cleared to work with children, and is an educational practitioner at a University.

All students' data are considered confidential, and individual students' data will not be shared with staff at the school. However, information based on the results of a group of participants may be shared. Only year 10 students who have parental permission, and who have themselves agreed to participate will be involved in the study. The child, and/or the parent may withdraw their permission at any time during the study by so indicating to the researcher. Should a student withdraw, their survey and interview data, if they did one, will be removed and destroyed. Statistical information provided by the analysis of their survey would remain integrated in the whole sample however any individual would be unidentifiable.

There are no known or anticipated risks to participation in this study. This research project has been reviewed and received clearance by the University of Leeds Research Ethics Committee. However, the final decision about participation is yours.

We would appreciate your permission for your child to participate in this project, as we believe it will contribute to furthering our knowledge of students' learning in physics. Please complete the Parent/Guardian Consent form, and return it to school by [Date two weeks from receiving]

Faculty of Education, Social Sciences and Law

UNIVERSITY OF LEEDS

Yours faithfully

Katy Bloom
Associate Principal Lecturer
Science Lead Tutor
Institute of Childhood and Education
Leeds Trinity University

Professor Jim Ryder
Department of Education
Faculty of Education, Social Sciences and Law
University of Leeds

Appendix 11. Guardian Consent Form

UNIVERSITY OF LEEDS

Physics Education Research Project

Parent/Guardian Consent Form

	Tick to Agree
I have been told enough about the project	
I understand how my child will be involved in the project	
I understand that my child will not be named in project reports	
I understand that nobody in school will be told what my child says	
I understand that I can ask my child to be withdrawn from the project at any time without giving a reason	
I agree to my child taking part in this project	

Parent/Guardian

Name: ....

Signature: ....

Date: ....

Researcher

Name: Katy Bloom....

Signature: ... KBloom ...

Date: ...date...

Appendix 12. Teacher Information Letter

Faculty of Education, Social Sciences and Law

UNIVERSITY OF LEEDS

School Information for Physics Research Project

Dear Teacher

I am writing to ask if you would like to take part in a University of Leeds research project on young peoples' attitudes to Physics. This project will be conducted at school over the next few months. Physics is a subject which can confer a considerable advantage in the jobs market, and is one of the preferred subjects for entry to the Russell Group of Universities, yet it's uptake post-16, particularly for girls, remains low, despite female achievement in the subject.

A short survey looking to measure aspects of student attitudes to, and beliefs about physics will be administered at the start and close of the study. Classroom observations will capture information about how the teacher interacts with the students regarding their physics learning. The researcher is interested in how feedback is used in physics teaching and learning. Some teachers will be asked to take part in some continuing professional development on feedback practices, and their practice observed to see what impact it has, both in terms of their own practice, and the learning of their students. During these observations, the researcher will ask if the teacher agrees to be ‘miked up’ so that quieter interactions with individuals may still be captured. The researcher would also like to observe the classroom interaction practice of teachers who do not undergo the professional development, so that enacted school policies of feedback can be built into later comparison and analysis.

Some students will be asked to meet with the researcher for a short lunchtime interview to gather more detailed data about their beliefs, which will be audio-recorded, and the researcher may ask them for permission to focus the observation on them during their physics lesson. However, apart from 10-15 minutes to complete the surveys at the start and end of the project, class time won’t be expended on the project, as it has been designed to minimise disruption to learning. In total, teachers taking part in the intervention will be observed a maximum of 3-4 times every month, and undergo some professional development of about 30-60 minutes at the start of the project, and have ‘coaching conversations’ at subsequent two-month intervals. The study will take place over a period of 4-6 months. Teachers not undertaking the professional development will administer the survey at the start and end of the project, and be the subject of 3-4 classroom observations only at the start of the study, for comparison purposes.

The researcher is cleared to work with children, and is an educational practitioner at a University. All students’ data are considered confidential, and individual students’ data will not be shared with staff at the school. However, information based on the results of a group of participants may be shared. All student, teacher and school data will be anonymous. Only year 10 students who have parental permission, and who have themselves agreed to participate will be involved in the study. The student, and/or the parent may withdraw their permission at any time during the study by so indicating to the researcher, as may any of the teachers. Should any student participant wish to withdraw, depending on the time at which it takes place, their data will be removed and destroyed. Statistical information from the survey analysis would have to remain included but would be non-identifiable. Should the teacher participant wish to withdraw at any time, the data from that school will not be used.

Faculty of Education, Social Sciences and Law

UNIVERSITY OF LEEDS

There are no known or anticipated risks to participation in this study. This research project has been reviewed and received clearance by the University of Leeds Research Ethics Committee. However, the final decision about participation is yours.

We would appreciate your agreeing to participate in this project, as we believe it will contribute to furthering our knowledge of students' learning in physics. Please complete the Teacher Consent form, and return it to Katy Bloom by email

Yours faithfully

Katy Bloom
k.bloom@leedstrinity.ac.uk
Associate Principal Lecturer
Science Lead Tutor
Institute of Childhood and Education
Leeds Trinity University

Professor Jim Ryder
Department of Education
Faculty of Education, Social Sciences and Law
University of Leeds

Appendix 13. Teacher Consent Form

UNIVERSITY OF LEEDS

Physics Education Research Project

Teacher Consent Form

	Tick to Agree
I have received enough information about the project	
I understand how I will be involved in the project	
I understand that the data obtained will be held in confidence and that if it is presented or published any personal details will be removed	
I understand that I am free to withdraw from the project at any time without giving any reason	
I agree to take part in this research project	

Teacher

Name: ……………………

Signature: ……………………

Date: ……………………

Researcher

Name: ……Katy Bloom……

Signature: … KBloom ……

Date: ……16/12/15……

Appendix 14. A broad comparison of IG/CG pre- and post-intervention survey outcomes.

		IG sample difference	Similarity in samples	CG sample difference
1	1. I learn things in Physics quickly.	Slightly higher self-concept		
2	2. Your intelligence is something about you that you can't change very much.		Similar mindset	
3	3. I often worry that it will be difficult for me in Physics classes..		Similar anxiety	
4	4. I can help classmates with Physics problems.			Slightly higher self-efficacy
5	5. You can learn new things, but you can't really change your basic intelligence.		Similar mindset	
6	6. In my Physics class, I understand even the most difficult work.		Similar self-concept	
7	7. I worry that I will get poor grades in Physics.	Slightly lower anxiety		
8	8. I can use calculations to work things out in Physics	Higher self-efficacy		
9	9. I can compare and contrast conduction and convection in heat transfer	Higher self-efficacy		
10	10. I feel helpless when doing a Physics problem		Similar anxiety	
11	11. No matter who you are, you can change your intelligence a lot	Less fixed mindset		
12	12. I get good grades in Physics.		Similar self-concept	
13	13. I get very tense when I have to do Physics homework	Slightly lower anxiety		
14	14. I can calculate the speed of an object	Higher self-efficacy		
15	15. You have a certain amount of intelligence and you can't do much to change it.	Less fixed mindset		
16	16. I have always done well in Physics			Higher self-concept
17	17. I get very nervous doing Physics problems.	Slightly lower anxiety		
18	18. I can interpret information in graphs to describe something in Physics	Slightly higher self-efficacy		
19 (-'ve)	19. I am just not good at Physics.		Similar self-concept	
20	20. I can use Physics concepts to explain real-life situations.		Similar self-efficacy	

mindset
anxiety
self-efficacy
self-concept

Appendix 14a.Table showing a broad IG and CG comparison of self-belief attributes from the **pre-**intervention surveys.

		IG sample difference	Similarity in samples	CG sample difference
1	1. I learn things in Physics quickly.	Higher self-concept		
2	2. Your intelligence is something about you that you can't change very much.		Similar mindset	
3	3. I often worry that it will be difficult for me in Physics classes..		Similar anxiety	
4	4. I can help classmates with Physics problems.	Higher self-efficacy		
5	5..You can learn new things, but you can't really change your basic intelligence.	Less fixed mindset		
6	6. In my Physics class, I understand even the most difficult work.	Higher self-concept		
7	7. I worry that I will get poor grades in Physics.	Lower anxiety		
8	8. I can use calculations to work things out in Physics	Higher self-efficacy		
9	9. I can compare and contrast conduction and convection in heat transfer	Higher self-efficacy		
10	10. I feel helpless when doing a Physics problem	Lower anxiety		
11	11. No matter who you are, you can change your intelligence a lot			Less fixed mindset
12	12. I get good grades in Physics.	Higher self-concept		
13	13. I get very tense when I have to do Physics homework	Lower anxiety		
14	14. I can calculate the speed of an object			Higher self-efficacy
15	15. You have a certain amount of intelligence and you can't do much to change it.		Similar mindset	
16	16. I have always done well in Physics			Higher self-concept
17	17. I get very nervous doing Physics problems.	Lower anxiety		
18	18. I can interpret information in graphs to describe something in Physics	Higher self-efficacy		
19 (-'ve)	19. I am just not good at Physics.	Higher self-concept		
20	20. I can use Physics concepts to explain real-life situations.	Higher self-efficacy		

(yellow)	mindset
(pink)	anxiety
(green)	self-efficacy
(light blue)	self-concept

Appendix 14b.Table showing a broad IG and CG comparison of self-belief attributes from the **post**-intervention surveys.

Appendix 15. Sample Frequencies of Interactions (Teacher 1)

Extract from observation findings spreadsheet

lesson #	N_T total interactions	O Other	N_L learning related interactions	I Instruction	Is Instruction statement	Ip Instruction prompt	N_F feedback interactions	Task T	Ts Task statement	Tp Task prompt	P Process	Ps Process statement	Pp Porcess prompt	R Regulation (total)	Rs self-regualtion statement	Rp self-regualtion prompt	Self S (all Ss)	s (total) of feedback interactions	p (total) of feedback interactions	B recipient male	%B of N_L
1	579	140	439	105	46	59	334	216	132	84	95	36	59	17	12	5	6	186	148	110	25.1
2	305	42	263	119	74	45	144	69	50	19	45	37	8	31	8	23	0	95	50	74	28.1
3	393	47	346	106	77	29	240	144	86	58	71	41	30	23	13	10	2	142	98	127	36.7
4	419	47	372	74	52	22	298	140	88	52	126	71	55	24	9	15	8	176	122	172	46.2
5	457	100	357	89	60	29	268	99	72	27	149	54	95	11	5	6	9	140	128	88	24.6
6	421	36	385	135	113	22	250	158	87	71	77	51	26	9	4	5	6	148	102	144	37.4
7	470	142	328	88	50	38	240	148	108	40	76	21	55	7	6	1	9	144	96	56	17.1
8	436	85	351	138	72	66	213	156	101	55	52	17	35	6	1	5	0	119	95	118	33.6
9	494	56	438	105	64	41	333	154	102	52	147	74	73	22	9	13	10	195	138	164	37.4
10	328	25	303	105	61	44	198	119	57	62	76	15	61	5	0	5	0	72	128	101	33.3
11	434	23	411	140	88	52	271	127	98	29	116	36	80	12	8	4	11	153	113	79	19.2
12	398	15	383	88	62	26	295	102	65	37	163	68	95	23	13	10	7	153	142	109	28.5
13	376	61	315	75	52	23	240	107	81	26	99	43	56	25	6	19	9	139	101	113	35.9
14	398	82	316	49	31	18	267	133	88	45	120	40	80	8	3	5	7	138	130	154	48.7
	5908	**901**	**5007**	**1416**	**902**	**514**	**3591**	**1872**			**1412**			**223**			**84**	**2000**	**1591**	**1609**	
		15.3		24.0				31.7			23.9			3.8			1.4			27.2	

G recipient female	%G of N_L	C recipient class	%C of N_L	T recipient group	%T of N_L	I as % of N_T	O as % of N_T	T as % of N_T	P as % of N_T	R as % of N_T	S as % of N_T	s as % of N_T	p as % of N_T	B as % of N_T	G as % of N_T	C as % of N_T	T as % of N_T
158	36.0	171	39.0	0	0.0	18.1	24.2	37.3	16.4	2.9	1.0	32.1	25.6	19.0	27.3	29.5	0.0
65	24.7	120	45.6	4	1.5	39.0	13.8	22.6	14.8	10.2	0.0	31.1	16.4	24.3	21.3	39.3	1.3
170	49.1	44	12.7	5	1.4	27.0	12.0	36.6	18.1	5.9	0.5	36.1	24.9	32.3	43.3	11.2	1.3
106	28.5	69	18.5	25	6.7	17.7	11.2	33.4	30.1	5.7	1.9	42.0	29.1	41.1	25.3	16.5	6.0
107	30.0	162	45.4	0	0.0	19.5	21.9	21.7	32.6	2.4	2.0	30.6	28.0	19.3	23.4	35.4	0.0
179	46.5	41	10.6	21	5.5	32.1	8.6	37.5	18.3	2.1	1.4	35.2	24.2	34.2	42.5	9.7	5.0
50	15.2	150	45.7	72	22.0	18.7	30.2	31.5	16.2	1.5	1.9	30.6	20.4	11.9	10.6	31.9	15.3
99	28.2	76	21.7	58	16.5	31.7	19.5	35.8	11.9	1.4	0.0	27.3	21.8	27.1	22.7	17.4	13.3
154	35.2	108	24.7	12	2.7	21.3	11.3	31.2	29.8	4.5	2.0	39.5	27.9	33.2	31.2	21.9	2.4
67	22.1	133	43.9	2	0.7	32.0	7.6	36.3	23.2	1.5	0.0	22.0	39.0	30.8	20.4	40.5	0.6
84	20.4	243	59.1	5	1.2	32.3	5.3	29.3	26.7	2.8	2.5	35.3	26.0	18.2	19.4	56.0	1.2
109	28.5	165	43.1	0	0.0	22.1	3.8	25.6	41.0	5.8	1.8	38.4	35.7	27.4	27.4	41.5	0.0
103	32.7	92	29.2	7	2.2	19.9	16.2	28.5	26.3	6.6	2.4	37.0	26.9	30.1	27.4	24.5	1.9
109	34.5	50	15.8	3	0.9	12.3	20.6	33.4	30.2	2.0	1.8	34.7	32.7	38.7	27.4	12.6	0.8
1560		**1624**		**214**			**14.7**	**31.5**	**24.0**	**3.9**	**1.4**	**33.7**	**27.0**	**27.7**	**26.4**	**27.7**	**3.5**
26.4		27.5		3.6													

I as % of N_L	% dialogic	% T of N_F	% Ts of N_F	% Tp of N_F	% P of N_F	% Ps of N_F	% Pp of N_F	% R of N_F	% Rs of N_F	% Rp of N_F	% S of N_F	total statements	total prompts
23.9	57.7	64.7	39.5	25.1	28.4	10.8	17.7	5.1	3.6	1.5	1.8	232	207
45.2	47.2	47.9	34.7	13.2	31.3	25.7	5.6	21.5	5.6	16.0	0.0	169	95
30.6	61.1	60.0	35.8	24.2	29.6	17.1	12.5	9.6	5.4	4.2	0.8	219	127
19.9	71.1	47.0	29.5	17.4	42.3	23.8	18.5	8.1	3.0	5.0	2.7	228	144
24.9	58.6	36.9	26.9	10.1	55.6	20.1	35.4	4.1	1.9	2.2	3.4	200	157
35.1	59.4	63.2	34.8	28.4	30.8	20.4	10.4	3.6	1.6	2.0	2.4	261	124
26.8	51.1	61.7	45.0	16.7	31.7	8.8	22.9	2.9	2.5	0.4	3.8	194	134
39.3	48.9	73.2	47.4	25.8	24.4	8.0	16.4	2.8	0.5	2.3	0.0	191	161
24.0	67.4	46.2	30.6	15.6	44.1	22.2	21.9	6.6	2.7	3.9	3.0	259	179
34.7	60.4	60.1	28.8	31.3	38.4	7.6	30.8	2.5	0.0	2.5	0.0	133	172
34.1	62.4	46.9	36.2	10.7	42.8	13.3	29.5	4.4	3.0	1.5	4.1	241	165
23.0	74.1	34.6	22.0	12.5	55.3	23.1	32.2	7.8	4.4	3.4	2.4	215	168
23.8	63.8	44.6	33.8	10.8	41.3	17.9	23.3	10.4	2.5	7.9	3.8	191	124
15.5	67.1	49.8	33.0	16.9	44.9	15.0	30.0	3.0	1.1	1.9	2.6	169	148
28.6	**60.7**	**52.6**	**34.1**	**18.5**	**38.6**	**16.7**	**21.9**	**6.6**	**2.7**	**3.9**	**2.2**	**2902**	**2105**
												49.1	35.6

References

Agasisti, T. et al. (2018), ‘Academic resilience: What schools and countries do to help disadvantaged students succeed in PISA.’ *OECD Education Working Papers, No. 167, OECD Publishing, Paris. http://dx.doi.org/10.1787/e22490ac-en*

Agra, E., Fischer, S. M. and Beilock, S. L. (2017). ‘The Role of Students’ Gender and Anxiety in Physics Performance’ *2017 PERC Proceedings, edited by Ding, Traxler, and Cao; American Association of Physics Teachers.*

Ahmed, W., Minnaert, A., Kuyper, H. and van der Werf, G. (2012). ‘Reciprocal relationships between math self-concept and math anxiety.’ *Learning and Individual Differences, vol 22 (3) 385-389*

Ajjawi, R., and D. Boud. (2017). ‘Researching Feedback Dialogue: An Interactional Analysis Approach.’ *Assessment and Evaluation in Higher Education 42:2: 252–265.*

Ajjawi, R. and Boud, D. (2018). ‘Examining the nature and effects of feedback dialogue.’ *Assessment and Evaluation in Higher Education, 43:7, 1106-1119.*

Alivernini, F. and Lucidi, F. (2011). ‘Relationship between social context, self-efficacy, motivation, academic achievement, and intention to drop out of high school: A longitudinal study.’ *The Journal of Educational Research, 104, 241–52.*

Andersson, C., and Palm, T. (2017). ‘The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme.’ *Learning and Instruction, 49, 92–102.*

Andersson, C. and Palm, T. (2018). ‘Reasons for teachers’ successful development of a formative assessment practice through professional development – a motivation perspective.’ *Assessment in Education: Principles, Policy and Practice, 25:6, 576-597.*

Angrosino, M.V. (2012). ‘*Observation-based research.’* In: *Arthur, J., Waring, M., Coe, R. and Hedges, V. (Eds) Research Methods and Methodologies in Education. London: Sage. pp165-169*

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., and Wong, B. (2010). ''Doing' science versus 'being' a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity.' *Science Education, 94:4, 617-639.*

Aronson, J., Fried, C.B. and Good, C. (2002). 'Reducing the Effects of Stereotype Threat on African American College Students by Shaping Theories of Intelligence.' *Journal of Experimental Social Psychology, 38:2, 113–125.*

Arslan, A. (2012). 'Predictive power of the sources of primary school students' self-efficacy beliefs on their self-efficacy beliefs for learning and performance.' *Kuram Ve Uygulamada Eğitim Bilimleri, 12:3, 1915-1920.*

Aschbacher, P. and Alonzo, A. (2006) 'Examining the Utility of Elementary Science Notebooks for Formative Assessment Purposes'. *Educational Assessment, 11:3-4, 179-203.*

Askew, S. and Lodge, C. (2000). 'Gifts, ping-pong and loops-linking feedback and learning.' In *S. Askew (Ed.). Feedback for Learning (pp.1-17). London: Routledge Falmer.*

Atwater, M. M., Wiggins, J., and Gardner, C. M. (1995). 'A study of urban middle school students with high and low attitudes toward science.' *Journal of Research in Science Teaching, 32:6, 665–677.*

Avci, F., and Kirbaslar, F.G. (2017). 'Determination of factors affecting the science anxiety levels of secondary school students.' *Electronic Journal of Science and Mathematics* Education, 11(1), 401-417.

Bandura, A. (1977). 'Self-efficacy: Toward a Unifying Theory of Behavioral Change.' *Psychological Review*, 84:2, 191-215

Bandura, A. (1982) 'Self-efficacy, mechanism in human agency'. *American Psychologist, 37, 122-147.*

Bandura, A. (1997). Self-Efficacy: The Exercise of Control. *New York: W.H. Freeman and Company.*

Bandura, A. (2001). 'Social Cognitive Theory: An Agentic Perspective.' *Annual Review of Psychology. 52,1-26*

Bandura, A. (2006). ‘Guide for constructing self-efficacy scales.’ *In Pajares, F., and Urdan, T. (eds) Self-efficacy beliefs of adolescents, 5, pp.307–337. Greenwich, CT; Information Age Publishing.*

Bandura, A. (2018). ‘Toward a Psychology of Human Agency: Pathways and Reflections.’ *Perspectives on Psychological Science, 13(2), 130–136.*

Bassey, M. (1999). *Case Study Research in Educational Settings*. Buckingham: Open University Press.

Bassey, M. (2000). ‘Fuzzy Generalisations and Best Estimates of Trustworthiness.’ *Paper presented at the ESRC Teaching and Learning Research Programme First Annual Conference - University of Leicester*.

Bassey, M. (2001). ‘A Solution to the Problem of Generalisation in Educational Research: Fuzzy prediction.’ *Oxford Review of Education, 27:1, 5-22, DOI: 10.1080/03054980123773*

Bates, B. (2016) *Learning theories simplified*. London: SAGE Publications.

Bedford, S. (2017). ‘Growth mindset and motivation: a study into secondary school science learning.’ *Research Papers in Education. 32:4, 424-443, DOI: 10.1080/02671522.2017.1318809*

Bennett, R. E. (2011). ‘Formative assessment: a critical review.’ *Assessment in Education: Principles, Policy and Practice, 18, 5-25. http://dx.doi.org/10.1080/ 0969594X.2010.513678.*

Bergh, van den, L., Ros, A. and Beijaard, D. (2013). ‘Teacher feedback during active learning: Current practices in primary schools.’ *British Journal of Educational Psychology (2013), 83, 341–362.*

Bergh, van den, L., Ros, A. and Beijaard, D. (2014). ‘Improving Teacher Feedback During Active Learning.’ *American Educational Research Journal. 51. 772-809. 10.3102/0002831214531322.*

Bergh, van den, L, Ros, A and Beijaard, D 2015, 'Teacher learning in the context of a continuing professional development programme: a case study', *Teaching and Teacher Education, vol. 47, pp. 142-150*. https://doi.org/10.1016/j.tate.2015.01.002

Berthold, K., Nuckles, M. and Renkl, A. (2007). 'Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts.' *Learning and Instruction,* 17, 564-577.

Black, P.J., and Wiliam, D. (1998a). 'Assessment and classroom learning.' *Assessment in Education: Principles, Policy and Practice,* 5(1), 7-73

Black, P.J., and Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment.* London: King's College London School of Education.

Blackwell, L.S., Trzesniewski, K.H. and Dweck, Carol Sorich, 2007. 'Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention.' *Child* development, 78:1, pp.246–63.

Boaler, J. (2002). 'The Development of Disciplinary Relationships: Knowledge, Practice and Identity in Mathematics Classrooms.' *For the Learning of Mathematics. 22:1, 42-47*

Boekaerts, M., and Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. Applied Psychology: An International Review, 54, 199–231. doi:10.1111/j.1464-0597.2005.00205.x

Bond, L., Smith, T., Baker, W., and Hattie, J. (2000). 'The certification system of the National Board for Professional Teaching Standards: A construct and consequential validity study.' *Washington DC: National Board for Professional Teaching Standards.*

Bong, M. (1996). 'Problems in Academic Motivation Research and Advantages and Disadvantages of Their Solutions.' *Contemporary Educational Psychology, 21;2, 149–165.*

Bong, M. and Clark, R. E. (1999). 'Comparison between self-concept and self-efficacy in academic motivation research.' *Educational Psychologist, 34:3, 139-153*

Bong, M. and Skaalvik, E.M., (2003). 'Academic Self-Concept and Self-Efficacy : How Different Are They Really?' *Educational Psychology Review, 15:1,1–40.*

Borko, H. (2004). 'Professional development and teacher learning: Mapping the terrain.' *Educational Researcher, 33, 3–15.*

Boud, D. and Molloy, E. (2012a). 'What is the Problem with Feedback? In: *Anonymous Feedback in Higher and Professional Education. London, Routledge, pp. 11-20.*

Boud, D. and Molloy, E. (2012b). *Feedback in Higher and Professional Education: Understanding it and Doing it Well*. London; New York, Routledge.

Boud, D. and Molloy, E. (2013). 'Rethinking models of feedback for learning: the challenge of design.' *Assessment & Evaluation in Higher Education, 38:6, 698-712.*

Brenner, P. S., Serpe, R. T., and Stryker, S. (2014). 'The Causal Ordering of Prominence and Salience in Identity Theory: An Empirical Examination.' *Social Psychology Quarterly, 77(3), 231–252. doi:10.1177/0190272513518337*

Brenner, P. S., Serpe, R. T., and Stryker, S. (2018). 'Role-specific Self-efficacy as Precedent and Product of the Identity Model.' *Sociological Perspectives, 61(1), 57-80. https://doi.org/10.1177/0731121417697306*

Britner, S. L. and Pajares, F. (2006). 'Sources of science self-efficacy beliefs of middle school students.' *Journal of Research in Science Teaching, 43, 485–99.*

Broadfoot, P. 1996. *Education, assessment and society.* Buckingham: Open University Press.

Brookhart, S. 2008. *How to give effective feedback to your students*. Alexandria, VA: Association for Supervision and Curriculum Development.

Brookhart, S. M. (2012). 'Teacher feedback in formative classroom assessment.' In *C. Webber and J. Lupart (Eds.), Leading student assessment (pp. 225-239). Dordrecht, Netherlands: Springer.*

Brooks, C., Carroll, A., Gillies, R. M., and Hattie, J. (2019). 'A Matrix of Feedback for Learning.' *Australian Journal of Teacher Education, 44(4). http://dx.doi.org/10.14221/ajte.2018v44n4.2*

Brown, N. (2107). Online article. *'In which science actually self-corrects, for once.'* http://steamtraen.blogspot.com/2017/01/in-which-science-actually-self-corrects.html

Brown, G.T., Peterson, E.R. and Yao, E.S. (2016). 'Student conceptions of feedback: Impact on self-regulation, self-efficacy, and academic achievement.' *British Journal of Educational Psychology.86(4):606-629. doi: 10.1111/bjep.12126.*

Brownlow, S., Jacobi, T., and Rogers, M. (2000). ‘Science anxiety as a function of gender and experience.’ *Sex Roles, 42(1-2), 119–131.*

Bryman, A. (2016). *Social Research Methods.* Oxford University Press.

Burnett, Paul. (2010). ‘Praise and Feedback in the Primary Classroom: Teachers' and Students' Perspectives.’ *Australian Journal of Educational and Developmental Psychology. 10. 145-154.*

Burnette, J., O'Boyle, E., Vanepps, E, Pollack, J. and Finkel, E. (2012). ‘Mind-Sets Matter: A Meta-Analytic Review of Implicit Theories and Self-Regulation.’ *Psychological Bulletin. 139. 10.1037/a0029531.*

Butler, D. L., and Winne, P. H. (1995). ‘Feedback and self-regulated learning: A theoretical synthesis.’ *Review of Educational Research, 65, 245–281. doi:10.3102/00346543065003245*

Campbell-Mapplebeck, A.L. (2019). ‘Divergent and convergent feedback: how science teachers conceptualise and practise oral feedback, and how students perceive it helps their learning.’ *PhD thesis, University of York.*

Carless, D. (2006). ‘Different perceptions in the feedback process.’ *Studies in Higher Education 31, no. 2: 219–33.*

Carless, D. (2009) ‘Trust, distrust and their impact on assessment reform.’ *Assessment and Evaluation in Higher Education, 34 (1), pp.79-89.*

Carless, D., Salter, D., Yang, M. and Lam, J. (2011). ‘Developing sustainable feedback practices.’ *Studies in Higher Education, 36 (4), pp.395-407.*

Carless, D. (2012) ‘Trust and its role in facilitating dialogic feedback.’ In: *Boud, D. and Molloy, E. (eds.) Feedback in Higher and Professional Education. London, Routledge, pp. 100-113.*

Carlone, H. (2004). ‘The cultural production of science in reform-based physics: Girls' access, participation, and resistance.’ *Journal of Research in Science Teaching, 41:4, 392-414.*

Carmichael, C. and Taylor, J.A. (2005). ‘Analysis of student beliefs in a tertiary preparatory mathematics course.’ *International Journal of Mathematical Education in Science and Technology 36:7, 713-719.*

Chan, J., and Lam, S. (2010). ‘Effects of different evaluative feedback on students' self-efficacy in learning.’ *Instructional Science, 38:1, 37-58.*

Chen, S., Yeh. Y., Hwang, F., and Lin, S. S. J. (2013). ‘The relationship between academic self-concept and achievement: A multicohort–multi-occasion study.’ *Learning and Individual Differences, 23, 172–8.*

Cleaves, A. (2005). ‘Formation of science choices in secondary school.’ *International Journal of Science Education, 27:4, 471-486.*

Coe, R. and Aloisi, C. and Higgins, S. and Major, L.E. (2014) 'What makes great teaching? Review of the underpinning research.', *Project Report. Sutton Trust, London.*

Coe, R., Waring, M., Hedges, L.V. and Arthur, J. (2017). *Research Methods and Methodologies in Education.* Sage Publishing.

Cohen, L., Manion, L., and Morrison, K. (2017) *Research Methods in Education* (8th ed.) Routledge.

Cook, T. D. and Shadish, W. R. (1994). ‘Social experiments: Some developments over the past fifteen years.’ *Annual Review of Psychology, vol. 45:1, 545-580.*

Craven, R. G., Marsh, H. W., and Debus, R. L. (1991). ‘Effects of internally focused feed- back and attributional feedback on enhancement of academic self-concept.’ *Journal of Educational Psychology, 83:1, 17–27.*

Creswell, J.W. (2013) *Qualitative Enquiry and Research Design: Choosing Among Five Approaches.* (3rd ed). Sage.

Dalgety, J., and Coll, R. K. (2006). ‘Exploring First-Year Science Students' Chemistry Self-Efficacy.’ *International Journal of Science and Mathematics Education, 4*:1, 97-116.

Daniels, L.M. Stupnisky, R.H. Pekrun, R. Haynes, T.L. Perry, R.P. and Newall, N.E. (2009). ‘A longitudinal analysis of achievement goals: from affective antecedents to emotional effects and achievement outcomes.’ *Journal of Educational Psychology, 101, 948-963, 10.1037/a0016096*

Dann, R. (2015). ‘Developing the foundations for dialogic feedback in order to better understand the ‘learning gap’ from a pupil’s perspective.’ *London Review of Education 13:3, 5-20.*

Dann, R. (2018). *Developing feedback for pupil learning: Teaching, learning and assessment in school.* Abingdon: Routledge.

Dann, R. (2019). ‘Feedback as a relational concept in the classroom.’ *The Curriculum Journal, 30:4, 352-374, DOI: 10.1080/09585176.2019.1636839*

Darling-Hammond, L., Chung Wei, R., Andree, A., Richardson, N. and Orphanos, S. (2009). ‘Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad.’ *National Staff Development Council.*

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B. and Osher, D. (2019). ‘Implications for educational practice of the science of learning and development.’ *Applied Developmental Science, DOI: 10.1080/10888691.2018.1537791*

Darling-Hammond, L., Hyler, M. E., Gardner, M. (2017). *Effective Teacher Professional Development*. Palo Alto, CA: Learning Policy Institute.

DCSF (2009). *Assessment for Learning in science. Unit 2a: Developing oral feedback in science.*

Denzin, N.K. and Lincoln, Y.S., (2011). *Handbook of Qualitative Research.* Thousand Oaks, California: Sage. 4th edition

DeWitt, J., Archer, L. and Moote, J. (2019). ‘15/16-Year-Old Students’ Reasons for Choosing and Not Choosing Physics at a Level’. *International Journal of Science and Mathematics Education. 17, 1071-1087*. https://doi.org/10.1007/s10763-018-9900-4

Drever, R. (2003). *Using Semi-Structured Interviews in Small-Scale Research*. Edinburgh: The Scottish Council for Research in Education.

Dupeyrat, C. and Mariné, C. (2005). ‘Implicit theories of intelligence, goal orientation, cognitive engagement, and achievement: A test of Dweck's model with returning to school adults.’ *Contemporary Educational Psychology 30, 43-59.*

Dweck, C.S. (1975). ‘The Role of Expectations and Attributions in the Alleviation of Learned Helplessness.’ *Journal of Personality and Social Psychology.* 31:4, 674-685

Dweck, C. S. (2000) *Self-theories: Their Role in Motivation, Personality, and Development (Essays in Social Psychology).* Psychology Press.

Dweck, C. S. (2006) *Mindset*. New York: Random House

Dweck, C. S. (2010). Talk. 'Challenging Learning' UK Tour, York.

Dweck, C. S., and Yeager, D. S. (2019). 'Mindsets: A View from Two Eras.' *Perspectives on Psychological Science, 14:3, 481–496.*

Education Endowment Foundation (2017) The Sutton Trust-EEF Teaching and Learning Toolkit. Retrieved from https://educationendowmentfoundation.org.uk/public/files/Toolkit/Technical_Appendix/EEF_Feedback_Technical_Appendix.pdf

Education Endowment Fund. (2018) *Embedding Formative Assessment*. Evaluation report and executive summary.

Elbra-Ramsay, C. M. (2019). 'Navigating the Pedagogical, Relational and Moral Economies of Assessment: An Exploration of the development of Student Teachers' Understanding of Feedback.' *Unpublished PhD thesis.*

Elliot, A. J. and Dweck, C. S. (2005). 'Competence and motivation: Competence as the core of achievement motivation.' In *A. J. Elliot and C. S. Dweck (Eds.), Handbook of competence and motivation (pp. 3–12). New York, NY: Guilford Press.*

Elliott, V., Baird, J., Hopfenbeck, T. N. Ingram, J., Thompson, I., Usher, N., Zantout, M., Richardson, J. and Coleman, R. (2016). 'A marked improvement? A review of the evidence on written marking.' *London: Education Endowment Foundation.*

Eraut, M. (2004). 'Informal learning in the workplace.' *Studies in Continuing Education, 26, 247–273.*

ERIC Digest (2003) *Try Harder? Motivational Effects of Effort Attributional Feedback.* ERIC Development Team.

Eva, K.W., Armson, H., Holmboe, E., Lockyer, J., Loney, E., Mann, K. and Sargeant, J. (2012). 'Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes.' *Advances in Health Sciences Education, 17:1, 15-26.*

Ezzy, D. (2002). *Qualitative analysis: Practice and Innovation.* London Routledge.

Feldman, A., and Capobianco, B. M. (2008). 'Teacher learning of technology enhanced formative assessment.' *Journal of Science Education and Technology, 17, 82-99. http://dx.doi.org/10.1007/s10956-007-9084-0.*

Fencl, H., and Scheel, K. (2005). 'Engaging students: An examination of the effects of teaching strategies on self-efficacy and course climate in a non-majors physics course.' *Journal of College Science Teaching, 35(1), 20–24.*

Ferguson, R. F., Phillips, S. F. Rowley, J. F. S and Friedlander, J. W. (2015). *The Influence of Teaching Beyond Standardized Test Scores: Engagement, Mindsets, and Agency. A Study of 16,000 Sixth through Ninth Grade Classrooms.* Harvard

Ferla, J., Valcke, M. and Cai, Y., (2009). 'Academic self-efficacy and academic self-concept: Reconsidering structural relationships.' *Learning and Individual Differences, 19:4, 499–505.*

Flammer, A. (1995). 'Developmental analysis of control beliefs.' *In A. Bandura (Ed.), Self-efficacy in changing societies (p. 69–113). Cambridge University Press.*

Furtak, E. M. (2012). 'Linking a learning progression for natural selection to teachers' enactment of formative assessment.' *Journal of Research in Science Teaching, 49, 1181e1210. http://dx.doi.org/10.1002/tea.21054.*

Gabelica, C., Van den Bossche, P., Segers, M., and Geijselaers, W. (2012). 'Feedback, a powerful lever in teams: A review.' *Educational Research Review, 7, 123–144.*

Gamlem, S. M., and Smith, K. (2013). 'Student perceptions of classroom feedback. *Assessment in Education: Principles, Policy and Practice.' 20, 150-169.* http://dx.doi.org/10.1080/0969594X.2012.749212.

Gamlem, S. and Munthe, E. (2014). 'Mapping the quality of feedback to support students' learning in lower secondary classrooms.' *Cambridge Journal of Education. 44. 75-92.* 10.1080/0305764X.2013.855171.

Gan, M. (2011). The Effects of Prompts and Explicit Coaching on Peer Feedback Quality. PhD Thesis, The University of Auckland.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York: Basic Books.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. London: Falmer Press.

Gipps, C., McCallum, B. and Hargreaves, E. (2000). *What makes a good primary school teacher? Expert classroom strategies.* London: Routledge Falmer.

Glover, C., and E. Brown. (2006). 'Written feedback to students: Too much, too detailed or too incomprehensible to be effective?' *Bioscience Education E-Journal 7: 7–3*. http:// www.bioscience.heacademy.ac.uk/journal/vol7/beej-7-3.pdf

Good, C., Aronson, J. and Inzlicht, M., (2003). 'Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat.' *Journal of Applied Developmental Psychology, 24:6, 645–662.*

Gottheiner, D. M., and Siegel, M. A. (2012). 'Experienced middle school science teachers' assessment literacy: investigating knowledge of students' conceptions in genetics and ways to shape instruction.' *Journal of Science Teacher Education, 23, 531-557*. http://dx.doi.org/10.1007/s10972-012-9278-z.

Grant, A. M. and Wall, T. D. (2009) 'The Neglected Science and Art of Quasi-Experimentation: Why-to, When-to, and How-to Advice for Organizational Researchers', *Organizational Research Methods, 12(4), pp. 653–686.*

Grant, L., Bultitude, K., and Daly, A. (2010) *Girls into Physics: Action Research: A practical guide to developing and embedding good classroom practice.* Institute of Physics.

Greene, J.C. (2004). 'Memories of a novice, learning from a master.' *American Journal of Evaluation. 29 (3), 222.*

Grix, J. (2002). 'Introducing students to the generic terminology of social research.' *Politics, 22(3). 175-86.*

Guba, E.G. and Lincoln, Y.S. (1995) 'Competing paradigms in qualitative research.' In *Denzin, N.K. and Lincoln, Y.S., (eds). Handbook of Qualitative Research. (pp. 105-117). Thousand Oaks, California: Sage.*

Gungor, A. Eryilmaz, A. and Fakioglu, T. (2007). 'The relationship of freshman's physics achievement and their related affective characteristics.' *Journal of Research in Science Teaching. 44(8) 1036-1056.*

Guskey, T. R. (2002a). ‘Does It Make a Difference? Evaluating Professional Development.’ *Educational Leadership. 59 (6). 45-51.*

Guskey, T. R. (2002b). ‘Professional development and teacher change.’ *Teachers and teaching: Theory and practice, 8, 381–391.*

Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran Smith, M., McDonald, M., and Zeichner, K. (2005). ‘How teachers learn and develop.’ In *L. Darling-Hammond and J. Bransford (Eds.), Preparing teachers for a changing world (pp. 358–389). San Francisco, CA: Wiley.*

Hammersley, M. (2017). ‘Research ethics.’ In *Coe, R., Waring, M. and Arthur, J. Research Methods and Methodologies in Education (2nd ed.) Sage.*

Handley, K., M. Price, and J. Millar. 2008. Engaging students with assessment feedback. FDTL5 Project Final Report, Oxford Brookes University.

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. Review of Educational Research, 52, 123-142. http://dx.doi.org/10.3102/00346543052001123

Hargreaves, E. (2013). ‘Inquiring into children’s experiences of teacher feedback: reconceptualising Assessment for Learning.’ *Oxford Review of Education, 39:2 229-246.*

Hargreaves, E. (2014). ‘The practice of promoting primary pupils’ autonomy: Examples of teacher feedback.’ *Educational Research, 56:3, 295-309.*

Hargreaves, E., McCallum, B. and Gipps, C. (2000). ‘Teacher feedback strategies in primary classrooms: New evidence.’ In *S. Askew (Ed.), Feedback for learning (pp. 21-31). London: Routledge Falmer.*

Hattie, J. (1992). *Self-concept*. Lawrence Erlbaum Associates, Inc.

Hattie, J. (1999) *Influences on Student Learning*. Inaugural Professorial Lecture, University of Auckland. http://www.education.auckland.ac.nz/webdav/site/education/shared/hattie/docs/influences-on-student-learning.pdf

Hattie, J. (2012a). *Visible Learning for Teachers: maximising impact on learning.* London: Routledge

Hattie, J. (2012b). 'Know Thy Impact.' *Educational Leadership 70(1), 18-23*

Hattie, J., and Clarke, S. (2019). *Visible Learning: Feedback*. London: Routledge.

Hattie, J. and Gan, M. (2011) Instruction based on Feedback. *Handbook of Research on Learning and Instruction. Ed. Richard E. Mayer, Patricia A. Alexander. Routledge*

Hattie, J. and Masters, D. (2012). *The assessment of feedback from a student viewpoint.* (personal communication).

Hattie, J. and Timperley, H. (2007). 'The Power of Feedback*.' Review of Educational Research March, Vol. 77(1), 81–112*

Havnes, A., Smith, K., Dysthe, O., and Ludvigsen, K. (2012). 'Formative assessment and feedback: making learning visible.' *Studies in Educational Evaluation, 38, 21-27.*

Hazari, Z., Sonnert, G., Sadler, P., and Shanahan, M.-C. (2010). 'Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study.' *Journal of Research in Science Teaching, 47:8, 978–1003.*

Heitink, M.C., Van der Kleij, F.M., Veldkamp, B.P., Schildkamp, K. and Kippers, W.B. (2016). 'A systematic review of prerequisites for implementing assessment for learning in classroom practice.' *Educational Research Review, 17, 50-62*

Hendriks, M. A., Scheerens, J., and Sleegers, P. (2014). 'Effects of evaluation and assessment on student achievement: a review and meta-analysis.' *In M. Hendriks (Ed.), The influence of school size, leadership, evaluation, and time on student outcomes (pp. 127-174). Enschede: University of Twente.*

Higgins, S., Kokotsaki, D. and Coe, R.J. (2011). *The Sutton Trust Toolkit of Strategies to Improve Learning: Student Premium Toolkit.* Durham University. Retrieved from http://www.cem.org/attachments/1toolkit-summary-final-r-2-.pdf

Hollins, M., Murphy, P., Ponchaud, B., and E Whitelegg (2006). *Girls in the Physics Classroom: A Teachers' Guide for Action* London, Institute of Physics.

Honicke, T. and Broadbent, J. (2016). 'The influence of academic self-efficacy on academic performance: A systematic review.' *Educational Research Review, 17, 63–84.*

Hopkins, D. (1993). *A Teacher's Guide to Classroom Research* (2nd ed). Open University Press.

Hounsell, D., McCune, V., Hounsell, J., and Litjens, J. (2008). 'The quality of guidance and feedback to students.' *Higher Education Research and Development, 27(1), 55-67.*

Howe, C. and Abedin, M. (2013) 'Classroom dialogue: a systematic review across four decades of research.' *Cambridge Journal of Education. 43:3, 325-356*.

Huang, C. (2011). 'Self-concept and academic achievement: A meta- analysis of longitudinal relations.' *Journal of School Psychology, 49, 505–528*

Hughes, G. (2011): 'Towards a personal best: a case for introducing ipsative assessment in higher education.' *Studies in Higher Education, 36:3, 353-367*

Institute of Physics. (n.d). Policy Data Centre https://www.iop.org/policy/statistics/uk-a-levels/gender/page_67095.html#gref

Joët, G., Usher, E. L., and Bressoux, P. (2011). 'Sources of self-efficacy: An investigation of elementary school students.' *Journal of Educational Psychology, 103, 649–63.*

Johnson, A. (2016). Unpublished PhD thesis. '*The implementation and evaluation of formative feedback on students' attitudes and perceptions on their learning: An action research.'*

Johnston, J. (2012). 'Deciding Paradigms and Methodology.' In Oversby, J. (ed) *ASE Guide to Research in Science Education*

Johnston, J. and Toplis, R. (2012). 'Research Methods.' In Oversby, J. (ed) *ASE Guide to Research in Science Education*

Joyce, B., and Showers, B. (2002). *Student achievement through staff development*. New York, NY: Longman.

Klem, A. M., & Connell, J. P. (2004). 'Relationships Matter: Linking Teacher Support to Student Engagement and Achievement.' *Journal of School Health, 74, 262-273.* http://dx.doi.org/10.1111/j.1746-1561.2004.tb08283.x

Kluger, A. N., and DeNisi, A. (1996). 'The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory.' *Psychological Bulletin, 119:2, 254–284.*

Knight, N. (2003). 'An evaluation of the quality of teacher feedback to students: A study of numeracy teaching in the primary education sector.' *Paper presented at the AARE/NZARE conference, Auckland, New Zealand*. Retrieved from http://educationgroup.co.nz/wp-content/uploads/2016/04/An-evaluation-of-the-quality-of-teacher-feedback-to-students.pdf

Korthagen, F. A. J. (2010). 'Situated learning theory and the pedagogy of teacher education: towards an integrated view of teacher behaviour and teacher learning.' *Teaching and Teacher Education, 26, 98-106.*

Kröner, S., and Biermann, A. (2007). 'The relationship between confidence and self-concept: Towards a model of response confidence.' *Intelligence, 35, 580–590*

Kunkel, E. (2012). *'Self-efficacy and Self-regulation.'* Internet PowerPoint.

Kwakman, K. (2003). 'Factors affecting teachers' participation in professional learning activities.' *Teaching and Teacher Education, 19, 149–170.*

Leahy, S., Lyon, C., Thompson, M., and Wiliam, D. (2005). 'Classroom assessment: Minute-by-minute and day-by-day.' *Educational Leadership*, 63(3), 18-24

Lee, H., Feldman, A., and Beatty, I. D. (2012). 'Factors that affect science and mathematics teachers' initial implementation of technology-enhanced formative assessment using a classroom response system.' *Journal of Science Education and Technology, 21, 523-539. http://dx.doi.org/10.1007/s10956-011-9344-x.*

Lee, J. (2009). 'Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries.' *Learning and Individual Differences,* 19:3, 355–365.

Li, S. (2011). '*Learning in a Physics Classroom Community: Physics Learning Identity Construct Development, Measurement and Validation*.' PhD thesis, Oregon State University.

Li, Y and Bates, T. 2017 'Does growth mindset improve children's IQ, educational attainment or response to setbacks? Active-control interventions and data on children's own mindsets' *SocArXiv, 1-27.* https://osf.io/preprints/socarxiv/tsdwy

Licht, B. G., and Dweck, C.S. (1984). 'Determinants of Academic Achievement: The Interaction of Children's Achievement Orientations with Skill Area.' *Developmental Psychology. 20, No 4, 628-636*

Lincoln and Guba (1985). *Naturalistic Enquiry*. Sage

Liu, N. and Carless, D. (2006) 'Peer feedback: the learning element of peer assessment.' *Teaching in Higher Education, 11 (3), pp.279-290.*

Lynch, J. (2018). Growth Mindset: *The Perils of a Good Research Story*. Online article https://medium.com/@quixotic_scholar/growth-mindset-the-perils-of-a-good-research-story-d6ce32a447d2

Madaus, G. (1994). 'A Technological and Historical Consideration of Equity Issues Associated with Proposals to Change the Nation's Testing Policy.' *Harvard Educational Review: 64:1, 76-96.*

Marsh, H.W., Hau, K. and Artelt, C., (2009). *OECD' s Brief Self-Report Measure of Educational Psychology ' s Most Useful Affective Constructs: Cross- Cultural, Psychometric Comparisons Across 25 Countries. pp.37–41.*

Marsh, H. W. and Martin, A. J. (2011). 'Academic self-concept and academic achievement: Relations and causal ordering.' *British Journal of Educational Psychology, 81: 59-77.* doi:10.1348/000709910X503501

Marsh, H. W., Pekrun, R., Parker, P. D. Murayama, K., Guo, J., Dicke, T., Arens, A. K. (2019). 'The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies.' *Journal of Educational Psychology, Vol 111(2), 331-353*

Marsh, H., Seaton, M., Dicke, T., Parker, P., and Horwood, M. (2019). 'The Centrality of Academic Self-Concept to Motivation and Learning.' *In* K. Renninger and S. Hidi (Authors), *The Cambridge Handbook of Motivation and Learning (Cambridge Handbooks in Psychology, pp. 36-62). Cambridge: Cambridge University Press.* doi:10.1017/9781316823279.004

McCallum, B., Hargreaves, E. and Gipps, C. (2000). 'Learning: The pupil's voice.' *Cambridge Journal of Education, 30 (2), 275-289.*

Mears, C. L. (2012). 'In-depth Interviews.' In *Arthur, J, Waring, M, and Coe, R (eds), Research Methods and Methodologies in Education, SAGE Publications, London.*

Mercer, N. (2007). ‘Dialogic teaching in science classrooms: Full Research Report.’ *ESRC End of Award Report, RES-000-23-0939-A. Swindon: ESRC*

Mercer, N. (2010). ‘The analysis of classroom talk: Methods and Methodologies.’ British Journal of Educational Psychology, 80, 1–14

Merolla, D.M., Serpe, R.T., Stryker, S., and Schultz, P.W. (2012). ‘Structural Precursors to Identity Processes: The Role of Proximate Social Structures.’ *Social Psychology Quarterly, 75, 149-172*

Mislevy, R.J. (2012). Online Manuscript. *Four metaphors we need to understand assessment.* https://www.ets.org/Media/Research/pdf/mislevy_four_metaphors_understand_assessment.pdf

Moeller, J., Salmela-Aro, K., Lavonen, J., and Schneider, B. (2014). ‘Does Anxiety in Science Classrooms Impair Math and Science Motivation?’ *Selected papers presented at the 2nd Network Gender and STEM Conference, 3–5 July 2014, in Berlin, Germany. International Journal of Gender, Science and Technology.*

Molloy, E., Borrell-Carrió, F. and Epstein, R. (2012). ‘The impact of emotions in feedback.’ In: *Boud, D. and Molloy, E. eds. Feedback in Higher and Professional Education. Abingdon, Routledge, pp. 60-81.*

Molloy, E. and Boud, D. (2013). ‘Changing conceptions of feedback.’ In: *Boud, D. and Molloy, E. eds. Feedback in Higher and Professional Education. Abingdon, Routledge*, pp.11-23.

Morony, S., Kleitman, S., Lee, Y. P. and Stankov, L. (2013). ‘Predicting achievement: Confidence vs self-efficacy, anxiety, and self-concept in Confucian and European countries.’ *International Journal of Educational Research, 58, 79–96.*

Mortimer, E. F. and Scott, P. H. (2003). *Meaning making in science classrooms*. Milton Keynes: Open University Press.

Murphy, P. and Whitelegg, E. (2006) *Girls in the Physics Classroom: A Review of the Research on the Participation of Girls in Physics* London, Institute of Physics.

Nicol, D. J. and Macfarlane-Dick, D. (2006). ‘Formative assessment and self-regulated learning: a model and seven principles of good feedback practice.’ *Studies in Higher Education, 31, 199—218.*

Nuthall, G. (2007). *The hidden lives of learners.* Wellington: NZCER Press.

OECD (2009), *PISA 2006 Technical Report*, PISA, OECD Publishing.

OECD (2010), *PISA 2006 Results: Technical Report,* OECD Publishing.

OECD. (2016). *PISA 2015 results (Volume II): Policies and practices for successful schools, PISA.* Paris: OECD Publishing. http://dx.doi.org/10.1787/978264267510-en

O'Mara, A. J., Marsh, H. W., Craven, R. G., and Debus, R. (2006). 'Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis.' *Educational Psychologist, 41, 181–206.*

Osborne, J. and Collins, S., (2001). 'Pupils' views of the role and value of the science curriculum: a focus group study.' *International Journal of Science Education, 23:5, 441-467.*

Osborne, J., Simon, S. and Collins, S., (2003). 'Attitudes towards science: a review of the literature and its implications.' *International Journal of Science Education, 25, 1049-1079.*

O'Shea, A., Cleary. J., and Breen, S. (2010). 'Exploring the role of confidence, theory of intelligence and goal orientation in determining a student's persistence on mathematical tasks.' *Joubert, M. and Andrews, P. (Eds.) Proceedings of the British Congress for Mathematics Education April 2010*

Ozan, C., and Kıncal, R. Y. (2018). 'The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills.' *Educational Sciences: Theory and Practice, 18, 85–118.* http://dx.doi.org/10.12738/estp.2018.1.0216

Pajares, F., 1996. 'Self-Efficacy Beliefs in Academic Settings.' *Review of Educational* Research, 66:4, 543–578.

Pajares, F. (1997). 'Current Directions in Self-efficacy Research.' in M. Maehr and P. R. Pintrich (Eds.). *Advances in motivation and achievement. Volume 10, 1-49. Greenwich, CT: JAI Press.*

Panadero, E., Jonsson, A., and Botella, J. (2017). 'Effects of self-assessment on self- regulated learning and self-efficacy: Four meta-analyses.' *Educational Research Review, 22, 74-98.* doi: https://doi.org/10.1016/j.edurev.2017.08.004

Patton, M. Q. (2014). Q*ualitative Research and Evaluation Methods: Integrating Theory and Practice* (4th ed.) Sage Publishing.

Pekrun, R., Goetz, T., Titz, W. and Perry, R.P. (2002). 'Academic emotions in students' self- regulated learning and achievement: A program of qualitative and quantitative research.' *Educational Psychologist, 37 (2), pp.91-105.*

Pekrun, R. Cusack, A. Murayama, K. Elliot, A.J. and Thomas, K. (2014). 'The power of anticipated feedback: Effects on students' achievement goals and achievement emotions.' *Learning and Instruction, 29, 115-124*

Pinxten, M., De Fraine, B., Damme, J., and D'Haenens, E. (2010). 'Causal ordering of academic self-concept and achievement: Effects of type of achievement measure.' *British Journal of Educational Psychology, 80, 689–709.*

Poulos, A., and M.J. Mahony. (2008). 'Effectiveness of feedback: The students' perspective.' *Assessment and Evaluation in Higher Education 33:2: 143–54.*

Price, M., Handley, K., Millar, J. and O'Donovan, B. (2010): 'Feedback: all that effort, but what is the effect?' *Assessment and Evaluation in Higher Education, 35:3, 277-289*

Pryor, J., and B. Crossouard. (2008). 'A socio-cultural theorisation of formative assessment.' *Oxford Review of Education 34:1, 1–20.*

Punch, K. and Oancea, A. (2014) *Introduction to Research methods in Education*. Sage Publications.

Ramaprasad, A. (1983). 'On the Definition of Feedback.' *Behavioural Science. 28. 4 - 13.*

Rattan. A., Good. C. and Dweck. C. S. (2012). '"It's ok – Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students.' *Journal of Experimental Social Psychology, 48:3, 731-737*

Reay, D. and Wiliam, D. (1999). ''I'll be a nothing': structure, agency and the construction of identity through assessment.' *British Educational Research Journal, 25: 343–354.*

Rimm-Kaufmann, S. and Sandilos, L. (n.d.) ‘Improving Students' Relationships with Teachers to Provide Essential Supports for Learning’. *American Psychological Association* website.

Ritchie, L. (2018). University of Chichester Teaching and Learning Conference. *Web article* https://www.lauraritchie.com/2018/04/11/linking-skills-feedback-assessment-to-develop-student-agency-deep-learning/

Robson, C. (2002). *Real World Research (2nd ed).* Blackwell Publishing.

Rogers, J. and Revesz, A. (2019). Experimental and quasi-experimental designs. In Simpson. J. (ed.) *The Routledge Handbook of Research Methods in Applied Linguistics*

Roos, B. (2004). ‘Learning and Assessment in the Knowledge Society.’ Paper presented at the *EDEN 2004 Annual Conference, Hungary*.

Ruiz-Primo, M. A. and Furtak, E. M. (2007). ‘Exploring teachers’ informal formative assessment practices and students’ understanding in the context of scientific enquiry.’ *Journal of Research in Science Teaching 44:1, 57-84.*

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144

Sadler, D. R. (1998). Formative assessment: revisiting the territory. *Assessment in Education, 5:1, 77-84.*

Sadler, R. (2010) Beyond feedback: developing student capability in complex appraisal, *Assessment and Evaluation in Higher Education, 35:5, 535–550.*

Sahin, M., Caliskan, S. and Dilek, U. (2015). ‘Development and Validation of the Physics Anxiety Rating Scale.’ *International Journal of Environmental and Science Education. 10(2), pp183-200*

Sahin, M. (2014). ‘The relationship between pre-service teachers’ physics anxiety and demographic variables.’ *Journal of Baltic Science Education, Vol. 13, No. 2, 200-213.*

Sawtelle, V. (2011). ‘A Gender Study Investigating Physics Self-efficacy’. *Doctoral Thesis.* Florida International University.

Sawtelle, V., Brewe, E. and Kramer, L. H. (2012). 'Exploring the relationship between self-efficacy and retention in introductory physics'. *Journal of Research in Science Teaching, 49: 1096–1121.*

Schneider, M. C., and Randel, B. (2010). 'Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment.' In *H. L. Andrade and G. J. Cizek (Eds.), Handbook of formative assessment (pp. 251–276). Abingdon: Routledge.*

Schreiner, C and Sjøberg, S. (2007). 'Science Education and Young People's Identity Construction – two mutually incompatible projects?' In *Corrigan, D., Dillon, J., and Gunstone, R. (Eds.), The Re-emergence of Values in the Science Curriculum. Rotterdam: Sense Publishers*

Schunk, D. H. (2003). *Try Harder? Motivational Effects of Effort Attributional Feedback.* ERIC Digest.

Schunk, D. H., & Ertmer, P. A. (1999). 'Self-regulatory processes during computer skill acquisition: Goal and self-evaluative influences.' *Journal of Educational Psychology, 91*(2), 251–260. https://doi.org/10.1037/0022-0663.91.2.251

Schunk, D. H., and Pajares, F. (2009). 'Self-efficacy theory'. In *K. R. Wenzel and A. Wigfield (Eds.), Educational psychology handbook series. Handbook of motivation at school (pp. 35-53). New York, NY, US: Routledge/Taylor and Francis Group.*

Schunk, D.H. and Pajares, F. (2010). 'Self-Efficacy Beliefs.' *International Encyclopaedia of Education. 668-672*. 10.1016/B978-0-08-044894-7.00620-5.

Schunk, D. H. and Swartz, C. W. (1993). 'Goals and progress feedback: Effects on self-efficacy and writing achievement.' *Annual Meeting of the American Educational Research Association. Atlanta, GA.*

Schunk, D. H., and Usher, E. L. (2013). 'Barry J. Zimmerman's theory of self-regulated learning.' In H*. Bembenutty, T. Cleary, and A. Kitsantas (Eds) Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman (pp.1–28). Charlotte, NY: Information Age Publishing.*

Schunk, D.H., and Zimmerman, B.J. (2007). 'Influencing Children's Self-Efficacy and Self-Regulation of Reading and Writing Through Modelling.' *Reading and Writing Quarterly, 23:1, 7-25*, DOI: 10.1080/10573560600837578

Schütze, B., Rakoczy, K., Klieme, E., Besser, M. and Leiss, D. (2017). 'Training effects on teachers' feedback practice. The mediating function of feedback knowledge and the moderating role of self-efficacy.' *Pedocs open access. ZDM 49 (2017) 3, S. 475-489*

Scott, P., and Ametller, J. (2007). 'Teaching science in a meaningful way: Striking a balance between 'opening up' and 'closing down' classroom talk.' *School Science Review, 88, 77–83.*

Scott, P.H., Mortimer, E.F., and Aguiar, O.G. (2006). 'The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons.' *Science Education, 90, 605–631.*

Scott., S.V. (2014). 'Practising what we preach: towards a student-centred definition of feedback', Teaching in Higher Education, 19:1, 49-57

Seon Ahn, H., and Bong, M. (2019). 'Self-Efficacy in Learning: Past, Present, and Future'. In *K. Renninger and S. Hidi (Authors), The Cambridge Handbook of Motivation and Learning (Cambridge Handbooks in Psychology, pp. 63-86). Cambridge: Cambridge University Press*.

Shanahan, M.-C. (2009). 'Identity in science learning: exploring the attention given to agency and structure in studies of identity.' *Studies in Science Education, 45(1), 43-64.*

Shavelson, R. J., Hubner, J. J., and Stanton, G. C. (1976). 'Self-concept: Validation of construct interpretations.' *Review of Educational Research, 46, 407–41.*

Shaw, K. (2004). 'The Development of a Physics Self-Efficacy Instrument for Use in the Introductory Classroom.' *AIP Conf. Proc. 720, 137.*

Shepard, L. A. (2001). 'The role of classroom assessment in teaching and learning.' *In V. Richardson (Ed.), Handbook of research on teaching (4th ed., pp. 1066–1101).* Washington, DC: American Educational Research Association.

Shute, V. (2007). *Focus on formative feedback*. Princeton: Educational Testing Service.

Shute, V. (2008). 'Focus on formative feedback.' *Review of Educational Research, 78 (1), 153-189.*

Simpson, M. and Tuson, J. *Using Observations in Small-Scale Research*. Edinburgh: The Scottish Council for Research in Education.

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L. and Macnamara, B. N. (2018). 'To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses.' *Psychological Science, 29:4: 549-571*

Skaalvik, E. M. (1997). 'Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety.' *Journal of Educational Psychology, 89(1), 71-81.*

Skovholt, K. (2018). 'Anatomy of a teacher-student feedback encounter.' *Teaching and Teacher Education 69,142-153*

Spencer, S.J., Steele, C.M. and Quinn, D. (1999). 'Stereotype Threat and Women's Math Performance.' *Journal of Experimental Social Psychology, 35:1, 4–28.*

Stankov, L. (1999). 'Mining on the "No Man's Land" between intelligence and personality.' In *P. L. Ackerman, P. C. Kyllonen, and R. D. Roberts (Eds.), Learning and individual differences: Process, trait, and content determinants (pp. 315–338). Washington, DC: American Psychological Association.*

Stankov, L., Lee, J., Luo, W., and Hogan, D. J. (2012). 'Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety?' *Learning and Individual Differences*, 22(6), pp.747–758.

Stevenson, H.W., Lee, S., Chen, C., and Stigler, J., Hsu, C.C., and Kitamura, S. (1990). 'Contexts of achievement: A study of American, Chinese, and Japanese children.' *Monographs of the Society for Research in Child Development*, 58, 1272-1285.

Stryker S., Serpe R.T. (1982) 'Commitment, Identity Salience, and Role Behaviour: Theory and Research Example.' In: *Ickes W., Knowles E.S. (eds) Personality, Roles, and Social Behavior. Springer Series in Social Psychology. Springer, New York, NY*

Svanes, K. I. and Skagen, K. (2016). 'Connecting feedback, classroom research and Didaktik perspective.' *Journal of Curriculum Studies, 49:3, 334-351.*

Tesch, R. (1990). *Qualitative research: Analysis types and Software Tools*. London: Falmer Press.

Thyer, B. A. (2012) *Quasi-Experimental Research Designs*. Oxford Scholarship online. DOI: 10.1093/acprof:oso/9780195387384.001.0001

Tice, D. M. and Wallace, H. (2003). 'The reflected self: Creating yourself as (you think) others see you.' In *M. R. Leary and J. P. Tangney (Eds.), Handbook of self and Identity: 91-105. New York: Guildford Press.*

Torrance, H. (2012). 'Formative assessment at the crossroads: conformative, deformative and transformative assessment.' *Oxford Review of Education, 38:3, 323-342.*

Torrance, H. and Pryor, J. (1998). *Investigating Formative Assessment. Teaching, Learning and Assessment in the classroom*. Buckingham, Open University Press

Trochim, William. (2007). *The Research Methods Knowledge Base.*

Tunstall, P. and Gipps, C. (1996). 'Teacher feedback to young children in formative assessment: A typology.' *British Educational Research Journal, 22 (4), 389-404.*

Tunstall, P, Gipps, C. and Harlen, W. (1996). 'How does your teacher help you to make your work better?' Children's understanding of formative assessment.' *The Curriculum Journal, 7 (2), p. 185-203.*

Ucak, E, and Say, S. (2018). 'Analysing the secondary school students' anxiety towards science course in terms of a number of variables.' *European Journal of Educational Research, 8:1) 63-71. doi: 10.12973/eu-jer.8.1.63*

Udo, M. K., Ramsey, G. P. and Mallow, J. V. (2004). 'Science Anxiety and Gender in Students Taking General Education Science Courses.' *Journal of Science Education and Technology, Vol. 13, 4 pp. 435- 446*

Usher, E. L. and Pajares, F. (2008). 'Sources of self-efficacy in school: Critical review of the literature and future directions.' *Review of Educational Research, 78, 751–96.*

Van der Schaaf, M., Baartman, L., Frans Prins, F., Oosterbaan, A. and Schaap, H. (2013). Feedback Dialogues That Stimulate Students' Reflective Thinking, *Scandinavian Journal of Educational Research, 57:3, 227-245, DOI: 10.1080/00313831.2011.628693*

Vercauteren, D. (2009). 'Do they get the picture? Feedback in primary classrooms.' *Practical Research for Education, 41, 29-36.*

Vermunt, J. D. and Verloop, N. (1999). 'Congruence and friction between learning and teaching.' *Learning and Instruction, 9, 257-280.*

Vignoles, A. (2012). 'Longitudinal research.' In *Coe, R., Waring, M. and Arthur, J. Research Methods and Methodologies in Education (2nd ed.) Sage.*

Voerman, L., Meijer, P.C., Korthagen, F. A. J. and Simons, R. J. (2012). 'Types and frequencies of feedback interventions in classroom interaction in secondary education.' *Teaching and Teacher Education, 28, 1107-1115.*

Voerman, L., Korthagen, F. A. J., Meijer, P.C. and Simons, R. J. (2014). 'Feedback revisited: Adding perspectives based on positive psychology. Implication for theory and classroom practice.' *Teaching and Teacher Education, 43, 91-98.*

Voerman, L., Meijer, P. Korthagen, F and Simons, P.R.J. (2015). 'Promoting effective teacher-feedback: From theory to practice through a multiple component trajectory for professional development.' *Teachers and Teaching. 21,1-20.*

Waring, M. (2012) Finding your theoretical position. In *Coe, R., Waring, M. and Arthur, J. Research Methods and Methodologies in Education (2nd ed.) Sage.*

Walker, C., Gleaves, A. and Grey, J. (2006). 'Can students within higher education learn to be resilient and, educationally speaking, does it matter?' *Educational Studies, 32:2, 251-264.*

Weaver, M.R. (2006). 'Do students value feedback? Students' perceptions of tutors' written responses.' *Assessment and Evaluation in Higher Education 31:3, 379–94.*

Weiner, B. (1985). 'An attributional theory of achievement motivation and emotion.' *Psychological review, 92:4, 548-573.*

Wenger, E. (1998). *Communities of Practice: Learning, Meaning and Identity.* Cambridge University Press.

Wiggins, G. (2014) *What feedback is and isn't*. Author blog. https://grantwiggins.wordpress.com/2014/04/15/what-feedback-is-and-isnt/

Wiliam, D. (2010). ‘An integrative summary of the research literature and implications for a new theory of formative assessment.’ *In H. L. Andrade and G. J. Cizek (Eds.), Handbook of formative assessment (pp. 18–40). Abingdon: Routledge.*

Wiliam, D. (2011) *Embedded Formative Assessment.* Solution Tree Press

Wiliam, D. (2014). ‘Formative assessment and contingency in the regulation of learning processes.’ *Paper presented at the meeting of the American Educational Research Association, Philadelphia, PA.*

Wiliam, D. (2018). ‘Feedback: At the heart of—but definitely not all of—formative assessment.’ *In A. A. Lipnevich and J. K. Smith (Eds.), Cambridge Handbook of Instructional Feedback. Cambridge, UK: Cambridge University Press.*

Wiliam, D. (2019). ‘Feedback: At the heart of—but definitely not all of—formative assessment.’ *PowerPoint presentation AERA, Toronto Canada*

William, D., Lee, C., Harrison, C. and Black, P. (2004). ‘Teachers developing assessment for learning: impact on student achievement.’ *Assessment in Education: Principles Policy and Practice, 11(1) pp. 49–65.*

Williams, M. M. and George-Jackson, C. E. (2014). ‘Using and doing science: gender, self-efficacy, and science identity of undergraduate students in STEM.’ *Journal of Women and Minorities in Science and Engineering. 20. 99-126.*

Winter, G. (2000). A Comparative Discussion of the Notion of 'Validity' in Qualitative and Quantitative Research. The Qualitative Report, Volume 4, Numbers 3 and 4, March 2000. (http://www.nova.edu/ssss/QR/QR4-3/winter.html)

Wolters, C. A., Yu, S. L., and Pintrich, P. R. (1996). ‘The relation between goal orientation and students’ motivational beliefs and self-regulated learning.’ *Learning and Individual Differences, 8, 211–38.*

Yeager, D.S. and Dweck, C.S. (2012). ‘Mindsets That Promote Resilience: When Students Believe That Personal Characteristics Can Be Developed.’ *Educational Psychologist, 47(4), 302-314*

Yang, M. and Carless, D. (2013). ‘The feedback triangle and the enhancement of dialogic feedback processes.’ *Teaching in Higher Education, 18, (3), 285-297.*

Yeager, D. and Walton, G. (2011). ‘Social-psychological interventions in education: They're not magic.’ *Review of Educational Research. 81. 267-301.* 10.3102/0034654311405999.

Yin, R.K. (2015). *Case Study Research Design and Methods*. (4th ed). Sage

Yin, Y., Tomita, M. K., and Shavelson, R. J. (2014). ‘Using formal embedded formative assessments aligned with a short-term learning progression to promote conceptual change and achievement in science.’ *International Journal of Science Education, 36, 531-552*. http://dx.doi.org/10.1080/09500693.2013. 787556.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B. and Shapley, K. (2007). ‘Reviewing the evidence on how teacher professional development affects student achievement.’ *Washington, DC: U.S. Department of Education*

Zaccarelli, F. G., Schindler, A., Borko, H. and Osborne, J. (2018). ‘Learning from professional development: a case study of the challenges of enacting productive science discourse in the classroom.’ *Professional Development in Education, 44:5, 721-737, DOI: 10.1080/19415257.2017.1423368*

Zeldin, A. L. and Pajares, F. (2000). ‘Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers.’ *American Educational Research Journal, 37, 215–46*

Zimmerman, B. J. (1998). ‘Academic studying and the development of personal skill: A self-regulatory perspective.’ *Educational psychologist, 33(2-3), 73-86.*

Where editions have been updated, I have checked page numbers and content, and updated the references to the latest editions.