

From: Stiles, W. B. (1992). *Describing talk: A taxonomy of verbal response modes*. Newbury Park, CA: Sage.

## Chapter 1

### Development of the Taxonomy

Jerry Goodman introduced me to response modes in the spring of 1969, while I was a graduate student in clinical psychology at UCLA and he was my clinical supervisor. He used the modes to describe my alternatives as I tried to listen and respond helpfully to my psychotherapy clients. He distinguished six categories of "help intended" communication: (1) Question, for gathering information, (2) Advisement, for guiding another's behavior, (3) Silence, for providing interpersonal space, (4) Interpretation, for explaining or classifying another's behavior, (5) Reflection, for expressing empathy, and (6) Disclosure, for revealing one's personal condition (Goodman & Dooley, 1976; Goodman & Esterly, 1988).

I was attracted to the response modes at first by the therapeutic power of Reflections. I was fascinated by the process-facilitating effects of "simply" repeating a client's communication, in comparison to, say, the process-deflecting effects of Questions. Reflections were followed by deeper exploration; Questions were followed by a change in direction. In the tapes of therapy, in my work as a therapist, and in my own experience of being listened to, response modes seemed to make a dramatic difference.

My encounter with Jerry and response modes occurred in the middle of my graduate training, and its influence on my practice and research was delayed a few years. By late 1973, however, as an assistant professor of psychology at the University of North Carolina, I had begun to study the response modes seriously. I had reread Rogers's (1951) *Client-Centered Therapy* and adopted a very non-directive style of psychotherapy, therapy supervision, and classroom teaching. For me, Reflections were a core technique. I decided to investigate the effects of response modes in psychotherapy, grinding the axe that Reflections were more therapeutic than other modes. (That's what I thought then, though I no longer think the world is so simple.) I hypothesized that client-perceived "high spots" in therapy sessions would be associated with Reflections more than with other therapist responses. I gathered tape-recorded therapy sessions and postsession questionnaires about client's in-session experiences. This project eventually foundered on the difficulty of identifying high spots from the client descriptions (but see Elliott & Shapiro, 1988, for a more successful procedure). Nevertheless, it focused my attention on classifying therapists' interventions.

From the outset, the Goodman response mode framework

seemed to me to carve nature at the joints. I assumed that each therapist response would fit in one and only one of the six modes. I expected to find "litmus tests" -- mental experiments for distinguishing between categories in difficult cases. (I use the term litmus test by analogy with the chemical test in which a strip of litmus paper dipped in a solution turns pink or blue depending on whether the solution is acidic or basic.) For example, I knew that Interpretation and Reflection were distinguished by the viewpoint, or frame of reference adopted in a therapist's intervention. Responses that take the therapist's viewpoint, such as "You seem to restrict your choices," are Interpretations, whereas responses that take the client's viewpoint, such as "You wish you could take more risks," are Reflections. Confronted with a difficult discrimination between Interpretation and Reflection, I could ask, "whose frame of reference is being used?" and solve the puzzle. Student coders and I identified a variety of litmus tests as we tried to improve coding reliability.

In October, 1974, Brian Premo (then a first-year graduate student) and I found a litmus test to distinguish between Interpretations and Advisements: whether the topic of the response was drawn from the experience of the therapist or of the client. Interpretations concerned the client's experience or ideas whereas Advisements (a directive category that includes commands, suggestions, permission, and prohibition) concerned the therapist's experience or ideas. For example "You have only hurt yourself by lying about it" concerns the client's feelings and is an Interpretation, whereas "You might as well tell the truth" concerns the therapist's idea of what to do and is an Advisement.

In a key insight, we realized that the principle underlying this litmus test, which we called *source of experience* (i.e., client's experience as topic versus therapist's experience as topic), applied equally to any therapist response and helped to describe all of the modes. Not only Interpretations, but also therapist Questions and Reflections clearly concerned the client's experience, whereas therapist Disclosures as well as Advisements clearly concerned the therapist's experience. (We used the term "experience" broadly, to include knowledge, ideas, beliefs, feelings, memories, perceptions, intentions, and even voluntary behaviors.) Silence was slightly problematic, because a silence didn't involve anybody's experience explicitly, but in the Goodman framework, therapist Silence was considered as providing interpersonal space for the client, so we classified it as concerning the client's experience.

Intrigued, we asked: are there other principles that, like source of experience, can distinguish between sets of modes? We realized we had been using such a principle already, frame of reference (i.e., client's viewpoint versus therapist's viewpoint), though we so far had applied it only to distinguish Reflections from Interpretations. Not only in Reflections, however, but also in Goodman-conceived Silences, therapists adopted the client's frame of reference, whereas in Advisements, Disclosures and Questions, as well as in Interpretations, therapists used their own frame of reference.

Cross-classifying the modes by these two *principles of classification*, source of experience and frame of reference, seemed to help characterize the modes. The principles described what therapists were doing in the interventions psychologically. In Reflection ("You feel proud of your promotion"), the therapist attempted to express the client's experience as viewed from within the client's own frame of reference -- seeing the experience as the client saw it; indeed, Rogers (1951) described Client-Centered Reflections in very similar terms. In Interpretation ("Perhaps you were over-reacting"), the therapist also talked about the client's experience, but placed it in some external frame of reference -- the therapists' own or perhaps some psychological theory -- sometimes bringing a new perspective to the client's experience. Therapist Questions ("How did that make you feel?") could be construed as a revealing a gaps in the therapist's frame of reference and seeking information from the client to fill the gaps. Therapist Disclosures ("I'm finding it hard to listen today") expressed the therapist's experience (thoughts or feelings) from the therapist's own viewpoint.

Representing the cross-classification of the six Goodman modes by the two principles as a 2x2 matrix (Table 1-1) revealed a striking (to me) incompleteness and asymmetry. There was a blank cell, which reminded me of missing elements in the chemical periodic table. And although there were two modes within each cell, the within-cell pairings involved surprisingly dissimilar modes. For example, Interpretation and Question were paired, whereas to me, Interpretation seemed intuitively more similar to Reflection or to Advisement than to Question. Perhaps, we thought, there is some third principle of classification that could distinguish between the modes in each pair. We spent several months identifying and refining our concepts of the third principle and two missing modes.

The third principle, which we identified before the seventh and eighth modes, distinguished Reflection, Interpretation, and Advisement from their Table 1-1 cell mates, Silence, Question, and Disclosure, respectively. The distinguishing feature was whether the therapist had to presume knowledge of the client's experience in order to make the response. In Reflection, Interpretation, and Advisement, the therapist must presume to know what the client's experience is, was, will be, or should

Table 1-1

Initial Classification of Therapist Response Modes by Source of Experience and Frame of Reference

	Frame of reference	
	Client	Therapist
Client	<b>Reflection</b>  <b>Silence</b>	<b>Interpretation</b>  <b>Question</b>
Therapist		<b>Advisement</b>  <b>Disclosure</b>

be; in the other modes, no such presumption is required. For example, the Interpretation, "You are exaggerating it's importance," presumes to know what the client's experience is, whereas the Question, "How important is this to you?" requires no such presumption (the speaker may make presumptions privately, but these are not necessary for the Question to have the meaning it has). The Advisement, "Tell me about your job," presumes to know what the client should do (in effect, imposing an experience), whereas the Disclosure, "I'm glad you told me that," does not. The Reflection, "You were frightened by his threats," presumes to understand the client's feelings, whereas a Silence, allowing space, does not.

The name for this presuming principle was problematic. We called it "full versus empty" at first, and we used the name "focus" for several years, but I think it is most informative to call it *presumption about experience*. Three of the modes in Table 1-1 required presumptions about the client's experience (i.e., Reflection, Interpretation, and Advisement), whereas the other three require no such presumption (Silence, Question, Disclosure) and could be said to require the therapist to presume knowledge only of his or her own experience.

The intersection of the three principles pointed us toward the missing modes. Both concerned the therapist's experience and used the client's frame of reference, but they differed in their presumptions -- one presuming and one not presuming to know what the client's viewpoint actually was. Although we began with only these characterizations, in the end these two modes seemed as distinct and coherent as the other six.

Therapist *Edifications* concern the therapist's experience but use the client's frame of reference while making no specific presumptions about the client. There is a paradox in using the client's frame of reference while presuming no knowledge of the client's experience. The resolution is that the frame of reference is a neutral, general viewpoint shared by *any* other person. Thus, Edifications are representations about objective

reality -- simple informational statements, such as "She told me she was going shopping," or "The clinic is on Oak Street." In retrospect, it seemed surprising that we had not included such a category earlier, as Edifications are common and recognizably distinct. They may have been left out of the Goodman framework because most therapists convey facts relatively rarely, and when they do, they often have some ulterior (e.g., interpretive or directive) purpose.

Naming this category was difficult, and our meaning for "Edification" differs from the common meaning of "edification," referring only to conveying objective information, without necessarily implying that the client is thereby improved. Alternative names such as "Assertion" or "Information" or "Report" seemed too broad.

Therapist *Confirmations* also concern the therapist's experience and use the client's frame of reference but *do* presume knowledge of the client's experience. These are expressions of agreement, disagreement, or shared experience, such as "We make a great team," or "We disagree about the importance of homework," or "We seem to have reached an impasse." Such responses convey the therapist's ideas or opinions (therapist's experience), but explicitly compare these to the client's viewpoint. To make such a comparison, the therapist must presume to understand this particular client's point of view.

The name "Confirmation" does not imply positivity or agreement; "We disagree" is just as much a Confirmation as "We agree." The name was chosen for its connotation of interpersonal connection; one confirms another person by using his or her viewpoint. Both agreement and disagreement entail presuming to know what the other person thinks.

As a final change in the list of modes, we switched from Silence to *Acknowledgment*. The Silence category had been a continuing problem. Alone among the modes, it was a nonverbal response, whereas all of the others were verbal. Coding pauses in the conversation had presented practical problems: When had a Silence taken place? (How long did the pause need to be?) How many Silences should be coded? (Should a long pause be coded as more than one response?) Who should receive credit? (When was a pause to count as a therapist Silence and when as a client Silence?) After much discussion, and with considerable reluctance, I became convinced that Silence belongs to another realm of phenomena than the other modes, a nonverbal or paralinguistic realm that might encompass interruptions, verbal crowding and allowing, and the management of attention in conversation. This realm is no less important, but it is conceptually separate (see chapter 3). The verbal response mode category that concerns the other's experience and uses the other's frame of reference but presumes no knowledge of either is better called *Acknowledgment*.

Acknowledgments are brief, contentless utterances such as "mm-hm" and "oh." We had had to confront such responses early in our work, because psychotherapists used them very frequently. We had reasoned that "mm-hm" and similar utterances were a sort of vocalized Silence, and we coded them as Silences (this now sounds like stretching, but it seemed reasonable at the time). They fit the principles well. The topic of a therapist's "mm-hm" was whatever the client had just said (client's experience). The frame of reference adopted was the client's (i.e., the "mm-hm" merely signaled receipt of the client's communication). And no specific presumption about the client's experience was entailed; in fact, therapists often said "mm-hm" before the client had completed a thought, so no coherent presumption was possible. Thus, when we changed names from Silence to Acknowledgment, we already had in mind a coherent category of verbalizations. The switch made it easier for us to accommodate contentless lexical utterances such as "yeah," "no," and "hello" in the category; these fit the principles and seemed appropriately described as Acknowledgments, whereas we had been unable to bring ourselves to call them Silences.

### Generalizing the Taxonomy

I had initially considered the response mode system as a classification of therapist interventions to clients in psychotherapy, or of helpers to help-seekers in circumstances analogous to therapy. I had begun by calling the categories as "listener response modes" (Stiles, 1975), designed to classify the "help-intended communication" by a "listener" who listened helpfully while a "discloser" revealed a personal concern (see discussion of Goodman system in chapter 3). We had called the values of the three principles "therapist" and "client," or even more confusingly, "listener" and "discloser." It gradually became clear that this construction was far too limited, and that the categories could be used to describe any verbal communication from one person (or collectivity) to another.

In recognition of this far greater generality, we began calling the values of the principles "speaker" and "other". For example, an Interpretation is described as concerning the other's experience and presuming knowledge of the other's experience, but using the speaker's frame of reference. Both members of a dyad (e.g., client as well as therapist) can be speakers in turn, and the taxonomy applies equally to both. Thus, "speaker" characterizes a momentary role or activity, rather than a fixed role (such as "therapist"). The "other" can be not just another person, but any entity to which talk is directed, including a group or an audience or even a nonhuman, inanimate, or abstract entity.

**Table 1-2**

**Taxonomy of Verbal Response Modes**

<i>Source of experience</i>	<i>Presumption about experience</i>	<i>Frame of reference</i>	
		<i>Other</i>	<i>Speaker</i>
Other	Other	<b>Reflection (R)</b>	<b>Interpretation (I)</b>
	Speaker	<b>Acknowledgment (K)</b>	<b>Question (Q)</b>
Speaker	Other	<b>Confirmation (C)</b>	<b>Advisement (A)</b>
	Speaker	<b>Edification (E)</b>	<b>Disclosure (D)</b>

We also began calling the categories "verbal response modes" (VRMs), in recognition of their applicability to any verbal communication. The categories are theoretically universal, insofar as any verbal communication must be from one center of experience (the speaker) to another (the other). As a general-purpose system, the VRM taxonomy permits direct comparisons across roles (e.g., therapist versus client, doctor versus patient, husband versus wife, parent versus child), as well as across types of relationships or types of discourse. The VRM system has been used to study verbal communication in medicine, law, politics, education, entertainment, advertising, and social and family relations, as well as in psychotherapy (see chapter 4), and many other applications are possible.

With the addition of Edification and Confirmation and the changes from Silence to Acknowledgment and from "therapist" and "client" to "speaker" and "other," the taxonomy took on a form that has remained stable in its essentials since about 1976. A published manual appeared in 1978 (Stiles, 1978a). This stable version is summarized in Table 1-2 and is fully elaborated in chapters 5 to 10.

An important change that emerged gradually while the system was being developed was a shift to *defining* the modes by the principles of classification. The modes began as intuitively recognized classes of utterances, and the principles began as coding aids, or litmus tests, intended to help make distinctions between modes. However, with the 2x2x2 taxonomy complete (Table 1-2), it became possible to reverse this and define the modes in terms of the principles. In effect, the system pulled itself up by its bootstraps. Coders can classify any utterance by answering three questions: Whose experience is the topic? Does the utterance require the speaker to presume knowledge of the other's experience? And whose frame of reference is used? These three forced choices place

any utterance in one of the eight VRM categories.

This changeover required surprisingly little alteration in the contents of the categories. The original Goodman modes were built on common intuition, extensive observation, and clinical experience, and they approximated categories that had been widely used in theory, research, and counselor training. The principle-defined categories are essentially identical, still easily recognizable and intuitively distinct.

I think that the principles of classification help to explain why these particular categories seem natural and hence why the principle-defined categories are essentially the same as the intuitively-defined ones. Each principle takes the values of "speaker" and "other," reflecting the most obvious line of demarcation in any interpersonal communication (there are no plausible intermediate values). Thus, the distinctness and coherence of the mode categories reflect the separateness of the people who are communicating.

The principles offer a basis for saying that the system is a mutually exclusive and exhaustive classification. By deciding (a) whether the source of experience is the speaker or the other, (b) whether the utterance required presumptions about the other, and (c) whether the speaker's or the other's frame of reference was used, a coder can place every utterance in one and only one category. Of course, "exhaustiveness" in this sense implies only that every utterance can be classified; it does not imply that all of the utterance's meaning is conveyed by the classification. Many distinctions are possible within each mode (see chapters 3 and 6). Indeed, it would be possible to make the same claim of exhaustiveness about a system that used only one principle. For example, a system that used only the source of experience principle could exhaustively classify utterances into informative (speaker's experience) and attentive (other's experience) categories.

## Forms and Mixed Modes

It had been clear all along that mixtures of the modes were possible. For example, Goodman and Dooley (1976) acknowledged the existence of "hybrid responses such as Question-in-service-of-Advisement" (p. 106), utterances that are grammatical questions but are intended as directives, for example, "Can you say that again with more feeling?" or "Is it a good idea to be wearing such a heavy sweater?"

The notion that each mode had a characteristic grammatical form, but that the grammatical form could be used to express other mode intents, was incorporated into the taxonomy systematically.

Very early in my work, I made an assumption that each mode had an associated set of grammatical features and that each utterance could be classified on the basis of these features alone, independently of its classification according to the principles. Thus, each utterance was classified twice, once for form and once for intent. The same eight mode names are used for both form and intent classification. Utterances in which form and intent coincide (e.g., Advisement in service of Advisement) are called *pure modes*; utterances in which form and intent differ (e.g., Disclosure in service of Advisement) are called *mixed modes*. The grammatical criteria for each of the eight form categories are presented and discussed in detail in chapter 7.

Each of the 64 possible form-intent combinations can be written as a two-letter code. Each mode is abbreviated by its first letter, except for Acknowledgment, which is abbreviated *K*. As a notational convention, the form symbol is written first, and the intent symbol second:

**Sit down. AA**

**Would you sit down? QA**

**I'd like you to sit down. DA**

In earlier publications, the intent symbol was enclosed in parentheses -- e.g., D(A) -- but this now seems unnecessary.

A grammatical form retains some of its force even when it is used to express a different intent. For example, the psychological force of the Question form of "Would you get me some coffee?" (QA) attenuates the Advisement intent, making it less presumptuous, more attentive, and hence relatively polite (in comparison to "Get me some coffee" AA). Such form-intent discrepancies seem often to signal conflicting social pressures. In general, directing someone to do something while minimizing the social imposition yields a realm of form-intent discrepancies that contribute to *politeness* (Brown & Levinson, 1978; Stiles, 1981). These involve using non-presumptuous forms (Question, Disclosure, Acknowledgment, Edification) to express presumptuous

intents (especially Advisement and Interpretation).

Similarly, the Acknowledgment form of yes/no answers contributes to their impact in a conversation. In a "yes" or "no" answer, a speaker can, in effect, use the other's words to convey a feeling, impart information, or give directions:

**(Are you angry with me?) No. KD (conveys feeling)**

**(Has it stopped raining?) No. KE (imparts information)**

**(Should I slow down?) No. KA (gives directions)**

Using Acknowledgment forms thus conveys attentiveness and acquiescence at a formal level, while the Disclosure, Edification, or Advisement intent conveys information or direction at an intentional level. The mixed modes seemed to me to capture this subtle, relational aspect of question/answer exchanges.

A few years after the taxonomy had been constructed, I did some reading in linguistics and philosophy (including Austin, 1975; Bach & Harnish, 1979; Grice, 1957, 1969; Searle, 1969) and realized that form and intent can be described as representing distinct levels of an utterance's meaning (Stiles, 1986b). The form codes represent an aspect of the utterance's *literal meaning*, while the intent codes represented an aspect of its *pragmatic meaning*, or, as Grice (1969) described it, its *occasion meaning* -- the meaning a speaker intended on a particular occasion. Both levels contribute to the psychological force of the utterance, as do other, off record levels of meaning (Stiles, 1986b; see chapter 5). Mixed modes thus represent indirect speech acts -- using one literal meaning to convey a different pragmatic meaning.

The long process of identifying which grammatical forms correspond to which intents has drawn on the collective linguistic intuition of the coders and collaborators who have worked on the research. By discussing problematic examples encountered in coding, we have gradually collected a set of grammatical features that characteristically express each of the eight intents. In each instance, we ask, "what is the literal meaning of this grammatical expression, when taken out of its context?" Because they are based on linguistic intuition, the form definitions are in one sense less anchored than the intent definitions, even though form is usually coded more reliably than intent. On the other hand, the present form specifications (chapter 7) rest on my discussions with literally hundreds of coders and collaborators who have coded more than two million utterances in varied types of discourse, and I have considerable confidence in them.

## Unitizing

Over the first year or two of using the VRM system, the scoring unit shrank from the *speech* (everything a speaker said

without being interrupted) to the *sentence* (everything between one initial capital letter and then next on our transcripts) to the *utterance*. As now constituted, an utterance is defined as an independent clause, a nonrestrictive dependent clause, an element of a compound predicate, or a term of acknowledgment, evaluation, or address. The scoring unit shrank as we found instances of the larger units whose sense demanded two or more different codes. The current definition has been developed to give a grammatical specification of a communicative unit for which there is one and only one VRM code. Details and examples are presented in chapter 8.

### Recent Developments

Several hundred coders have now coded several million utterances using the VRM system. Each new application of the system raises new coding issues and requires new adjustments; nevertheless, grounded in the principles of classification, the categories have remained notably stable across time and topics.

Since the publication of the 1978 manual (Stiles, 1978a), I have encountered many examples of utterances that appeared difficult to code at first. Usually coders have brought these to research project staff meetings. Our approach has been to discuss these problematic utterances in terms of the principles of classification as informed by our understanding of what the speaker meant in context. Almost always we have found a resolution that further articulates the principles without contradicting them. Typically, resolving the coding of one problematic utterance has clarified the coding of many other similar utterances. Thus, extending the taxonomy into new areas has resulted in an elaboration of earlier work rather than a contradiction of it. One product of this work is a body of "lore" dealing with how to apply the principles to particular linguistic constructions.

The changes that contradict the 1978 manual have been few and relatively minor and technical. An example of one of the largest contradictions will put this in perspective: "Well" is now coded as an Acknowledgment form, whereas it was cited as an example of a noncommunicative noise (like "uhhh" or throat-clearing, which are not coded) in 1978. Psychologically, "well" seems to Acknowledge the other's communication while holding the floor for the speaker, whereas other Acknowledgments ("mm-hm," "yeah") typically relinquish the floor to the other. But both meet the definition of Acknowledgment (Table 1-2).

**A: Where did you go after you left us last night? QQ**

**B: Well, KK**

**I'm not sure I want to say. DD**

In retrospect, this (and the other) contradicting changes appear to correct earlier misconceptions. That is, rather than being arbitrary changes, they seem to represent a better understanding of what expressions mean, and hence of how

language mediates interpersonal relationships.

Chapters 5 to 10 of this book reflect the fruits of this process of articulation and elaboration. In comparison with the 1978 manual, the formal characterization of the principles in chapter 5 is more comprehensive and, I hope, clearer. The unitizing rules in chapter 8 encompass many grammatical constructions that we had not considered in 1978. And chapters 9 and 10 in particular represent the lore and examples from coding projects that have been conducted since 1978, the result of time, attention, and thought by many coders.

## Chapter 2

### Overview of the Taxonomy

This chapter briefly describes the verbal response mode (VRM) taxonomy and lists its technical features. More detailed descriptions of the principles of classification, the VRM intent categories, and the VRM form categories (and mixed modes) are given in chapters 5, 6, and 7, respectively.

#### Principles, Intents, and Forms

Every utterance from one person to another can be considered to concern either the speaker's or the other's experience, where "experience" is understood broadly to include thoughts, feelings, perceptions, and intentional actions. For example, in the informative utterance, "I want to go fishing," the *source of experience* is the speaker, whereas in the attentive utterance, "Do you want to go fishing?" the source of experience is the other.

Further, in making an utterance, the speaker may or may not make presumptions about the other's experience. That is, the utterance may or may not require the speaker to presume to know what the other person is, was, will be, or should be thinking, feeling, perceiving, or intending. To illustrate, in saying "I want to go fishing" or "Do you want to go fishing?" the speaker need not make such a presumption; however, in saying "Go fishing" the speaker presumes to know what the other should do -- in effect, he or she seeks to impose an experience (the intention to go fishing) on the other. The former, nonpresumptuous utterances are said require a *presumption about experience* of the speaker only, whereas the latter, presumptuous utterance requires a presumption about the experience of the other in order to have the meaning it has.

Finally, in making an utterance, the speaker may represent the experience either from his or her own personal viewpoint or from a viewpoint that is shared or held in common with the other. To illustrate, "I want to go fishing," "Do you want to go fishing?" and "Go fishing" all use the speaker's *frame of reference* -- the experience is understood (or, in the case of the question, is to be understood) from the speaker's viewpoint and may be described as directive. By contrast, the more acquiescent utterance, "You want to go fishing" takes the other's frame of reference, in effect representing the experience as the other views it.

These three principles of classification -- source of experience, presumption about experience, and frame of reference -- form the basis of the VRM taxonomy described in this book. The three principles are dichotomous -- each can take the value "speaker" or "other" -- and they are orthogonal in the sense that all eight (2x2x2) combinations of them are possible. Each combination defines one of the modes, as shown in Table 2-1. For example, Questions are defined as utterances that concern the others' experience, presume knowledge of the speaker's experience only (no presumptions about the other's experience

required), and use the speaker's frame of reference. The eight VRM categories are Disclosure (D), Edification (E), Advisement (A), Confirmation (C), Question (Q), Acknowledgment (K), Interpretation (I), and Reflection (R). An Uncodable (U) category is used only for utterances that coders cannot understand or hear clearly.

Speaker and other may be individuals, or they may be collectivities, as when a speaker addresses an audience, or they may even be inanimate or imaginary entities, as when a person speaks to the heavens or when one fictional character speaks to another. The VRM codes describe the relationship between the speaker and the other, whoever or whatever they are.

The VRM taxonomy is thus a conceptually-based, general-purpose system for coding speech acts. The taxonomic categories are mutually exclusive, and they are exhaustive in the sense that every comprehensible utterance can be classified. The mutual exclusivity and exhaustiveness derive from the modes' basis in the three dichotomous principles of classification (see chapters 1 and 5 for further discussion of the principles). The abbreviation "VRM" is used in this book to refer to this particular taxonomy of verbal response modes. (See chapter 3 for references to other classification systems.)

The names of the eight modes are commonly used words; however, in this VRM taxonomy, the modes are defined by the intersection of the three principles and not by the colloquial meanings or the dictionary meanings of the names. The descriptions of the modes in the body of Table 2-1 are meant as characterizations of the categories, which are defined by that particular combination of "speaker" and "other" values on the principles. More detailed definitions and examples are given in chapter 6. In this book, the mode names are capitalized as a way of indicating that the taxonomic meaning may differ from the colloquial meaning in some cases.

In applying the taxonomy, each utterance is coded twice, once with respect to its grammatical form, or literal meaning, and once with respect to its communicative intent, or pragmatic meaning. Thus, the taxonomy includes 64 possible form-intent combinations, eight *pure modes*, in which form and intent coincide, and 56 *mixed modes*, in which they differ. As a notational convention, the form abbreviation is written first and intent second. The relation of form to intent is expressed, "in service of." For example, "Would you roll up your sleeve?" is coded QA, which is read as *Question in service of Advisement*, i.e., Question form (inverted subject-verb order) but Advisement intent (guiding the other's behavior). A brief summary of the forms is given in Table 2-2. Forms and mixed modes are more fully described in chapter 7.

**Table 2-1**  
**Taxonomy of Verbal Response Modes**

<i>Source of experience</i>	<i>Presumption about experience</i>	<i>Frame of reference</i>	<i>Mode</i>
Speaker	Speaker	Speaker	DISCLOSURE (D) Reveals thoughts, feelings, perceptions, or intentions.
		Other	EDIFICATION (E) States objective information.
	Other	Speaker	ADVISEMENT (A) Attempts to guide behavior; suggestions, commands, permission, prohibition.
		Other	CONFIRMATION (C) Compares speaker's experience with other's; agreement, disagreement, shared experience or belief.
Other	Speaker	Speaker	QUESTION (Q) Requests information or guidance.
		Other	ACKNOWLEDGMENT (K) Conveys receipt of or receptiveness to other's communication; simple acceptance, salutations.
	Other	Speaker	INTERPRETATION (I) Explains or labels the other; judgements or evaluations of other's experience or behavior.
		Other	REFLECTION (R) Puts other's experience into words; repetitions, restatements, clarifications.

UNCODABLE (U) is used only for incomprehensible utterances.

**Table 2-2**  
**Summary of Verbal Response Mode Form Criteria**

<i>Disclosure:</i>	Declarative; first person ("I") or first person plural ("we") where other is not a referent.
<i>Edification:</i>	Declarative; third person (e.g., "he," "she," "it" or a noun).
<i>Advisement:</i>	Imperative, or second person with verb of permission, prohibition, or obligation.
<i>Confirmation:</i>	First person plural ("we") where referent includes other (i.e., "we" refers to both speaker and other).
<i>Question:</i>	Interrogative, with inverted subject-verb order or interrogative words.
<i>Acknowledgment:</i>	Nonlexical or contentless utterances; terms of address or salutation.
<i>Interpretation:</i>	Second person ("you"); verb implies an attribute or ability of the other; terms of evaluation.
<i>Reflection:</i>	Second person ("you"); verb implies internal experience or volitional action.

## So What?

Each VRM category represents a type of *microrelationship* -- a way that a speaker can be related to an other for one utterance. Insofar as verbal microrelationships combine with each other and with other attributes to form human relationships, VRM coding is broadly useful for understanding and describing relationships between people (e.g., intimacy, status, social roles) as conveyed in language (see chapter 3 for a discussion). Qualitatively, VRM concepts link observable response categories (e.g., Question, Disclosure) with general psychological principles (e.g., whose experience or frame of reference is used). Aggregating VRM codes across some stretch of discourse yields numerical indexes that describe relationships (see chapter 11), permitting quantitative research on topics in which the relationship between people talking to each other figures as important. Chapter 4 illustrates some of the sorts of questions VRM coding can address, but many more uses are possible. (Investigators are continually thinking of new ones.) As noted there, aggregating utterances according to the VRM principles of classification offers a bridge between the relatively molecular level of coded utterances and the relatively molar level of relationship dimensions (e.g., attentiveness, directiveness, presumptuousness) and of theoretical concepts (e.g., theories of psychotherapy).

### Technical Features of the VRM Taxonomy

Here is a quick sketch of the VRM taxonomy's characteristics.

#### *Applicability*

The taxonomy is applicable to any discourse and has already been applied to psychotherapy, medical interviews, informal conversations, letters, family interaction, job interviews, university lectures, political speeches, labor negotiations, courtroom interrogations, radio call-in programs, and television advertisements (see chapter 4 and the reference list). Coding requires only that there be a "speaker" and an intended audience, or "other". Speaker and other may be collectivities rather than individuals. For example, the opposing sides of a labor/management negotiation (each consisting of several individuals) have been construed as alternating as "speaker" and "other" (Hinkle, Stiles, & Taylor, 1988) and television advertisements have been construed as emanating from a "speaker" addressing the collective television audience (Rak & McMullen, 1987).

#### *Access Strategy*

The VRM system uses an observational access strategy, not self-report. The system codes what people *do* in verbal interaction -- their speech acts -- not what they think.

#### *Communication Channel*

The VRM system codes verbal behavior. Other channels, including paralinguistic and kinetic communication, may be useful for identifying the mode of an utterance, but these are not coded directly (see chapter 3).

#### *Data Format*

VRMs can be coded from written documents, transcripts, audiotapes, videotapes, or live interactions. Most utterances in natural interaction can be coded accurately from verbatim transcripts; however, patterns of emphasis, timing, and inflection evident on audiotape can clarify the intent of some utterances. The additional (visual) information available from videotape or live interactions may be needed in some circumstances, for example, (a) in conversations involving several individuals, where visual information may be needed to distinguish which person is speaking and which is being addressed; (b) when speech is indistinct, where the added redundancy provided by visual information may aid in understanding what was said; or (c) when speech accompanies ongoing activity, such as children playing, where the meaning of some utterances draws on the physical activities being performed.

Live coding of complex or rapidly moving interaction is difficult and therefore likely to be inaccurate. Coders working from tapes of such interactions often need to rewind and replay the recordings to catch speakers' meanings.

Coding directly from audiotapes requires more skill than coding from transcripts; however good intercoder reliability can be obtained from audiotape coding (e.g., McDaniel, Stiles, & McGaughey, 1981; Stiles, Au, Martello, & Perlmutter, 1983; Stiles, Shapiro, & Firth-Cozens, 1988). Coders typically comment that audiotapes make it easier to understand what the speaker meant.

#### *Scoring Unit*

A scoring unit is the stretch of text to which individual codes are assigned. The VRM scoring unit is the *utterance*, defined as a simple sentence, an independent clause, a nonrestrictive dependent clause, an element of a compound predicate, or a term of acknowledgment, evaluation, or address. Details and examples are presented in chapter 8.

#### *Contextual Unit*

Contextual unit refers to the amount of text that must be considered by a coder in deciding how to classify an utterance. VRM form codes are based on grammatical features (chapter 7), so in perfectly grammatical, non-elliptical speech, utterances can be coded in isolation. In natural speech, elliptical, incomplete, and ungrammatical utterances require reference to context, though usually the immediately preceding two or three utterances are sufficient. See chapters 7 and 10

for discussion and examples.

VRM intent codes classify the speaker's intended meaning on a particular occasion and thus must always be considered in context. For virtually any utterance it is possible to imagine some other context in which its meaning would be different. To illustrate, "It's starting to rain" would be EE in most contexts, but it could be ER if it repeated something the other just said, or EA if it was understood that the other was required to close the windows, or EQ if it was spoken with an upward inflection at the end. Many of the examples given later (e.g., chapters 9 and 10) illustrate the importance of context.

In practice, VRM intent can usually be coded accurately in a context of a few preceding utterances. However, some utterances may be understandable only in the context of much earlier events. Probably this is most common in longstanding relationships, for example, among family members. Like any outsider, VRM coders may overlook meanings that are clear to participants. On the other hand, since the VRM system is usually applied only to "on record" meaning (chapter 5; Stiles, 1986b, 1987b), coders' ignorance of the more subtle, off record meanings of some utterances may not impair the validity of VRM codes.

#### *Summarizing Unit*

The summarizing unit is the stretch of discourse over which summary indexes are calculated. The VRM system allows any size of summarizing unit, from a single utterance to a whole class of relationships. Most VRM studies have summarized over an encounter or a segment of an encounter (e.g., the medical history-taking segment of a medical interview). Utterances have most often been aggregated separately for each speaker (e.g., the percentage of a client's utterances coded as Disclosure). However, it is also possible to aggregate by dyad or by larger groups. See chapter 11 for a discussion of summary indexes produced by the VRM system.

#### *Level of Measurement*

VRM categories are nominal measures. However, aggregate VRM indexes used to characterize an episode of communication may be ordinal, interval, or even ratio scales. For example, the frequency of client Disclosures in a psychotherapy session is a ratio measure (to illustrate: 80 Disclosures is twice as high as 40 Disclosures). See chapters 4 and 11 for illustrations and descriptions of VRM aggregate indexes.

#### *Transcribing*

The VRM system does not require special transcribing procedures; as indicated above, any discourse, written or spoken, can be coded. Nevertheless, following a few consistent transcribing principles will improve the accuracy and reliability of coding and unitizing. A few rules that I have found useful for transcribing are given in chapter 11. For more

detailed transcribing rules, I recommend those of Auld and White (1956). Fully phonetic transcriptions are not necessary and may be distracting; however, researchers interested in very intensive study of discourse (including other coding in addition to VRMs) may wish to consider the transcription system developed by Jefferson (e.g., Sacks, Schegloff, & Jefferson, 1974).

#### *Coder Qualifications*

Applying the VRM system requires high verbal aptitude, interest in interpersonal communication, patience with details, and intensive training and practice. Competence in basic grammar is essential. Psychological mindedness is helpful, but professional training in psychology is not required. Suggestions for training coders are offered in chapter 11.

VRM coding is a skill that is acquired gradually, through practice. One cannot become proficient simply by reading descriptions of the system, such as this book. Accuracy of coding -- and hence intercoder reliability -- can continue to improve for a considerable time. Improvement is facilitated by practice, regular feedback (e.g., comparisons with the work of other coders), and opportunity for discussion of difficult codes.

#### *Computer-Assisted Training Program*

A computer-assisted training program for VRM coders (on computer disk) accompanies this book. This program introduces the principles of classification, forms, and intents, and it offers practice to train new coders in the basic skills needed to code unitized transcripts. As described in chapter 11, additional practice is required for coders to become proficient in unitizing (following rules described in chapter 8) or in direct coding of untranscribed material (audiotapes, videotapes, or live interactions).

## References

- Auld, F., Jr., & White, A. M. (1956). Rules for dividing interviews into sentences. *Journal of Psychology*, *42*, 273-281.
- Austin, J. L. (1975). *How to do things with words*, second edition. Oxford, England: Clarendon Press.
- Bach, K., & Harnish, R. M. (1979). *Linguistic communication and speech acts*. Cambridge, MA: MIT Press.
- Brown, P., & Levinson, S. (1978). Universals in language usage: Politeness phenomena. In E. N. Goody, (Ed.), *Questions and politeness: Strategies in social interaction* (pp. 56-324). Cambridge, England: Cambridge University Press.
- Elliott, R., & Shapiro, D. A. (1988). Brief structured recall: A more efficient method for identifying and describing significant therapy events. *British Journal of Medical Psychology*, *61*, 141-153.
- Goodman, G., & Dooley, D. (1976). A framework for help-intended interpersonal communication. *Psychotherapy: Theory, Research and Practice*, *13*, 106-117.
- Goodman, G., & Esterly, G. (1988). *The talk book: The intimate science of communicating in close relationships*. New York: Rodale Press (reprinted as a Ballentine paperback).
- Grice, H. P. (1957). Meaning. *Philosophical Review*, *66*, 377-388.
- Grice, H. P. (1969). Utterer's meaning and intentions. *Philosophical Review*, *78*, 147-177.
- Hinkle, S., Stiles, W. B., & Taylor, L. A. (1988). Verbal processes in a labour/management negotiation. *Journal of Language and Social Psychology*, *7*, 123-136.
- McDaniel, S. H., Stiles, W. B., & McGaughey, K. J. (1981). Correlations of male college students' verbal response mode use in psychotherapy with measures of psychological disturbance and psychotherapy outcome. *Journal of Consulting and Clinical Psychology*, *49*, 571-582.
- Rak, D. S., & McMullen, L. M. (1987). Sex-role stereotyping in television commercials: A verbal response mode and content analysis. *Canadian Journal of Behavioral Science*, *19*, 25-39.
- Rogers, C. R. (1951). *Client-centered therapy*. Boston, MA: Houghton-Mifflin.
- Searle, J. R. (1969). *Speech acts: An essay in philosophy of language*. Cambridge, England: Cambridge University Press.
- Sacks, H., Schegloff, E. A., & Jefferson, G. A. (1974). A simplest systematics for the organization of turn-taking in dyadic conversation. *Language*, *50*, 697-735.
- Stiles, W. B. (1975). *Listener response modes rating manual*. Unpublished manuscript. University of North Carolina, Chapel Hill, NC.
- Stiles, W. B. (1978). *Manual for a taxonomy of verbal response modes*. Chapel Hill, NC: Institute for Research in Social Science, University of North Carolina at Chapel Hill.
- Stiles, W. B. (1981). Classification of intersubjective illocutionary acts. *Language in Society*, *10*, 227-249.
- Stiles, W. B. (1986b). Levels of intended meaning of utterances. *British Journal of Clinical Psychology*, *25*, 213-222.
- Stiles, W. B. (1987b). Some intentions are observable. *Journal of Counseling Psychology*, *34*, 236-239.
- Stiles, W. B., Au, M. L., Martello, M. A., & Perlmutter, J. A. (1983). American campaign oratory: Verbal response mode use by candidates in the 1980 American Presidential primaries. *Social Behavior and Personality*, *11*, 39-43.
- Stiles, W. B., Shapiro, D. A., & Firth-Cozens, J. A. (1988). Verbal response mode use in contrasting psychotherapies: A within-subjects comparison. *Journal of Consulting and Clinical Psychology*, *56*, 727-733.