2 LCT in qualitative research
Creating a translation device for studying constructivist pedagogy

*Karl Maton and Rainbow Tsai-Hung Chen*

*Transcending the divide between theory and data.*

**Introduction**

Qualitative researchers often experience two moments of crisis: when they move from discussing theory to collecting data, and when they move from collecting data to analysis. Too frequently they find their chosen theory lends itself neither to enactment in substantive research nor to engagement with empirical findings. They sense a gap between their theory and data but lack the means of translating between them. Thus, theoretical ideas and arguments that make sense in the abstract often unravel when faced with the real world of research. Moreover, this gap is repeated in the places to which they might turn for guidance. On the one hand, philosophies of social science are typically silent on practical questions. They address ontology and epistemology and may offer meta-theoretical tenets for research but rarely demonstrate their implications for substantive studies. On the other hand, methods textbooks typically offer concrete strategies divorced from the explanatory frameworks that would give them meaning. They explain research techniques, but not why they should be employed. The problem of enabling productive dialogue between theory and data thus remains keenly felt by many scholars of social science.

As Basil Bernstein (2000: 132) highlighted, a key source of this problem resides in the form taken by theories themselves. Bernstein distinguished between the ‘internal language of description’ of a theory or how the constitutive concepts are interrelated, and its ‘external language of description’ or how those concepts relate to referents beyond the theory. Each language can be stronger or weaker. An internal language is stronger where concepts are tightly interrelated and weaker where they are less integrated. An external language is stronger where concepts and referents are related in relatively unambiguous ways and weaker where these relations are vague or unclear. In these terms, qualitative researchers often face frameworks with stronger internal languages but weaker external languages: they make sense in their own terms, but their enactment in empirical research is problematic. Crucially, they struggle to engage with data.
This chapter addresses how Legitimation Code Theory (LCT) can be used to overcome this divide in qualitative research. Specifically, we discuss how to develop an ‘external language of description’ or translation device between theory and data. We ground our discussion in the example of a major study that enacted the LCT concepts of specialization codes (Chapter 1, this volume) to explore how constructivist pedagogy shapes the educational experiences of students (Chen 2010). First, we elaborate on Bernstein’s notion of an ‘external language’ – its rationale, its role in research, and ways it has been interpreted – to clarify the nature of a ‘translation device’. Second, we introduce the study we use to exemplify how such a device can be evolved. Third, we analyse the evolving process of that study. There are few published examples of ‘external languages’; there is even less public discussion of how they can be developed. Publications typically reveal the products of research; here we reveal the process as well as the product, to make explicit part of the craft of LCT (Chapter 1, this volume). We analyse the study as an unfolding narrative, focusing on how relations between theory and data were negotiated in the development of an external language of description. Last, we introduce the resulting translation device, discuss how it enables dialogue between theory and data, and consider the nature of the process more generally.

We should emphasize that this chapter is intended to be neither a definitive guide nor a template for enacting LCT. More widely, it aims neither to normatively define how theory and data should be related nor to restrict diversity in how this can be achieved. As we discuss, there are several interpretations of ‘external languages’, and, as other chapters in this volume illustrate, there are many ways of using LCT and developing translation devices. Rather, by focusing in detail on one study we hope to shed some illustrative light on how the framework can be used in qualitative research to generate explanatory power through fostering dialogue between theory and data.

A discursive gap

As Bernstein insisted, all research involves a theory of some kind, the question is how explicit that theory is made:

we all have models, some are more explicit than others; we all use principles of descriptions, again some are more explicit than others; we all set up criteria to enable us both to produce for ourselves, and to read the descriptions of others, again these criteria may vary in their explicitness. Some of our principles may be quantitative whilst others are qualitative. But the problem is fundamentally the same. In the end whose voice is speaking? My preference is to be as explicit as possible. Then at least my voice may be deconstructed.

(Bernstein 2000: 209)
By ‘voice’, Bernstein meant not a reductive reading of identity, whereby knowledge claims are reduced to ‘voicing’ social categories such as class, gender or ethnicity. Rather, he highlighted the often neglected voice of researchers qua researchers – the basis of claims as a sociologist, as an educationalist, as a linguist, etc. The point is that without making explicit one’s theory and the principles of its enactment, and in ways that enable others to recreate the analysis for themselves, the veracity of one’s knowledge claims remains obscured.

A key aspect of this issue concerns relations between the theoretical and the empirical. All research involves what Bernstein (2000: 209) called a ‘discursive gap’ between theory and data, but frameworks differ in whether and how this gap is traversed. Most fail to either recognize or overcome this gap – they may ‘have a powerful and persuasive internal conceptual language but reduced powers to provide externally unambiguous descriptions of the phenomena of their concern’ (Bernstein 2000: 208). This constrains their capacity for building epistemologically powerful knowledge. Substantive studies using such theories are less able to relate their findings to one another as their processes of analysis are less visible and thus less open to scrutiny by others. They thus lack what Bernstein (2000: 168) called ‘a crucial resource for either development or rejection’ of concepts or the explanations generated by their enactment, leaving the theory at risk of becoming frozen in time.

Reinforcing this problem is a tendency for research into education and society to portray explicit means of enacting concepts as imposing theory onto data in a ‘cookie-cutter’ model that ignores the particularities of objects of study. While such an approach would indeed be deaf to data, so is denial of the discursive gap. Failure to recognize that relations between theory and data are not immediate or unproblematic but rather require an explicit means of translation typically leads to theory becoming deaf to data, for nothing seems to fall outside the theory. Conversely, belief in the possibility of purely inductive explanations, free of pre-existing theory, is a fantasy that renders invisible and thus unquestionable the implicit theories held by researchers. Bernstein’s notion of ‘external languages of description’ suggests an alternative to the false dichotomy of either imposing theory on data or miraculously deriving theories from data. It acknowledges the discursive gap but offers a means for traversing that gap through dialogue by positing the possibility of a translation device.

Interpretations of ‘external language’

What is an ‘external language’ and how can one be developed? Bernstein offered somewhat abstract criteria for external languages but few published examples (2000: 131–41), a paucity that has encouraged divergent interpretations. Though potentially confusing in all using the term ‘external language’, these interpretations have engendered complementary tools for
research. We shall distinguish three such tools as what we shall term *data instruments, mediating languages, and translation devices*. We should emphasize our aim here is to clarify our focus – a means of enabling dialogue between theory and data – rather than to normatively define the term ‘external languages’ or definitively interpret Bernstein’s meaning. Moreover, while ‘translation devices’ will form our focus in this chapter, the other two represent valuable tools for research and all three may be fruitfully used together.

First, *data instruments* provide a methodological guide to a project by delineating how concepts suggest foci for data collection and questions for analysis. They make explicit the movement *from* theory *towards* data. For example, in a study of Australian teachers’ pedagogic strategies with Taiwanese students, Dooley (2001) includes data instruments for classroom discourse and interview data. Each comprises plans for enacting concepts from Bernstein’s framework by identifying key issues they highlight for analysis (e.g. ‘instructional discourse’ is enacted within classroom talk as such categories as ‘dialogic structures’ and ‘monologue’) and questions to ask data (e.g. ‘Is speech key to the activity or is spoken discourse minimized?’). Similarly, in research into the professional knowledge base of Anglophone teachers working overseas in Indonesia, Exley (2005) includes a ‘data generation instrument’ of sample interview questions (e.g. ‘How did you organize your content?’) and a ‘data analysis instrument’ of questions for organizing and interrogating data (e.g. ‘What did teachers say about the specialization of this content?’). Such data instruments offer tools for methodological engagement by illustrating the kinds of questions posed by concepts when exploring specific objects of study. They differ from what we are describing as ‘external languages of description’ in that they concern the process rather than the product of research. In effect, they formalize and condense the kind of narrative we shall unfold below but not the translation device it leads to. Crucially here, they do not systematically show how concepts are instanciated in data and how data can be read across to their conceptualization. Nonetheless, such data instruments valuably make explicit the methodological rationale and workings of research practice.

Second, *mediating languages* take the form of typologies or networks that distinguish sub-categories for the realizations of a concept to create a more empirically sensitive instrument and avoid ‘the dominance of high-level structuring concepts’ (Brown 2006: 140). The main concept is divided into or reconceptualized as categories which, through engagement with data, are recursively divided into sub-categories until the network is able to account for all data in the study. For example, in research into parental participation in mathematics education, Brown (1999) delineates ‘a system of conceptually consistent categories organized in the form of a network’ (2006: 140). In a simpler example, Straehler-Pohl and Gellert (2013) suggest three foci for applying ‘classification’. Such tools are not part of the ‘internal language’ (L1) of a theory because they represent means for
enacting concepts to describe something beyond the theory. However, they typically have a far broader focus than what we are calling ‘external languages’ (L2). For example, multi-level typologies developed for calibrating strengths of ‘semantic gravity’ and ‘semantic density’ conceptualize different types of wording, clasuring and sequencing in English discourse as a whole rather than, say, a specific corpus of observation data.\(^1\) Thus they represent what can be termed mediating languages (or ‘L1.5’): while offering a basis for developing translation devices, their broad focus and wide range of possible data means they may require extension by external languages to engage with a determinate problem-situation. In short, mediating languages comprise a network of potential forms that researchers may encounter and external languages comprise a translation device for recreating the analysis of a specific study. (Put another way, a mediating language can be described as a more general, and an external language as a more specific form of translation device.)\(^2\) However, when set within a broader context of exemplars such typologies and networks may form part of external languages (while also offering more transposable types). Brown (1999), for example, provides detailed descriptions of distinctions and empirical examples that make explicit how to translate between his network and data.

A third interpretation is our focus in this chapter. The notion of external languages as translation devices was first exemplified by studies enacting Bernstein’s (1977) concepts of ‘classification’ (strength of boundaries between contexts and categories) and ‘framing’ (strength of control within contexts or categories). By conceptualizing organizing principles of practices, these code concepts operate at a relatively high level of abstraction and condensation. They require external languages to describe how they are realized within any specific study, such as what boundaries ‘classification’ refers to and what ‘strong’ or ‘weak’ classification looks like in the data. Such devices typically comprise a concept, indicators of how the concept is realized empirically, criteria against which relations with data can be decided, and examples from the data. An example of sustained development of such languages is offered by Morais, Neves and colleagues (e.g. Morais et al. 2004; Neves et al. 2004), who focus on classroom interactions. Table 2.1 presents an extract from an external language for analysing ‘the discursive rule sequence’ in terms of ‘framing’. It includes an indicator (‘exploring/discussing themes under study’), a four-point scale of strengths of ‘framing’, descriptions of criteria for these strengths, and two examples of classroom interactions for a ‘F++’ and a ‘F−−’ value. The external language of description we shall discuss in this chapter includes similar components. (We explain how they translate between theory and data, further below.)

Creating a translation device

If this begins to exemplify what an ‘external language’ looks like as a translation device, the next question concerns how to develop one. Existing
First, an external language is not simply an extension of the internal language of a theory but rather arises from its engagement with the specificities of an object of study. The intention is to enable new or unexpected information to emerge from the data that may reshape both the way concepts are enacted and, potentially, the concepts themselves. Evolving an external language thus requires a measure of distance from the internal language and immersion in the data of the study.

Second, this immersion is necessary because concepts are often realized differently in different problem-situations. For example, the LCT concepts of epistemic relations and social relations are being enacted across a wide variety of social fields and practices. In studies of the low uptake of school qualifications in music, ‘social relations’ are realized in curriculum documents as musical aptitude, attitude and personal expression (Lamont and Maton 2008, 2010); in research into internationalized online education, they are realized in online student postings as personal experience of national cultures (Doherty 2010); and in studies of standpoint theories, they are realized in knowledge claims as membership of social categories.

Table 2.1 An external language of description for the discursive rule sequence (Morais et al. 2004: 79)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>F++</th>
<th>F+</th>
<th>F-</th>
<th>F--</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploring/discussing themes under study</td>
<td>The teacher explores contents according to a rigid order which is never altered even when children intervene</td>
<td>The teacher explores contents according to a given order but accepts children’s interventions at the level of the microsequence</td>
<td>The teacher explores contents altering the microsequence, and occasionally the macrosequence, as a result of children’s interventions</td>
<td>The teacher explores contents, even changing the macrosequence, as a result of children’s interventions</td>
</tr>
</tbody>
</table>

Examples of transcripts

F++ Ronaldo reads aloud the material needed to the realization of an experiment planned by his group.

David, who is part of another group, wants to ask a question.

‘No, sorry, we are leaving doubts to the end.’ (Teacher).

F-- Children made a variety of experiments about several state changes of various substances.

The description of the experiences and the presentation of the results are done according to an order chosen by children.

Teacher’s questions intends to clarify some aspects referred to by children, but do not suggest any sequence to work presentation.
While the concept always refers to relations between knowledge practices and actors, these relations take different empirical forms within each problem-situation. Thus immersion in the data is essential—one cannot impose pre-existing categories from a theory’s internal language (though one can begin from or adapt an existing external language).

Third, immersion in data is not enough: one also needs to move beyond the specificities of the object and back towards theory. For example, the studies just mentioned do not remain locked within their specific foci but relate data to concepts of greater context-independence and condensation, namely social relations and thence specialization codes. This enables their disparate explanations of diverse problem-situations to be brought into relation through the shared internal language of the theory. One can, for example, compare the roles stronger social relations play in subject choices, online learning, and research.

These general characteristics suggest that evolving an external language involves iterative movements between theory and data. As Moss (2001: 18) describes:

> The researcher must be ‘prepared to live with the muddle which is the unordered data, and enjoy the pleasure of its potential, in order to be able to generate the theoretical apparatus which is specific to it’ (Bernstein, personal communication). Get in there too soon with the theory and it will overwhelm the data, limiting its potential to say something new. Delay pulling back from the data too long, and the researcher runs the risk of ending up submerged in the specifics, with no way of identifying the general principles which underpin the whole.

The timing of these movements is thus a matter of judgement. If the product seems more ‘scientific’, the process is more akin to ‘art’ or ‘craft’ (using these terms loosely). This is to say that, once developed, an external language is an objectified vehicle for inter-subjective meaning-making: anyone can use the device to read the data and reconstruct the analysis for themselves. In contrast, as the quote implies, the process of creating such languages is fluid, subjective and requires guidance by experience. In short, it is a matter of gaze. Nonetheless, as Chapter 1 (this volume) emphasizes, such craft work can be made explicit and visible. Thus, analysing the processes involved in creating a translation device forms our focus for the remainder of the chapter.

**Case study and frameworks**

Our case study involves a doctoral researcher, Rainbow Tsai-Hung Chen, and her supervisors, Sue Bennett (University of Wollongong) and Karl Maton (University of Sydney). We shall describe this study as a narrative involving ourselves as protagonists (using surnames), based on contemporaneous
correspondence, documents and notes, supplemented by cross-checked recolllections. The research explored the effects of online constructivist pedagogy on Chinese students at an Australian university. The context to this focus was threefold. First, the dramatic expansion in international students in recent decades had outpaced research into the effects of different forms of pedagogy on their educational experiences. The study focused on Chinese learners as they represented a high proportion of Australia’s international students (Australia Education International 2012). Second, existing research into Chinese students overseas had overwhelmingly focused on the students alone, attributing both their comparative success or failure to proclaimed culturally based attributes. This reductive and non-relational approach thereby exhibited ‘knowledge-blindness’ (Maton 2014b) by largely ignoring the role played by the educational contexts in which students were situated. Third, in recent years literature proclaiming the suitability of constructivism for online education had grown exponentially. It was also widely proclaimed that this approach empowers learners. However, relatively little research had been conducted into learners’ experiences.

Chen’s study focused on: the experiences and expectations brought to Australia by Chinese learners from their previous education; the curricular and pedagogic practices they encountered at the Australian university attended; and how the learners negotiated these educational practices. The data collection comprised: three focus groups with Chinese international students at the university; eight interviews with teachers of online courses in the university’s Faculty of Education and analysis of their course outlines; and multiple in-depth interviews with seven Chinese students in the Faculty, conducted at regular intervals through the course of their studies.

The research drew on the LCT dimension of Specialization, specifically the concepts comprising specialization codes (Maton 2014b; Chapter 1, this volume). At the time Chen commenced the research (2006–7), these concepts were being enacted in a growing number of projects but few extant external languages were available for adoption or adaptation. Thus, a translation device had to be developed from scratch. The research itself is reported more fully in Chen (2010) and summarized in Chen et al. (2011). Here we discuss in four principal stages how the external language was evolved, highlighting in each stage wider lessons for enacting LCT in qualitative research. First, we engage with the process of choosing theoretical frameworks for the study. The research did not begin from LCT but rather came to LCT to address questions left unanswered by frameworks more closely connected to the substantive topic. Second, we discuss how Chen’s scholarly gaze was shaped through an extended immersion in theory, and how she then began enacting and refining this gaze in data collection through immersion in her object of study. Third, we describe moves back from data towards theory during an extended phase of analysis, including the many iterative movements between the two that evolved the external language. In each of these stages we highlight common temptations
awaiting researchers and how these potential obstacles to enabling dialogue between theory and data can be avoided. Last, we turn from processes to products of research by discussing how the external language works as a translation device and its role in the substantive theory advanced by the completed thesis. Our narrative is, of course, simplified. No account can do justice to the numerous movements between theory and data involved in this study. Instead, our aim is to provide, in broad brushstrokes, a basic sense of the dialogic nature of the process whereby LCT was enacted in qualitative research.

Choosing frameworks

When beginning her doctoral study, Chen knew from previous research training that she needed a theory to help construct her research problem. However, existing studies into her chosen topic offered little to build on. A voluminous literature on Asian learners’ experiences of ‘Western’ education has emerged over recent decades but lacks adequate theoretical foundations. Typically, disparate studies of different educational contexts and groups of learners have generated empirical descriptions of various cultural attributes said to characterize Chinese students. Given Chen’s proposed research concerned people moving to another culture, she cast a wider net to review work on intercultural exchange more generally. This literature suggested Berry’s acculturation framework (2005) as a starting point. The model identifies factors affecting cross-cultural adaptation, including ‘the heritage culture’, ‘the host culture’, and ‘contact’ between them that leads to possible outcomes. Chen adapted the model for educational settings: ‘heritage culture’ became Chinese contexts and practices that shaped students’ educational dispositions; ‘host culture’ became the educational contexts and practices they encountered in Australia; and the consequences of ‘contact’ became the students’ educational experiences.

Berry’s framework provided a starting point for the study (see Chen et al. 2008). However, it did not address specifically educational practices. For this Chen turned to Bernstein’s notion of education as comprising three ‘message systems’: curriculum, pedagogy, and assessment (1977). This provided a means of dividing each of Berry’s three main components (‘heritage culture’, ‘host culture’, and ‘contact’) into three sub-categories of curriculum, pedagogy, and assessment, generating nine issues to explore. From her supervisors Chen came to understand that this combined model offered an ‘organizing framework’ but not an ‘analytic framework’ (Maton 2014b: 210). That is, it provided a list of what to explore but not the conceptual means for doing so. Chen still required concepts for analysing the organizing principles of dispositions, practices and contexts across the heritage culture, the host culture, and contact outcomes, in order to reveal similarities, variations and differences among them. Descriptive data analyses, such as ‘Chinese education values learning a large amount of content knowledge’
(on curriculum in the heritage culture) and ‘Australian online courses emphasize learners deciding which part of the content is important to them’ (on curriculum in the host culture), do not by themselves allow systematic comparison. They describe the realizations of organizing principles but not the principles themselves. To move beyond such empirical descriptions required an ‘analytic framework’ of concepts that capture those organizing principles.

It is common for researchers to begin, as Chen did, with models that focus on their own substantive concerns. This flows from the typical starting point for research: a review of existing literature on the topic. Berry directly related to intercultural exchange; Bernstein directly described educational issues. However, a shared substantive topic does not guarantee a framework will offer explanatory power for one’s own study. Conversely, that an alternative theory has yet to be used to explore an object of study does not mean it cannot do so or would not offer comparatively greater explanatory power. It is thus invaluable for scholars to explore frameworks from beyond the often narrow confines of their substantive concerns.

This became clear to Chen upon discovering Legitimation Code Theory. Through attending seminars on theory in educational research, she learned about the framework and met Maton, who was later appointed to help supervise her doctoral research. LCT helped reveal the limits of her existing framework, not through explicit critique of that framework but in comparison. As outlined in Chapter 1 (this volume), LCT comprises a multidimensional conceptual toolkit, each dimension offering concepts for analysing a set of organizing principles underlying practices as a species of legitimation codes. At the time of this research (2006–7) Specialization was the most developed and enacted dimension of LCT (e.g. Maton 2000a, 2000b, 2004, 2006, 2007). Concepts comprising Specialization – including specialization codes, epistemic–pedagogic device and knowledge–knower structures – extend and integrate concepts from Bernstein’s code theory (Maton 2014b: 54–7, 196–205). Put simply, they systematically conceptualize not only knowledge but also knowers. It is beyond our scope here to discuss this development, except to note that this enables constructivist practices (the focus of Chen’s study) to be more fully explored by capturing their emphases on learners’ attributes, dispositions and experiences (Chen et al. 2011; Maton 2010).

The concepts of specialization codes offered three additional features crucial for Chen’s project. First, they provide a means of conceptualizing the organizing principles of dispositions brought by learners, their current educational contexts, and their practices in negotiating these contexts. Second, in doing so, they enable these three factors to be systematically related, in contrast to the exclusive focus on learners’ dispositions characterizing existing studies on overseas Chinese students. Third, by transforming empirical descriptions into analysis of their organizing principles, the concepts reach beyond the specificities of any specific project, enabling
the findings of Chen’s study to be valuable for research using the same concepts to explore very different issues and contexts. Thus, while the combined Berry–Bernstein model provided an organizing framework for her research, concepts from LCT offered an analytic framework.

**Evolving an external language of description**

*From theory to data*

After selecting Specialization as the analytic framework for her study, Chen immersed herself in theory for many months, reading about LCT, frameworks that are extended and integrated by LCT, and studies enacting concepts from all these approaches. This reading sharpened Chen’s sense of the focus of her data collection in two principal ways. First, it explicitly guided her methods, such as generating questions for focus groups, interviews and analyses of course materials. In a similar manner to the ‘data instruments’ discussed earlier in this chapter, the concepts foregrounded key issues to address. For example, Specialization highlights, *inter alia*, the role of ideal knowers in shaping the basis of achievement, generating questions such as ‘What kind of student is considered a good student in China?’ and ‘What kind of student tends to do better in your course?’.

Second, the immersion in theory helped shape the way Chen thought about her object of study and research more generally. This kind of influence is not always as obvious to the noviciate. Unlike explicit questions, it is difficult to see or measure, particularly from the viewpoint of the person whose gaze is developing. Yet it is a crucial aspect of apprenticeship into the gaze embodied by an approach. A set of questions alone could limit seeing or hearing what the object of study is showing or saying, leading to a shallow and semi-mechanical application of theory. To be alive to the possibilities of the data additionally requires the flexibility of a refined gaze. In Chen’s case, this involved learning to think in a realist and relational manner, such as in terms of organizing principles as well as empirical descriptions, relative strengths as well as dichotomous types, and topologies as well as typologies (Chapter 1, this volume). For this, LCT was invaluable, as the concepts themselves embody these attributes and thus can help shape, enact and sustain the gaze. Doing so required Chen not only to read but also to write about the theory, alongside numerous discussions with her supervisors, and to write always in relation to her problem-situation rather than as a metadiscourse of theoreticist comparisons among theories. However, reshaping dispositions is not quickly or easily achieved: it emerges from an extended apprenticeship, of which immersion in theory is but a starting point. We return to how this process continued through data analysis, further below.

While immersion in theory is crucial, it is also a phase apprentice scholars can be reluctant to move beyond. Reading about ideas and discussing
hypothetically their potential enactment in one’s study can feel safe compared to the uncertain whirl of data collection. Indeed, some doctoral candidates never reach beyond this stage, either remaining ruminants or leaving their studies. Nonetheless, while one may keep returning to theory throughout the research process, immersion in theory must end if substantive research is to begin. This is something Bennett and Maton made clear to Chen, explicitly directing her to initiate data collection and immerse herself in the object of study itself. Knowing when to make such movements between theory and data is itself a craft skill to be learned through guided experience.

The data collection lasted ten months, during which Chen concentrated on interviewing participants, analysing course materials, and transcribing and translating data (focus groups and interviews were conducted in Mandarin). She continually wished to connect data to theory, often feeling apprehensive at being unable to digest the large amount of data her research was generating. Again, this is a common experience. Having invested time and energy in theoretical immersion and knowing that the data will have to be analysed in the near future can create an impatient desire to make it definitively explicable as soon as possible. Learning to live in the primeval chaos of data and feel at home there, to paraphrase Ludwig Wittgenstein (1980), is not easy. However, heeding her supervisors’ regularly repeated advice, Chen remained patient and immersed herself in collecting and understanding the data on its own terms. To reiterate the quote from Bernstein, to ‘live with the muddle’ of unordered data is also to ‘enjoy the pleasure of its potential’, for it speaks its own language, one that is not that of the internal language of the theory.

From data to theory (and back again)

The temptation to reach for theory too soon is perhaps strongest when moving from data collection to analysis. Having spent considerable time and energy ‘in the field’, the desire to know that you have not wasted your efforts can be overwhelming. The temptation is further fuelled by the capacity of well-established researchers to glance at data and declare X an ‘integrated code’, Y a ‘cultivated knower code’ or Z as revealing a ‘middle-class habitus’. Such realizations of a well-seasoned gaze may appear to newcomers like divine instruction but provide little guidance for employing theory in analysis.

Whether from enthusiasm or anxiety, novice researchers may thus begin imposing concepts on data before it has a chance to speak. This often involves extracting fragments of data (such as brief quotes) for analysis, shorn of the broader context which underpins their meaning. In this study, Chen began declaring specialization codes, peppering descriptions of data with LCT annotations of ‘ER+’, ‘SR−’, etc. When Chen eagerly sought confirmation of her analysis from Maton, much to her initial disappointment he
refused to agree or disagree, arguing that these conceptual proclamations were unsupported by and obscuring the data. As Chen’s supervisors emphasized, the rich qualitative data had stories to tell that were being smothered by concepts. They argued that only by immersing herself in the data and then moving slowly from within that data towards theory, and even then via categories that emerged from the data itself and not from the internal language of the theory, could Chen enable those stories to be heard. The analysis thus became staged into a thematic analysis of data, arrangement of that coded data into a descriptive account using the organizing frameworks, and analysis of this descriptive data using LCT. We discuss each in turn.

**Empirical thematic analysis**

Bernstein suggested that in evolving external languages, the ‘first step’ is ‘to ignore the theory and model… Crucial to the procedure is that it is constructed independently of the L1, that is, independent of the theory and the derived model’ (2000: 137–8). Accordingly, Chen began by simply coding the raw data with labels based on simple empirical descriptions, without using any concepts. The purpose of this thematic analysis was to establish a series of basic categories with which the data could begin to be organized. Initially, the analysis generated more than 300 categories, such as ‘quantity of knowledge’ and ‘teacher control’, which became eventually pared down to 26 as overlaps, similarities or overarching categories emerged. While not theory-determined in the sense of explicitly employing concepts, such analysis is theory-laden, thanks to the researcher’s gaze. For example, when explaining how Chinese students’ fear of ‘losing face’ prevented them from expressing their thoughts in class, a student stated:

> If today you are studying, say, lesson five, the teacher will expect you to know everything in the previous four lessons before you come to class. And they will give you a tongue-lashing if you ask a question about lessons 1–4. So you gradually lose confidence in asking simple questions.  
> (Focus group 1)

Rather than labelling this comment as simply ‘losing face’ or ‘reluctance to speak’, Chen also categorized it as concerning ‘the collective pace of learning’, reflecting an issue – pacing – raised by her readings and discussions about code theory. While not always obvious to the noviciate, one’s specialized gaze is always active in research.

**Organizational coding**

Thematic analysis allowed Chen to concentrate on the potential meanings emerging from the data rather than attempting to fit data into pre-existing
concepts. The resulting categories were then arranged in a second stage which we termed ‘organizational coding’. This more explicitly involved theory, specifically utilizing Berry’s acculturation model and Bernstein’s ‘message systems’ as organizing frameworks. They served to arrange data into three main issues of ‘heritage educational culture’, ‘host educational culture’, and ‘contact’ (students’ experiences), and three sub-themes of curriculum, pedagogy and assessment. Using this structure, Chen wrote three preliminary ‘chapters’ reporting the data in a highly descriptive manner, copiously including quotes from focus groups, teacher interviews, unit of study outlines, and student interviews. Indeed, her supervisors actively encouraged Chen to include far more data than could possibly be included within a completed doctoral thesis. For example, while one quote may eventually become exemplary for illustrating a theme, Chen was encouraged to include as many as possible and to not be concerned about including the same quotes in more than one place.

At the end of the curriculum, pedagogy and assessment sections of each ‘chapter’, Chen identified broad themes emerging from the section, discussing them descriptively and in non-LCT terms. For example, in the ‘chapter’ on ‘host educational culture’, the themes emerging from teachers’ accounts of their courses can be summarized as follows:

- **Curriculum**: emphasis on personal knowledge and experience; downplaying of content knowledge; personal interpretations of content knowledge.
- **Pedagogy**: emphasis on learning over teaching; beliefs in collaborative learning; implementation of individualized learning.
- **Assessment**: implicit evaluative criteria; emphasis on helping students develop individualized knowledge; downplaying of students demonstrating content knowledge.

As this list of knowledge- and knower-related themes suggests, the nascent gaze Chen had begun to acquire through immersion in theory was already engaging with the specificities of the data. While remaining at the level of empirical description, the theoretical analysis that would later become more explicit can already begin to be discerned. Nonetheless, while Chen now had organized and ‘thick descriptions’ of students’ educational dispositions, the courses, and students’ experiences, she was not yet able to compare them in a systematic manner. The organizing principles underlying these dispositions, contexts and practices were not yet explicit. This formed the focus of a final main stage of moving further into theory.

**Analytic coding**

To explore these organizing principles Chen drew on the LCT concepts of *specialization codes*, comprising modalities of strengths of *epistemic relations*
(ER) between knowledge practices and their proclaimed objects of study, and social relations (SR) between knowledge practices and their actors, authors or subjects (Chapter 1, this volume). Practices may more strongly (+) or weakly (−) emphasize each relation and these two strengths give four principal specialization codes (see Figure 1.2, page 12). At their simplest, these declare that legitimacy depends on ‘what you know’ (knowledge codes; ER+, SR−), ‘who you are’ (knower codes; ER−, SR+), both specialist knowledge and knower attributes (élite codes; ER+, SR+), and neither (relativist codes; ER−, SR−). As discussed earlier above, these relations are realized differently in different problem-situations. Thus, Chen interrogated each theme by asking:

1. what form do epistemic relations and social relations take here?;
2. what form do stronger or weaker epistemic relations and stronger or weaker social relations take here?; and
3. does this theme indicate stronger or weaker epistemic relations and/or social relations?

Such questions are typically not answered definitively at first but rather involve repeatedly referring to data across the project, particularly given that strengths of relations are relative. An albeit much simplified example of this complex process can be given using the following excerpt from a teacher interview:

The assignments try to be authentic. We try to situate the assignment in the context in which these people work and live. So if they are a teacher teaching cabinet-making, then they have to think about how their students are learning that task. If they’re a university teacher teaching science, then they have to think about their students learning science.

(Teacher E)

This quote was included in the ‘curriculum’ section of Chen’s empirical description of the ‘host educational culture’, addressing the nature of the Australian online teaching practices explored in the project. It illustrates how teachers emphasized the need for students to make connections between knowledge they may learn at university and their own work contexts. Stronger social relations in knowledge practices are defined as reflecting an emphasis on the dispositions of actors, whether innate, cultivated or socially based. Chen thus tentatively judged the quote as reflecting stronger social relations realized as an emphasis on personal knowledge and experience. The teacher highlights that what is valued is the knowledge learners bring to the educational context by virtue of their existing experiences. Conversely, stronger epistemic relations are often realized as an emphasis on principles or procedures. Here, however, such specialized knowledge or skills are absent. Taken in concert with numerous other teacher statements
and course materials echoing this theme, Chen thus judged the quote to exhibit relatively weak epistemic relations (realized as downplaying specialist content knowledge) and relatively strong social relations (realized as the notion of students as already legitimate knowers by virtue of personal experience).

However, these strengths of epistemic relations and social relations for curriculum were not realized empirically in the same ways when the same participants addressed pedagogy or assessment. In discussions of pedagogy, teachers focused not on ‘personal knowledge and experience’ but rather on individual learners’ preferred ways of learning and downplayed their own teaching practices. For example, one teacher explained that she expected students to ‘negotiate to learn in a way that suits them . . . it’s constructing your own learning in a way that is helpful for you’ (Teacher B). This and many other examples suggested that social relations in discussions of pedagogy were realized as personal dimensions of the learning process and epistemic relations as principles and procedures for teaching content knowledge. In discussion of assessment, social relations were realized in terms of students evaluating their own learning – for example: ‘What’s valid for you and what’s valid for me are two different things, aren’t they?’ (Teacher C) – and epistemic relations were realized as explicit evaluative criteria for judging student performances. In summary, though the same concepts were used to analyse these participants’ beliefs and practices concerning curriculum, pedagogy and assessment, each theme was analysed in its own terms.

Chen conducted similar analyses on the themes from her ‘chapters’ on the ‘heritage educational culture’ and ‘students’ experiences in the host culture’. Each time the forms taken by relations were explored and their relative strengths discussed and compared. Gradually, the different realizations of specialization codes in the three message systems of curriculum, pedagogy and assessment came to light for the three principal themes. Over time this evolved into the external language of description represented in Table 2.2 (further below).

As we emphasized at the outset, this process was not as simple or linear as the above examples might suggest. The evolving external language and emerging analysis were repeatedly refined. Often coding began from a hunch or ‘sense’ based on Chen’s understanding of the concepts and data. One of the benefits of the immersion we have described is a ‘feel’ for the data, an ambient sense of its semantic potential. A search for supporting and disconfirming data would follow, leading to adjustment of initial judgements and further returns to data for support or disconfirmation. In this process one writes the analysis in pencil, as it were, always ready to erase and rewrite, leaving open the possibility that one’s judgement may be wrong. It thus moves through repeated shifts between initial thoughts, the data, and what the concepts themselves suggest. This process also involves recurrent movements between wide-angle and soft-focus analysis of the entire problem-situation in fuzzier analytic terms and
telephoto and hard-focus analysis of more delimited instances with greater precision (see Chapter 5, this volume). By offering a descriptive account of the data as a whole, a wide-angle and soft-focus analysis enables a general sense of the codes involved to emerge and provides a context for more detailed explorations. Through rigorous studies of specific examples, telephoto and hard-focus analysis enables more precise understanding of the diverse realizations across the data of the codes and concretizes the more holistic picture. Chen was encouraged to repeatedly shift between these mutually-informing forms in her analytic coding. Her supervisors asked for any specific instances, such as the examples given above, to be contextualized within a general account of the study or compared to other examples, and for broader claims to be exemplified concretely. In this way, both the wood and the trees remained in view.

These iterative and recurrent movements between theory and data and between general and detailed analyses were thus situated within a social context of discussions with other researchers. Chen repeatedly refined her judgements after feedback from her supervisors, until a kind of equilibrium was reached between the data and conceptual redescriptions. Through these processes of shifting between data and theory, zooming between wide-angle and telephoto visions, and refocusing between soft- and hard-focus analyses, one’s image of the problem-situation becomes sharper and one’s ‘feel’ for the codes becomes codified, culminating in a completed (though always conjectural) analysis and an external language of description.

A translation device

The external language for Chen’s study is recreated in Table 2.2 and comprises: the forms taken by epistemic relations and social relations in discussions of curriculum, pedagogy and assessment; indicators for whether data exhibits stronger or weaker relations; and quotes from data as examples. It includes sections for epistemic relations and for social relations. Each section is structured so that when read from left to right it translates theory into data, and when read from right to left it translates data into theory. The former shows how concepts are enacted in this particular object of study; the latter shows how data can be conceptualized as exemplifying strengths of epistemic relations and social relations. For example, in the curriculum row of ‘epistemic relations’, the quote ‘The information in the textbook – decided by the teacher – was what a study unit was all about’ suggests content knowledge is being highlighted as the determining form of legitimate knowledge, which represents stronger epistemic relations, and so is coded as exhibiting ‘ER+’. Conversely, the quotes exemplify the kinds of data coded as differing strengths of relations, giving insight into how further data should be conceptualized.

Though there is no single form external languages can take, a simple table offers a more portable and synoptic instrument than a prosodically
<table>
<thead>
<tr>
<th>Concept manifested as emphasis on:</th>
<th>Indicators</th>
<th>Example quotes from empirical data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curriculum</td>
<td>content knowledge ER+</td>
<td>Content knowledge is emphasized as determining form of legitimate educational knowledge. The information in the textbook – decided by the teacher – was what a study unit was all about.</td>
</tr>
<tr>
<td>Pedagogy</td>
<td>the teaching of content knowledge ER+</td>
<td>Procedures for learning content knowledge are explicit to learners and emphasized as determining form of pedagogy. [The teacher] extracts the best things from what he or she knows and gives this to you in class, and then offers you instructions on the tasks you need to complete.</td>
</tr>
<tr>
<td>Assessment</td>
<td>explicit criteria ER+</td>
<td>Explicit evaluative criteria are emphasized in judging student performances. When a Chinese child paints the moon blue, the teacher will correct the child, saying that the moon shouldn’t be blue.</td>
</tr>
<tr>
<td></td>
<td>explicit criteria ER−</td>
<td>Explicit evaluative criteria are less significant in judging student performances. It’s not like learning medicine, you’ve got to get it right [otherwise] the patient will die. It’s not like that. It’s more open to interpretation.</td>
</tr>
</tbody>
</table>

### SOCIAL RELATIONS (SR)

<table>
<thead>
<tr>
<th>Concept manifested as emphasis on:</th>
<th>Indicators</th>
<th>Example quotes from empirical data</th>
</tr>
</thead>
<tbody>
<tr>
<td>personal knowledge and experience SR+</td>
<td>Personal experience and opinions are viewed as legitimate educational knowledge. [Students] actually come with a whole range of background and experience . . . what they need is a framework to download that.</td>
<td></td>
</tr>
<tr>
<td>personal dimension of the learning process SR+</td>
<td>Individual learners’ preferences are explicitly emphasized as determining form of pedagogy. So negotiate to learn in a way that suits them . . . it’s constructing your own learning in a way that is helpful for you.</td>
<td></td>
</tr>
<tr>
<td>self-evaluation SR+</td>
<td>Evaluation of legitimacy of student performances resides in beliefs of individual learners. Even if your question is brilliant, the teacher still might not answer you because he or she wants to teach something else first.</td>
<td></td>
</tr>
<tr>
<td>SR−</td>
<td>Student performances are judged against shared criteria external to the learner. I am a ‘test-taker.’ If the teacher doesn’t give me a standard, I don’t know what to do.</td>
<td></td>
</tr>
</tbody>
</table>
LCT in qualitative research

scattered prosaic description of conceptual enactment. Nonetheless the form
taken by such a table or figure is not set in stone. Other external languages
for LCT are simpler than Table 2.2, with columns for: code concepts (such
as ER/SR or ‘semantic gravity’), indicators for a range of relative strengths,
forms these take in the specific data, and examples from data (see Maton
2014b). In Chen’s case, the need to distinguish curriculum, pedagogy and
assessment makes the table slightly more complex, with two sections. There
is also work to be done in developing innovative forms of presentation that
can, for example, relate data to the Cartesian planes used to embody LCT
concepts. However, what is crucial is less the precise form of external lan-
guages than their capacity to act as translation devices between concepts
and data.

Crucially, an external language is neither evolved for its own sake nor the
only product of the process we have described. As a translation device, an
external language is primarily intended to serve the analysis of the problem
with which the research is concerned. Chen’s aim was to explore the effects
of online constructivist pedagogy on Chinese international students. Con-
comitant with evolving the external language, Chen generated a substantive
theory concerning these effects, one which both attends to the specificities
of the object of study and reaches beyond them.⁴

The ‘heritage culture’ of Chinese educational practices was conceptual-
ized as exhibiting a knowledge code (ER+, SR−): in short, content know-
ledge, teaching procedures and explicit evaluative criteria were emphasized
as the bases of legitimacy. In contrast, the ‘host culture’ of constructivist
pedagogies encountered by these students in Australia exhibited a knower
code (ER−, SR+): learners’ personal experiences, learning preferences and
self-evaluation were constructed as significant to achievement. This down-
played everything Chinese education had socialized the students to value,
and valorized what they had previously learned to consider as unimportant.
The ‘code clash’ that resulted from ‘contact’ between these dispositions
and practices shaped the students’ experiences. With their knowledge-code
dispositions (ER+, SR−), the students did not recognize the stronger social
relations of these knower-code practices (ER−, SR+) as legitimate. They
did not, for example, view personal experience as valid educational know-
ledge. Given these practices also exhibited relatively weak epistemic rela-
tions, the students thereby experienced them as a relativist code (ER−,
SR−), a kind of ‘anything goes’. This is to say they could not recognize the
stronger social relations and keenly felt the lack of stronger epistemic rela-
tions, such as the absence of sufficient content, explicit teaching and clear
assessment criteria. The result was a felt lack of legitimacy: students
described being ‘in a vacuum’, ‘no one cares what I’m doing’, ‘lonely’, and
feeling inferior, insecure, anxious, frustrated, helpless, guilty and depressed
(Chen 2010). A typical response was to continue their knowledge-code
practices, such as synthesizing ‘personal experiences’ from the literature
they read.
Space precludes extensive discussion of this substantive theory (Chen 2010; Chen et al. 2011), but the summary above highlights several features relevant to our focus. First, as mentioned, evolving an external language was not the principal aim of the research; rather, the key issue was to explore the effects of constructivist pedagogy on these Chinese students. The external language helped clarify, systematize and codify the analysis that generated the explanation of these effects. Second, as a translation device, the external language makes explicit the basis of this explanation. The device thereby makes research more accountable to other researchers in the field: they can use it to critically inspect and recreate the analysis. The external language acts a kind of a key or decoder to the analysis. Third, the device makes the outcomes of the study more available to other researchers in the field: not only can they build on the findings and the substantive theory, they can also adopt or adapt the external language for their own studies. Though likely to need modification, it provides a valuable starting point that enables work to feed into one another.5

Last, the translation device, combined with the capacity of LCT concepts to explore the organizing principles underlying dispositions, practices and contexts, gives the study relevance beyond the specific topic. One can condense a key conjecture arising from the thesis as: knowledge-code learners are likely to experience knower-code practices, where this code is not made clear, as a relativist code, leading to a felt loss of legitimacy and deleterious educational and psychological outcomes. While the study is rich with empirical detail and deeply immersed in the concrete particularities of its object of study, this conjecture reaches beyond such specificities as ‘Chinese learners’, ‘online education’, ‘constructivist pedagogies’ and ‘Australian universities’ to offer a starting point for comparative study of similar or contrasting cases, such as learners with knower-code dispositions entering educational contexts characterized by knowledge-code practices. The external language thereby helps provide a gateway to a wider range of research, enabling more integrative and cumulative forms of knowledge-building.

Conclusion

Anxieties felt by scholars when enacting a theory in collecting and analysing data are often well founded. Most approaches do not possess a conception of ‘external languages of description’. They often deny or admit defeat to what Bernstein termed the ‘discursive gap’ between theory and data. Put bluntly, most theories fudge the issue, offering little insight into how to negotiate these relations in research. By building on Bernstein’s notion of ‘external languages’, LCT aims to help overcome these issues; by describing how an external language was evolved within a substantive study, this chapter aims to help shed light on the process and the product.

As we have discussed, evolving an external language defies the false dichotomy of theory/data that bedevils research into education and society.
Against this ‘either-or’, LCT posits more than a ‘both-and’: it emphasizes both theory and data and relations between the two. Put another way, enacting LCT involves immersion in and getting a ‘feel’ both for the theory and for the data that then enables them to be brought together through a translation device. It thereby enables both thick description and thick explanation, both empirical fidelity and explanatory power. Rather than ‘either-or’, LCT thereby enables studies to trace a series of semantic waves or recurrent movements between context-dependent and specific meanings and context-independent and complex meanings, between minute particulars and condensed abstractions (Chapter 1, this volume). The resultant translation device then extends the semantic range of the framework to not only reach from descriptions to theorizations but also explicitly reveal how to move between them, in both directions. Studies of research using LCT suggest these characteristics are critical for enabling the building of cumulative and epistemologically powerful knowledge (Maton 2014b: 125–47).

Evolving an external language is, however, not easy. As Brown (2006: 130) highlights, it is ‘a time-consuming process that requires extensive, and thus expensive, engagement with empirical texts’. This is one reason why evolving external languages is often the province of postgraduate research or major projects. Nonetheless, as we discussed, extant external languages can be adopted or adapted and existing models, taxonomies and typologies can often be recruited to serve as the basis for developing translation devices (Maton 2014b: 210–12). Not every study need begin again from scratch.

As well as requiring effort and energy, it is also risky. Enabling others to see the basis of one’s analysis is to open up one’s explanations and conjectures to critical discussion. Where most approaches play it safe by using vague, ill-defined and woolly concepts, this is to put one’s analytic cards on (and literally in) the table. However, the prize is worth the risk. Making explicit the basis of one’s analysis enables more scholars to engage with one’s work and allows the possibility of expanding the sphere of debate and extend the community. Moreover, creating and refining external languages is a key contribution to research. To quote Basil Bernstein:

L₂ is equally an imaginative act as L₁ but is rarely constructed to warrant that adjective. It is essentially what research is about. The rest can be done in an armchair. Armchairs do not change one, only accommodate. Research is the means of change.

(Personal letter to Karl Maton)

Put more specifically:

Though the term ‘external’ may appear to suggest a secondary role, such languages of description represent a crucial catalyst to development. An external language provides a means for translating between theory and
data that other studies can adopt or adapt – to develop an external language is to extend the framework into a new problem-situation.

(Maton 2014b: 206)

The evolution of external languages is thus a crucial means for bringing together studies of a growing range of disparate problem-situations – they enable not only dialogue between theory and data but also dialogue between studies of diverse phenomena by translating among different data through the theory. In short, such translation devices are central to cumulative knowledge-building. No theoretical framework should be without translation devices.

Notes

1 See Maton and Doran (2015a) for a mediating language for analysing English discourse with semantic density; mediating languages for enacting semantic gravity in analyses of English discourse and images will be available in future publications (see LCT website at www.legitimationcodetheory.com).

2 In Chapter 12 of this volume, Maton (page 243) defines several kinds of translation device or ‘means of relating concepts to something beyond a theoretical framework’:

   - external languages of description for translating between theory and empirical data within a specific problem-situation;
   - external languages of enactment for translating between theory and practice;
   - mediating languages for translating between theory and all empirical forms of a phenomenon (i.e. a non-specific external language).

Our concern in this chapter is with enacting concepts in substantive research studies of specific problems-situations; our focus is thus external languages of description.

3 It was the impression upon a younger colleague of magical intuition when analysing data which motivated Maton to develop examples for workshops in LCT that explicitly illustrate a series of stages. Each comprises: a ‘raw’ transcript of classroom practices; the transcript annotated with concepts; an ‘analytical narrative’ that describes the practices, lightly using LCT concepts wherever appropriate; and a ‘conceptual redescription’ that rewrites the narrative as a synoptic theoretical explanation, with minimal reference to empirical content, and generates conjectural explanations. The key point of these examples is to make more visible different stages of the cooking process whereby raw data becomes theoretical explanation.

4 On relations between substantive theories generated by empirical studies and explanatory frameworks such as LCT, see Chapter 1, this volume.

5 Similarly, because LCT extends and integrates Bernstein’s code theory, existing external languages developed for inherited concepts, such as ‘classification’ and ‘framing’, can be extended to develop external languages for specialization codes (which include ‘classification’ and ‘framing’).
References

References


References


References

References


References


References


References


References


South African Department of Basic Education (2010) National Senior Certificate English Home Language Paper 2, Cape Town: Department of Basic Education.


References


Wolfe, J. (2007) ‘“You’ll have to start early if you want to be on time for the F sharp!” Language and the study of music: Implications for international students studying in tertiary music programs in Australia’, paper presented at ISANA International Conference, Adelaide, Australia.

