Chapter 2

Construct Maps

2.0 CHAPTER OVERVIEW AND KEY CONCEPTS

construct
construct maps

This chapter concentrates on the concept of the *construct map* introduced in the previous chapter. The aim is to introduce the reader to this particular approach to conceptualizing a construct—an approach found to be useful as a basis for measuring. There is no claim being made here that this approach will satisfy every possible measurement need (this point is expanded on at the end of the chapter). However, both for didactic purposes and because it will prove a useful tool in many applications, this chapter concentrates on just this one type of construct, as does the rest of the book. It consists of a series of construct maps, illustrating the main different types: respondent maps, item-response maps, and construct maps. All of the examples are derived from published applications. The reader can also find examples of construct maps within each of the cases in the cases archive in the compact disk included with this book. These contain both instances where the measurer has shared both the initial ideas and images of the construct map, as well as construct maps that have been through several iterations.

2.1 THE CONSTRUCT MAP

The type of construct described in this chapter is one that is particularly suitable for a visual representation—it is called a *construct map*. Its most important features are that there is (a) a coherent and substantive definition for the content of the construct; and (b) an idea that the construct is composed of an underlying continuum—this can be manifested in two ways—an ordering of the respondents and or an ordering of item responses. The two different aspects of the construct—the respondents and their responses—lead to two different sorts of construct maps: (a) a respondent construct map, where the respondents are ordered from greater to less, and qualitatively may be grouped into an ordered series of levels; and (b) an item-response construct map, where the item responses are ordered from greater to less, and qualitatively may also be grouped into an ordered series of levels.

A generic construct map is shown in Fig. 2.1. The variable being measured is called "X." The depiction shown here is used throughout this book, so a few lines are used to describe its parts before moving on to examine some examples. The arrow running up and down the middle of the map indicates the continuum of the construct, running from "low" to "high." The left-hand side indicates qualitatively distinct groups of respondents, ranging from those with high "X" to those with low "X." A respondent construct map would include only the left side. The right-hand side indicates qualitative differences in item responses, ranging from responses that indicate high "X" to those that indicate low "X." An item-response construct map would include only the right side. A full construct map has both sides represented.

Note that this depicts an idea rather than being a technical representation. Indeed, later this idea is related to a specific technical representation, but for now just concentrate on the idea. Certain features of the construct map are worth highlighting.

1. There is no limit on the number of locations on the continuum that could be filled by a student (or item-response label). Of course one might expect that there will be limitations of accu-

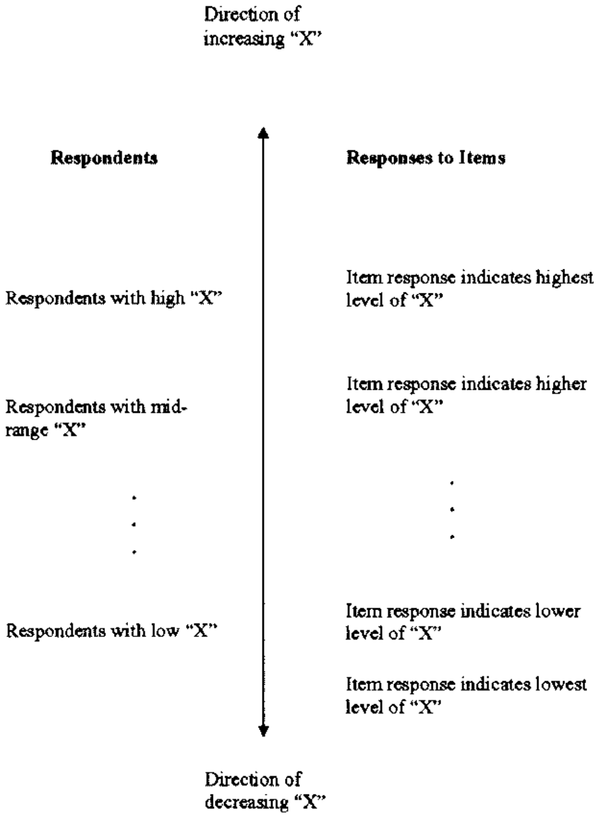

FIG. 2.1 A generic construct map in construct "X."

racy, caused by limitations of data, but that is another matter (see chaps. 5 and 6).

2. The item labels are actually summaries of responses. Although one might tend to reify the items as phenomena in their own right, it is important to keep in mind that the locations of the labels are not the locations of items per se, but are really the locations of certain types of responses to the items. The items' locations are represented via the respondents' reactions to them.

Of course words like *construct* and *map* have many other usages in other contexts, but in this book they are reserved for just this purpose. Examples of constructs that can be mapped abound: In attitude surveys, for example, there is always something that the respondent is agreeing to or liking or some other action denoting an ordering; in educational testing, there is inevitably an underlying idea of increasing correctness, of sophistication or excellence; in marketing, there are some products that are more attractive or satisfying than others; in political science, there are some candidates who are more attractive than others; and in health outcomes research, there are better health outcomes and worse health outcomes. In almost any domain, there are important contexts where the special type of construct that can be mapped is important.

A construct can be most readily expressed as a construct map, where the construct has a single underlying continuum—implying that, for the intended use of the instrument, the measurer wants to array the respondents from high to low, or left to right, in some context. Note that this does not imply that this ordering of the respondents is their only relevant feature. Some would see that measurement can only be thought of in such a context (e.g., Wright, 1977). There are good reasons for taking such a position, but the arguments involved are not necessary to the development in these chapters. In this book, the argument is that this is a good basis for instrument construction—the argument is not carried through to show that such an assumption is required.

There are several ways in which the idea of a construct map can exist in the more complex reality of usage—a construct is always an ideal; we use it because it suits our theoretical approach. If the theoretical approach is inconsistent with the idea of mapping a construct, it is hardly sensible to use a construct map as the fundamental approach—an example would be where the theory was based on an unordered set of latent classes. There are also constructs that are more complex than construct map, yet contain construct maps as a component. Probably the most common would be a multidimensional construct (e.g., the three LBC strands). In this sort of situation, to use the construct mapping approach, it is necessary merely to focus on one dimension at a time. Another common case is that where the construct can be seen as a partially ordered set of categories, such as where learners use different solution strategies to solve a problem. In this situation, the partial ordering can be used to simplify the problem so that it is collapsed into a construct map. In this case, there will be a loss of information, but this simplified construct may prove useful, and the extra complications can be added back in later. For other examples of more complex structures, see the Resources section at the end of this chapter.

Consider the LBC example introduced in the previous chapter. Here the construct described in Fig. 1.1 can be re-expressed as a construct map as in Fig. 2.2. The levels given in Fig. 1.1 are essentially different levels of student thinking, so consequently they are given on the left-hand side of the construct map.

2.2 EXAMPLES OF CONSTRUCT MAPS

The idea of a construct map is natural in the context of educational testing. It is also just as amenable to use in other domains where it is less common. For example, in attitude measurement one often finds that the underlying idea is one of increasing or decreasing amounts

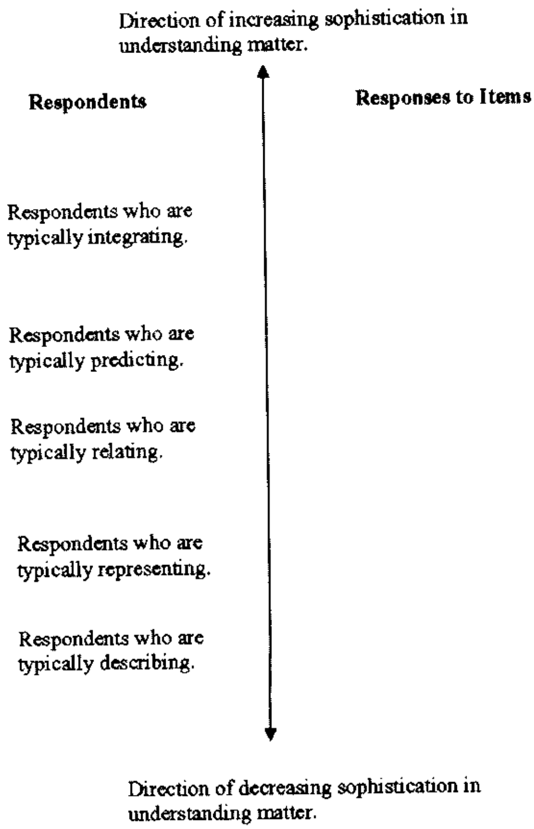

FIG. 2.2 A sketch of the construct map for the matter construct of the LBC instrument.

of something, and that something might be satisfaction, liking, agreement, and so on. The construct map is also applicable in a wide variety of circumstances as illustrated next.

2.2.1 The Health Assessment (PF-10) Example

An example of a self-report attitude-like construct that can be mapped in this way is the Physical Functioning subscale (PF-10; Raczek et al., 1998) of the SF-36 health survey (Ware & Gandek, 1998). This instrument is used to assess generic health status, and the PF-10 subscale assesses the physical functioning aspect of that. The items of the PF-10 consist of descriptions of various types of physical activities to which the respondent may respond that they are *imited a lot*, *a little*, or *not at all*. The actual items in this instrument are given in Table 5.2. An initial construct map for the PF-10 is shown in Fig. 2.3. Note the sequence of increasing ease of physical functioning as indicated by the order of the item responses. This sequence ranges from very much more strenuous activities, such as those represented by the label "Vigorous Activities," down to activities that take little physical effort for most people. Note that the order shown indicates the relative difficulty of reporting that the respondents' activities are *not limited at all*.

2.2.2 The Science Assessment (IEY) Example

This example is an assessment system built for a middle school science curriculum, "Issues, Evidence and You" (IEY; Science Education for Public Understanding Program, 1995). The SEPUP at the Lawrence Hall of Science was awarded a grant from the National Science Foundation in 1993 to create year-long issues-oriented science courses for the middle school and junior high grades. In issues-oriented science, students learn science content and procedures, but they are also required to recognize scientific evidence and weigh it against other community concerns, with the goal of making informed choices about relevant contemporary issues or problems. The goal of this approach is the development of an understanding of the science and problem-solving approaches related to social issues without promoting an advocacy position. The course developers were interested in trying

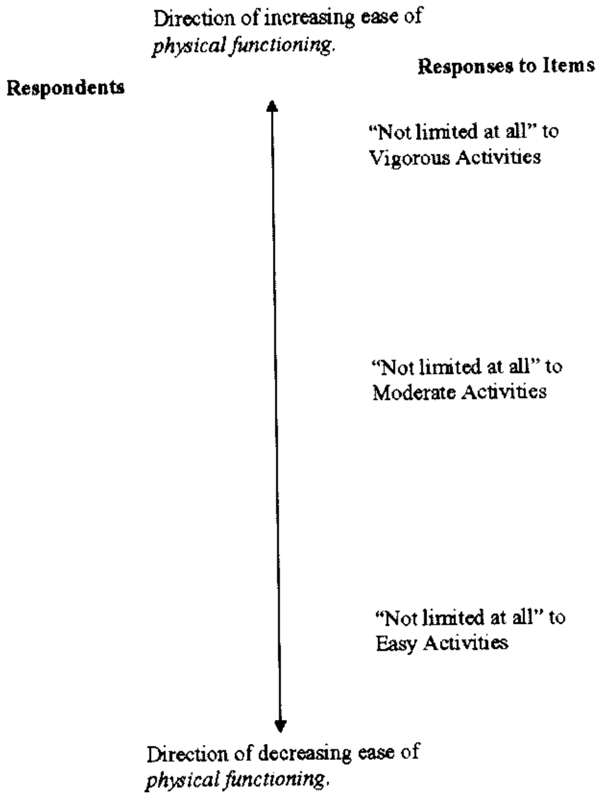

FIG. 2.3 A sketch of the construct map for the Physical Functioning subscale (PF-10) of the SF-36 Health Survey.

new approaches to assessment in the *Issues, Evidence, and You* course materials for at least two reasons. First, they wanted to reinforce the problem-solving and decision-making aspects of the course—to teachers and to students. Traditional fact-based chapter tests would not reinforce these aspects and, if included as the only form of assessment, could direct the primary focus of instruction away from the course objectives the developers thought were most important. Second, the developers knew that, to market their end product, they would need to address questions about student achievement in this new course, and traditional assessment techniques were not likely to demonstrate student performance in the key objectives (problem solving and decision making).

Both the IEY curriculum and its assessment system is built (which, like the LBC example, uses the *BEAR Assessment System* as its foundation; Wilson & Sloane, 2000) on four constructs. The *Understanding Concepts* construct is the IEY version of the traditional "science content." The *Designing and Conducting Investigations* construct is the IEY version of the traditional "science process." The *Evidence and Trade-offs* construct is a relatively new one in science education. It is composed of the skills and knowledge that would allow one to evaluate, debate, and discuss a scientific report such as an environmental impact statement and make real-world decisions using that information. The *Communicating Scientific Information* construct is composed of the communication skills that would be necessary as part of that discussion and debate process. The four constructs are seen as four dimensions on which students will make progress during the curriculum and are the target of every instructional activity and assessment in the curriculum. The dimensions are positively related because they all relate to science, but are educationally distinct.

The Evidence and Trade-offs (ET) construct was split into two parts (called *elements*) to help relate it to the curriculum. An initial idea of the Using Evidence element of the ET construct was built up by considering how a student might increase in sophistication as he or she progressed through the curriculum. A sketch of the construct map for this case is shown in Fig. 2.4. On the right side of the continuum is a description of how the students are responding to the ET items.

2.2.3 The Study Activities Questionnaire (SAQ) Example

An example of a construct map in a somewhat different domain can be found in the Study Activities Questionnaire (SAQ; Warkentin, Bol, & Wilson, 1997). This instrument is designed to survey students' activities while studying; it is based on a review of the literature in the area (Thomas & Rohwer, 1993) and the authors' interpretation of the study context. There are several dimensions mapped out in the instrument; but the focus here is on the "Learning Effectiveness" dimension of the "Effort Management" hierarchy. The authors referred to it as a *hierarchy* because they saw that each successive level could be built on the previous one—note that the hierarchy in this case is not necessarily seen as one that is inevitable—students could engage in planning without self-monitoring—but the authors saw this ordering as being the most efficacious. For the purposes of this instrument, *effort management* is the set of metacognitive and self-

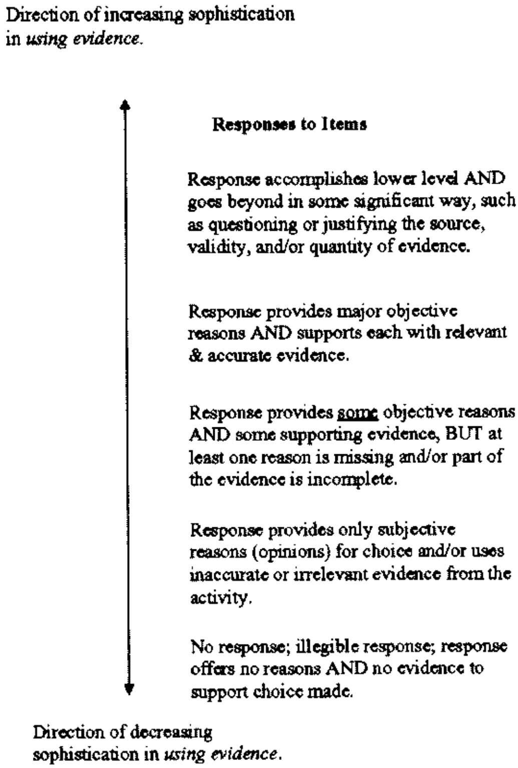

FIG. 2.4 A sketch of the construct map for the Using Evidence construct of the IEY ET constructs.

regulatory processes involved in planning and evaluating one's concentration, time, and learning effectiveness. The instrument posits four levels of increasing proficiency in effort management that form a continuum of proficiency, with each higher level subsuming lower level activities (see Fig. 2.5).

The first level is *monitoring*—being aware of one's learning effectiveness. For example, students might note how well they are learning the ideas in a paragraph by stopping at the end of the paragraph and recalling the main points. The second level, *self-regulation*, involves using the self-knowledge gained from monitoring to redirect or adjust one's behaviors. For example, if students noted that there seemed to be something missing in recalling the main points of the paragraph, they might re-read the paragraph or make a list of the main points. The third level, ***planning***, occurs when students develop a plan (before or during study) to manage or enhance their efforts. For example, students might decide to always stop at the end of each paragraph to see how well they had understood the content. Finally, at the fourth level, ***evaluation***, students would pause at the end of a study effort, reflect on the success of their plan, and consider alternatives. For example, they might find that they had indeed been understanding all the major points of each paragraph, and thus might conclude that the constant interruptions to the reading were not warranted. The questions were administered on a computer, and the administration of subsequent items was sometimes dependent on answers to previous ones—for example, if students said that they did not monitor,

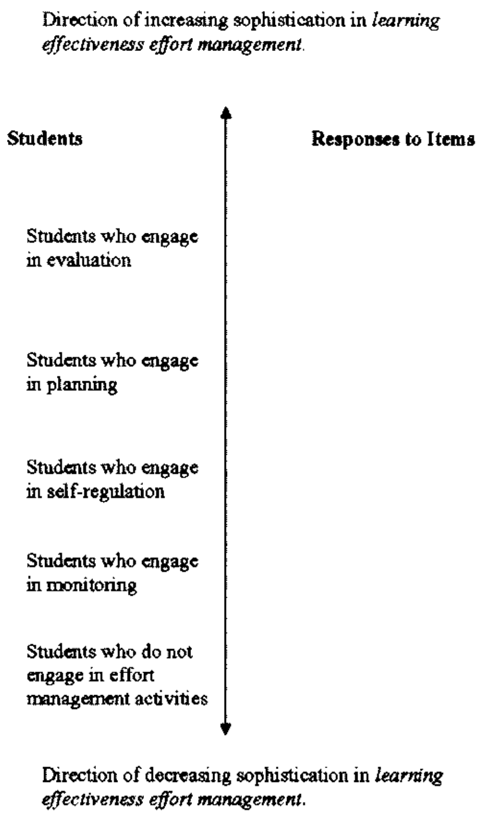

FIG. 2.5 A sketch of the construct map for the "Learning Effectiveness" construct in the "Effort Management" part of the SAQ.

then they were not asked about self-regulation (although they were still asked about planning and evaluation; see Fig. 2.5).

2.2.4 The Interview (Good Education) Example

Interviews can also be used as the basis for developing construct maps. Dawson (1998) used the Good Education Interview in a clinical interview format developed by Armon (1984) to investigate the complexity of arguments used by adults discussing issues about quality of education. She used questions such as, "What is a good education?" and "What are the aims (goals, purposes) of a good education?," along with probes such as "Why is that good?," to explore the interviewees' thinking. Interviewees' responses were divided into scorable arguments (Stinson, Milbrath, & Reidbord, 1993), and these were then scored with Commons' Hierarchical Complexity Scoring System (HCSS; Commons et al., 1983, 1995). The resulting construct map is shown in Fig. 2.6. The respondent levels on the left-hand side are stages in the HCSS scheme. The responses on the right-hand side show typical statements made by people at the corresponding levels. Note that this is the first example shown where both sides of the construct map are populated.

2.2.5 A Historical Example: Binet and Simon's Intelligence Scale

The earliest example I have found of a construct map was in Binet and Simon's (1905) description of their Measuring Scale of Intelligence. Binet and Simon identified tasks that they considered to be examples of differing levels of "intelligent behavior" and that could be easily administered and judged. By grouping these into sets that could typically be successfully performed by children of varying ages (and adults), they could set up a set of expectations for what a "normal" child should be able to do as she progressed toward adulthood. An example of such an item is "Arrangement of weights." They described it thus (Note that they have included a scored outcome space for this item in their description.):

> Five little boxes of the same color and volume are placed in a group on a table. They weigh respectively 2, 6, 9, 12, and 15 grams. They are

Direction of increasing *Argument Complexity.*

Respondents	**Responses:** *A good education …*
Metasystematic: The notion that good learning takes place in social interactions is coordinated with the idea that learning is discursive. Learning is viewed as a dialectical process in which teacher and student (or student and student) get caught up in the playful activity of learning. Testing—as a continuous spiral of feedback—is one way of conceptualizing this playful back-and-forth. This dialectic defines the learning process.	is one in which teaching involves constant testing requires a dialectical engagement with the learning process
Systematic: The notion of active participation in learning is coordinated with the idea that intellectual engagement can be increased through social interaction to produce the idea that good learning takes place in a discursive, participatory context. This is not the defining context for learning, but is the most enjoyable.	includes conversation/discussion includes group activities
Formal: Active engagement in learning is central to the learning process. The concept, interest, is differentiated into concepts like involvement, engagement, and inspiration, all of which point both to the absorption of the learner. Inspiration, stimulation, involvement, and engagement are generated by others (teachers). Social interaction is important insofar as it enhances engagement.	is one in which students are encouraged to ask questions is stimulating/involving/engaging includes social interaction includes active/experiential learning
Abstract: Knowledge acquisition is enhanced when students have fun. Interest motivates learning because it makes learning fun. Fun learning is interesting learning. Certain fun or playful activities are explicitly seen as educational. It is the teacher's job to make learning interesting.	includes playing games/doing fun things includes learning through play is one in which subjects/teachers are interesting is one in which learning is fun
Concrete: A good school (for concrete children, education equals school) is one in which you get to play and have fun. The child does not connect concepts of fun and play to learning.	includes play is one in which students have fun

Direction of decreasing *Argument Complexity.*

FIG. 2.6 A sketch of the construct map for the items in the Good Education Interview.

shown to the subject while saying to him: "look at these little boxes. They have not the same weight; you are going to arrange them here in their right order. Here to the left first the heaviest weight; next, the one a little less heavy; here one a little less heavy; here one a little less heavy; and here the lightest one."

There are three classes to distinguish. First the subject who goes at random without comparing, often committing a serious error, four degrees for example. Second the subject who compares, but makes a

> slight error of one or two degrees. Third the one who has the order exact. WE propose to estimate the errors in this test by taking account of the displacement that must be made to re-establish the correct order. Thus in the following example: 12, 9, 6, 3, 15,—15 is not in its place and the error is of 4 degrees because it must make 4 moves to find the place where it belongs. All the others must be changed one degree. The sum of the changes indicates the total error which is of eight degrees. (pp. 62–63)

The corresponding construct map is shown in Fig. 2.7. Sets of tasks that children tended to perform successfully at approximately the same age are shown on the right, and the corresponding ages (being descriptions of the respondents) are shown on the left. Binet and Simon used this construct to describe children with developmental

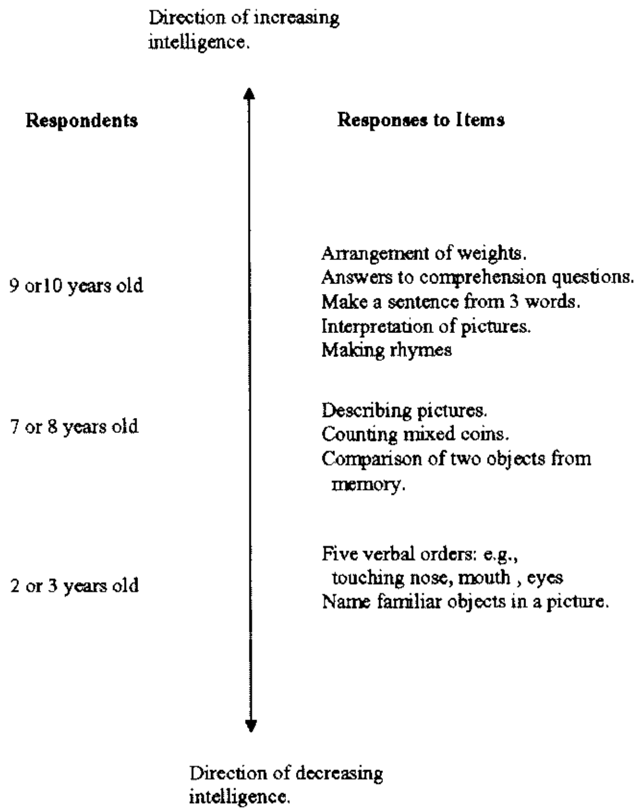

FIG. 2.7 A sketch of the construct map for Binet and Simon's (1905) Measuring Scale of Intelligence.

problems in French asylums: Those who could not succeed at the 2- to 3-year-old tasks were classified as "idiots," those who could succeed at that level but could not succeed at the 7-to 8-year-old tasks were classified as "imbeciles," and those who could succeed at that level but could not succeed at the next level were classified as "debile." Interestingly enough, they found children in French asylums who had been diagnosed into these classifications, but were actually succeeding at a level above that typical of their age.

2.3 USING CONSTRUCT MAPPING TO HELP DEVELOP AN INSTRUMENT

The central idea in using the construct mapping concept at the initial stage of instrument development is for the measurer to focus on the essential feature of what is to be measured—in what way does an individual show more of it and less of it—it may be expressed as from "higher to lower," "agree to disagree," "weaker to stronger," or "more often to less often," the particular wording dependent on the context. However, the important idea is that there is a qualitative order of levels inherent in the construct—and underlying that there is a continuum running from more to less—that allows it to be thought of as a construct map. One successful way to approach it is to think of the extremes of that continuum (say "novice" to "expert," or in the context of an attitude toward something, "loathes" to "loves"), make them concrete through descriptions, and then develop some intermediate stages or levels between the two extremes. It is also helpful to start thinking of typical responses that respondents at each level would give to first drafts of items (more of this in the next chapter).

Before this can be done, however, the measurer often has to engage in a process of "variable clarification," where the construct to be measured is distinguished from other, closely related, constructs. Reasonably often the measurer finds that there were several constructs lurking under the original idea—the four building blocks method can still be applied by taking them one at a time.

In creating a construct map, the measurer must be clear about whether the construct is defined in terms of who is to be measured, the respondents, or what responses they might give—the item responses. Eventually both will be needed, but often it makes sense in a specific context to start with one rather than the other. For instance, on the one hand, when there is a developmental theory of how individuals increase on the construct or a theory of how people array themselves between the extremes of an attitude, the respondent side is probably developed first. On the other hand, if the construct is mainly defined by a set of items and the responses to those items, it is probably easier to start by ordering the item responses.

2.4 RESOURCES

Examples of construct maps are given in the series of references cited in the Resources section of chapter 1. However, few of them incorporate both the respondent and item response sides of the continuum.

One important issue is that one needs to distinguish constructs that are amenable to the use of construct mapping and constructs that are not. Clearly any construct that is measured using a single score for each person can be a candidate for mapping. If a construct is a series of such, then each in turn can be seen as a construct map. Also exemplified earlier were constructs that are partially ordered—these too can be simplified so that they can be treated as construct maps.

The major type of construct that is not straightforwardly seen as a candidate for mapping is one where there is no underlying continuum—where, for example, there is assumed to be just a set of discrete, unordered categories. This is seen in areas such as cognitive psychology, where one might assume that there are only a few strategies available for solving a particular problem. Latent class analysis (e.g., Collins & Wugalter, 1992) is an approach that posits such a construct; it should be used when the measurer is seriously wanting to use that as the basis for reporting.

When there is an order (perhaps partial) between the latent classes, such as an increasing complexity in the nature of the strategies, then other possibilities arise. For example, one could have the strategies treated as observed categories with an underlying latent continuum of increasing sophistication (e.g., Wilson, 1992a, 1992b).

One could also try and combine the two types of constructs, adding a construct map within classes (e.g., Wilson, 1989; Mislevy & Wilson, 1996) or adding a dimension as a special class (e.g., Yamamoto & Gitomer, 1993). Increasingly complex combinations of all of these are also possible, leading to some complicated possibilities (see Junker, 2001; National Research Council, 2001).

2.5 EXERCISES AND ACTIVITIES

(following on from the exercises and activities in chap. 1)

1. Lay out the different constructs involved in the area in which you have chosen to work. Clarify the relationships among them and concentrate on one.
2. For your chosen construct, write down a brief (1–2 sentences) definition of the construct. If necessary, write similar definitions of related constructs to help distinguish among them.
3. Describe different levels of the construct—start with the extremes and then develop qualitatively distinguishable levels in between. Distinguish between levels among the respondents and levels in potential item responses. Write down the successive levels in terms of both aspects, if possible at this point.
4. Take your description of the construct (and any other clarifying statements) to a selected subset of your informants and ask them to critique it.
5. Try to think through the steps outlined earlier in the context of developing your instrument, and write down notes about your plans.
6. Share your plans and progress with others—discuss what you and they are succeeding on and what problems have arisen.